

Dynamic atlas of histone modifications and gene regulatory networks in endosperm of bread wheat

Received: 20 January 2023

Accepted: 7 October 2024

Published online: 06 November 2024

 Check for updates

Chao He^{1,2}, Siteng Bi^{1,2}, Yuqi Li^{1,2}, Chengxiang Song¹, Heping Zhang¹, Xintong Xu¹, Qiang Li¹, Sulaiman Saeed¹, Wei Chen¹, Chunjie Zhao¹, Caixia Lan¹, Handong Su¹, Hailiang Mao¹✉ & Wenhao Yan¹✉

Dissecting the genetic basis of seed traits in wheat is impeded by limited genetic polymorphisms and significant variations caused by environmental conditions and seed position in a spikelet. Seed performance is largely determined by endosperm development controlled by spatiotemporal variation in gene activities, which is greatly affected by chromatin status. Here, we map genome-wide dynamic distributions of H3K27me3, H3K4me3 and H3K9ac modifications and profile gene transcription across wheat endosperm development. The combinatorial effects of active and repressive marks ensure spatiotemporal dynamic gene expression, especially for starch biosynthesis. By scanning the transcription factor binding motifs in the ATAC-seq peaks, hub regulators are identified from the regulatory network. In addition, we observe significant correlations between sequence polymorphisms of hub regulators and variations in seed traits in a germplasm population. Thus, the analysis of genomic regulatory activities together with genetic variation provides a robust approach to dissect seed traits in bread wheat.

Hexaploid wheat (*Triticum aestivum* L., AABBDD) is a world-wide staple food crop. Seed traits greatly determine the yield and processing quality of wheat. However, performance of seed in wheat is largely affected by environmental conditions like temperature, which changes grain filling rate¹. Even within one spike, different position of spikelet and flowers in a spikelet which usually contains more than three flowers produces different sizes of seeds². In addition, the composition of three sub-genomes results in low frequency of recombination and genetic polymorphism in wheat³. The above facts bring difficulties to dissect genetic basis of seed traits in wheat.

Performance of seed depends on the development of endosperm⁴. The development of endosperm involves several important events. 0–4 day-post-anthesis (DPA) is defined as endosperm coenocytes, which undergoes rapid proliferation of nuclei without cytokinesis⁵. Subsequently, cellularization of the coenocyte (5–7 DPA) starts and all free nuclei are surrounded by newly formed cell wall.

Repeated rounds of mitosis produce cells that eventually occupy the central vacuole completely^{6,7}. From 7 DPA on, the endosperm begins to differentiate into specialized cell types, such as aleurone, starchy endosperm and so on. This period lasts for one week (7–14 DPA) and it is featured by the initiation of starch and protein accumulation. Afterwards, endosperm enters effective filling period, accompanied by grain maturity and drying⁸.

The synthesis and accumulation of starch and seed storage protein (SSP) have been extensively studied in rice and maize^{9,10}, while the knowledge in wheat is relatively limited. In addition to the transcriptomic analysis during development and maturation of wheat endosperm^{8,11,12}, only a few genes such as *TabZIP28*, *TaRSR1*, *TaSPA-B*, *TaNAC019-A1* and *TaNAC019-B1* have been reported to regulate starch and SSP synthesis and accumulation^{13–17}. *TabZIP28* and the homologue *TubZIP28* are highly expressed in the endosperm during the grain filling stage and promote starch synthesis by upregulating *AGPase* in

¹National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan 430070, China. ²These authors contributed equally: Chao He, Siteng Bi, Yuqi Li. ✉ e-mail: maohailiang@mail.hzau.edu.cn; yanwenhao@mail.hzau.edu.cn

wheat¹³; the AP2/EREBP family transcription factor TaRSR1 negatively regulates multiple genes encoding starch synthesis-related enzyme¹⁴; TaSPA-B¹⁵ and seed-specific TaNAC019-A1 and TaNAC019-B1^{16,17} can control the expression of starch- and SSP-related genes. Although such fragmental information is there, it is of great significance to systematically identify genes that are crucial for wheat endosperm development for a better overview of seed traits formation in wheat.

Histone modifications are well-known to be involved in the regulation of gene transcription¹⁸. For example, trimethylation on the fourth lysine and acetylation on the ninth lysine of histone 3 (H3K4me3 and H3K9ac) are both associated with active gene transcription, whereas H3K27me3 is enriched in repressed genes^{19–21}. H3K4me3 is usually present in promoters of actively transcribed genes, while H3K9ac is present not only in promoters but also in distal regulatory elements such as enhancers²². Moreover, DNase I hypersensitivity sites are largely concurrent with peaks of H3K4me3 and H3K9ac in rice²³. Thus, both H3K4me3 and H3K9ac peak regions can be treated as *cis*-element enriched regions for predicting upstream regulators. Evidence in rice and maize has depicted the role of histone modifications in regulating the transcriptional activity of endosperm-specific genes^{24–28}. For instance, a large number of genes involved in nutrient metabolic pathway are directly regulated in the endosperm by OsFIE2-catalyzed H3K27me3. Up-regulation of cell cycle-related genes in maize opaque mutant *opaque18* correlates with increased levels of H3K4me3²⁸. Moreover, a majority number of paternally expressed imprinted genes and transposable elements are marked by H3K27me3 in maize and *Arabidopsis* endosperm^{29–31}. However, the dynamic changes of histone modifications and how it affects gene transcription during wheat endosperm development remain largely unknown.

Here, we perform expression analysis and describe dynamics of histone modifications (H3K4me3, H3K27me3 and H3K9ac) in wheat endosperm at four developmental time-points (4, 7, 14 and 18 DPA). By integrating the information of gene transcription, TF footprint, chromatin modification and genotype/phenotype variation of candidate genes and seed traits in a germplasm population, we construct gene regulatory networks to control endosperm development and uncover hub-genes that affect endosperm development, thus lead to changes in seed traits.

Results

Dynamic gene transcription in developing wheat endosperm

To reveal the dynamic change of gene transcription during endosperm development, we performed RNA-seq using developing endosperm from four developmental stages, i.e., 4 DPA, 7 DPA, 14 DPA and 18 DPA after removing the embryo manually before extracting RNA from endosperm (Fig. 1a). The high correlation (Pearson correlation coefficient *R* value over 0.96 and *P* < 0.05) among replicates at each stage and clear difference between samples from different developmental stages indicate the high quality of the data (Supplementary Fig. 1a and Supplementary Table 1). Principal component analysis (PCA) analysis revealed that the expression pattern of genes at different stages of endosperm development are distinct from that in the embryo^{32,33} (Fig. 1b). We investigated the expression of genes orthologous to those endosperm-specific genes, including *OsNY-YBI*³⁴, *ZmICE1*³⁵, *AtAAPI*³⁶ and *VIMS*³⁷ and found that they are specifically expressed in wheat endosperm (Fig. 1c). In addition, we assessed the expression levels of orthologous to embryo-specific genes in *Arabidopsis*, rice and maize. We found that the expression of the wheat orthologous of maize genes is generally high in embryo³⁸ (Supplementary Fig. 1b). Specifically, the wheat orthologous of *AtS2* and *Zm3896* show high expression levels in wheat embryos and are barely detectable in the endosperm (Fig. 1c).

About 56.67% (61,140/107,891) of the genes were expressed and about 40.75% (43,967/107,891) were constantly expressed at all the four different developmental stages (Fig. 1d, e). These commonly expressed genes were enriched in basic cell metabolic processes like intracellular

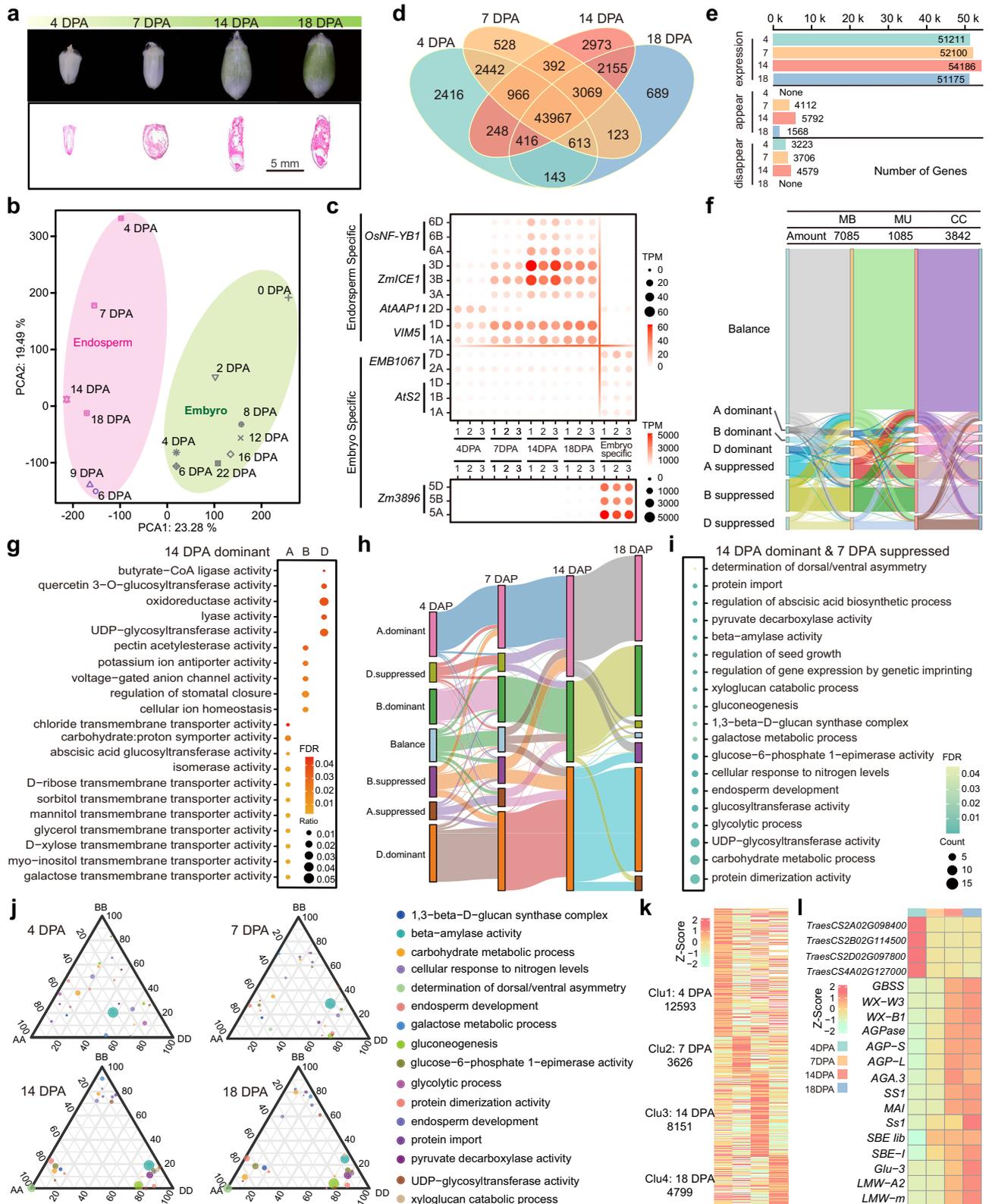
protein transport and protein folding (Supplementary Fig. 1c). The largest number of expressed genes were detected at 14 DPA (Fig. 1e), in which the genes that regulate response to environmental change as well as seed and fruit development were enriched (Supplementary Fig. 1d). This is consistent with the previous finding that 14 DPA is featured by rapid endosperm filling and grain expansion¹².

We found almost equal contribution of the three sub-genomes to the number of genes expressed at each developmental stage (Supplementary Fig. 1e), which is similar to the previous report¹¹. Interestingly, 41.02% of the triads showed asymmetric expression and most of them (31.98%) changed their asymmetry patterns dynamically along endosperm development (Fig. 1f). 7085 triads that maintained balanced expression were mainly involved in fundamental biological processes, such as DNA replication, Golgi vesicle transport and protein folding. 1085 triads that were expressed in an unbalanced manner but had a constant asymmetric pattern among the stages were mainly involved in glycosyl compound metabolic and sulfur compound catabolic pathways, while 3842 triads with changing sub-genome preference across the four stages were mainly involved in processes like DNA repair, cell cycle and stress response (Supplementary Fig. 1f).

The 14 DPA is characterized by the initiation of starch synthesis and has the largest number of expressed genes and stage preferentially expressed genes, which inspired us to further dissect the gene expression at 14 DPA. We found that the A sub-genome preferentially expressed genes at 14 DPA were transporters, while B and D sub-genome preferentially expressed genes were related to maintenance of internal environmental homeostasis or glycosyltransferases (Fig. 1g). Among all the sub-genome preferentially expressed genes at 14 DPA, 89.9% of the sub-genome dominant genes were already expressed at 4 DPA and their sub-genome preference tended to change dynamically. Particularly, 30.4% and 15.8% of sub-genome dominant genes were shown to be sub-genome suppressed before (7 DPA) and after (18 DPA) 14 DPA, respectively (Fig. 1h). GO analysis of sub-genome dominantly expressed genes among triads at 14 DPA suggested that nutrient accumulation related processes, including protein import, gluconeogenesis, galactose metabolic process, glycolytic process, glycosyltransferase activity, beta-amylase activity, glucose-6-phosphate 1-epimerase activity were crucial at 14 DPA (Fig. 1i). Genes that control dorsal/ventral asymmetry showed increased A dominant pattern while the expression of beta-amylase activity and gluconeogenesis related genes were more and more D dominant along endosperm development (Fig. 1j).

Time series analysis revealed that the expression of a total of 53,586 genes changed significantly (*P* < 0.05) during endosperm development. 43.17%, 12.43%, 27.94% and 16.45% genes were preferentially expressed at 4 DPA (Cluster 1), 7 DPA (Cluster 2), 14 DPA (Cluster 3) and 18 DPA (Cluster 4), respectively (Fig. 1k). GO and KEGG analysis showed that the cell cycle related chromosome organization and translation regulator were enriched at the early endosperm development stages (4 DPA and 7 DPA), while starch and sucrose metabolism related genes were enriched at the later stages (14 DPA and 18 DPA) (Supplementary Fig. 2a, b). Genes encoding the α -L-Arabinofuranosidases (*TraesCS2A02G098400*, *TraesCS2B02G114500*, *TraesCS2D02G097800* and *TraesCS4A02G127000*), which are involved in hydrolysis of the cell wall components, were preferentially expressed at the endosperm coenocytic stage (4 DPA) and were suppressed later (Fig. 1l and Supplementary Fig. 2a). This coincides with the observation that there is mainly cell proliferation instead of cell wall formation at endosperm coenocytic stage¹¹.

We found that many sucrose metabolism related genes were highly expressed at 14 DPA and 18 DPA (Fig. 1l and Supplementary Fig. 2a), such as *granule-bound starch synthase I* (*GBSS*, *WX-W3* and *WX-B1*), *sucrose synthase I* (*SsI*), *glycosyltransferase* (*SS1*) and *alpha-amylase inhibitor* (*MAI*). *1,4-alpha-glucan-branching enzyme* (*SBE Iib* and *SBE-I*) started to express at 7 DPA and reached its highest level at 18



DPA. Interestingly, most of the genes encoding glucose-1-phosphate adenylyl transferase including *AGPase*, *AGP-S* and *AGP-L* were highly expressed at both 14 DPA and 18 DPA, while *AGA.3* was only highly expressed at 14 DPA but decreased at 18 DPA. The expression of genes encoding low molecular weight glutenin subunit (*Glu-3*, *LMW-A2* and *LMW-m*) started to increase at 14 DPA and got higher at 18 DPA (Fig. 1l). The stage-dependent expression patterns of starch and glutenin

synthesis-related genes indicate a spatiotemporal transcription regulation behind.

Dynamic gene transcription during wheat endosperm development is correlated with changed level of histone modifications
 Histone modifications have been reported to be highly correlated with gene transcription and thus specify spatiotemporal gene activities to

Fig. 1 | Dynamic changes in gene expression at different endosperm development stages of wheat. **a** The anatomical dissection pictures at different sampling stages. DPA indicates day-post-anthesis. **b** PCA analysis of endosperm and embryo RNA-seq data. Endosperm RNA-seq data of 6 DPA and 9 DPA in the light pink oval were downloaded from the Sequence Read Archive (SRA) under the accession number of SRP013449³³. Embryo RNA-seq data of 0 DPA, 2 DPA, 4 DPA, 6 DPA, 8 DPA, 12 DPA, 16 DPA and 22 DPA in the ash green oval were downloaded from SRA under the accession number of PRJCA008382³². **c** The expression patterns of endosperm-specific and embryo-specific genes in the sampling endosperm and publicly available embryo datasets (PRJEB25639)³⁸. **d, e** Expressed genes (TPM > 0.5) in different endosperm developmental stages. **f** The dynamic changes of homoeology bias expressed genes in the three sub-genomes homoeologous triads of wheat. MB, MU and CC indicate maintain balanced, maintain unbalanced and constantly changing, respectively. **g** The GO enrichment results of 14 DPA dominant genes. The *P* values

were calculated using hypergeometric distribution and then adjusted using the FDR. **h** The expression pattern of 14 DPA dominant homoeology triads at 4 DPA, 7 DPA, 14 DPA and 18 DPA. **i** GO enrichment results of 14 DPA dominant while 7 DPA suppressed homoeology gene triads. The *P* values were calculated using hypergeometric distribution and then adjusted using the FDR. **j** Ternary plot showing relative expression abundance of the 14 DPA dominant while 7 DPA suppressed homoeology gene triads at 4DPA, 7DPA, 14DPA and 18DPA. Each circle represents a gene triad with an A, B, and D coordinate consisting of the relative contribution of each homoeologs to the overall triad expression. Triads in vertices correspond to single-subgenome-dominant categories, whereas triads close to edges and between vertices correspond to suppressed categories. Circle size represents the maximum expression level in the A, B and D homoeolog triad. **k** The expression patterns of the stage preferentially expressed genes. **l** The expression patterns of α -L-Arabinofuranosidases, starch metabolism and SSP associated genes. Source data are provided as a Source Data file.

control plant development³⁹. H3K27me3 is a marker for gene repression while H3K4me3 and H3K9ac both mark actively transcribed genes. In addition, H3K9ac overlaps with distal regulatory elements^{22,23}. We mapped the genome-wide distribution of H3K27me3, H3K4me3 and H3K9ac modifications at the above four endosperm developmental stages (Supplementary Data 1). The high signal-to-noise ratios (the average Signal Portion of Tags score over 0.5) indicated the quality of the data and the replicates were highly correlated (Pearson correlation coefficient *R* value over 0.8, Supplementary Fig. 3a–c). A large number of histone modification regions were located in the intergenic regions (Supplementary Fig. 3d). Overall, H3K27me3 was enriched in lowly expressed genes and high level of H3K4me3 and H3K9ac were presented in highly expressed genes. H3K27me3 was obviously antagonistic to H3K4me3 and H3K9ac in association gene expression level (Fig. 2a and Supplementary Fig. 4a). For example, in the three sucrose synthesis genes mentioned above, the levels of H3K27me3 modification were found to be gradually decreased, while those of H3K4me3 or H3K9ac increased from 4 DPA to 18 DPA (Fig. 2b).

The level of histone modifications (only gene body and its surrounding regions were considered) varied over time (Fig. 2c). H3K4me3 was lower at 4 DPA and 7 DPA, but was obviously elevated at 14 DPA. A similar trend was observed for H3K9ac. On the contrary, the highest level of H3K27me3 was observed at 4 DPA and slightly decreased at 7 DPA. Then, it sharply dropped at 14 DPA and maintained low level (Fig. 2c). Interestingly, higher expression of early stage (4 DPA and 7 DPA) specifically expressed genes was associated with not only higher level of H3K4me3 and H3K9ac but also high level of H3K27me3. Furthermore, at the later stage, those stage specifically expressed genes possessed either high (14 DPA) or moderate (18 DPA) level of active modifications but always low level of H3K27me3 (Fig. 2d). All these findings indicated that active and repressive histone marks cooperate to maintain temporal gene transcription and more dramatic fold change of H3K27me3 intensity compared with relative fold change of H3K4me3 and H3K9ac between early (4 and 7 DPA) and later stages (14 and 18 DPA) may indicate the pivotal role of repressive chromatin state for endosperm development.

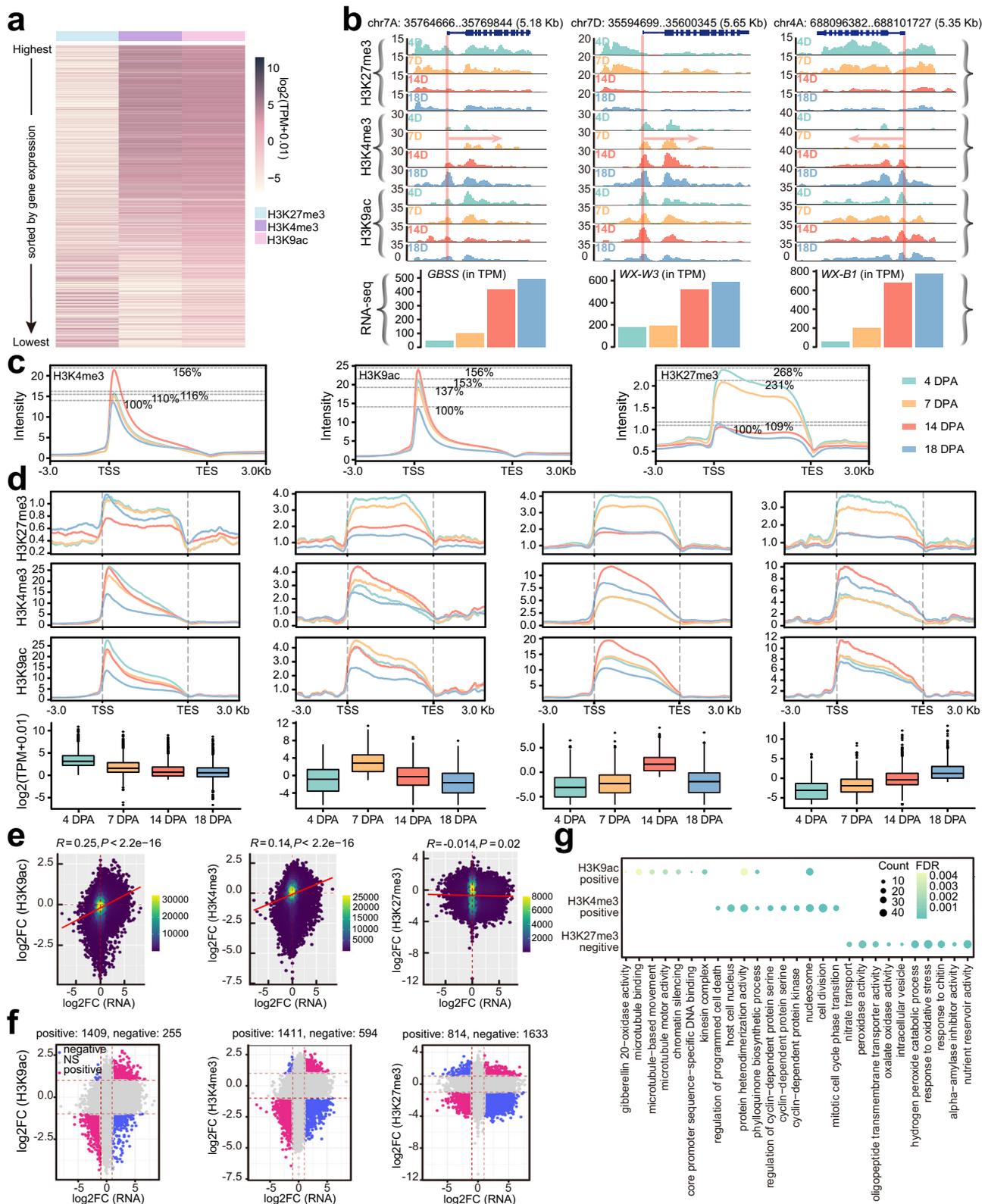
We again focused on 14 DPA to further probe the regulatory role of histone modifications in sharpening gene transcription. As expected, changes in gene expression between 7 DPA and 14 DPA were significantly (Pearson correlation coefficient *R* value over 0.14, $P < 2.2e-16$) positively correlated with the changes in H3K4me3 and H3K9ac, while negatively correlated with the fold change in H3K27me3 (Fig. 2e). Those genes with expression fold change > 2 (*q* value < 0.05) were regarded as differentially expressed genes (DEGs) and DEGs with histone modification fold changes > 2 were regarded as histone modified DEGs (hDEGs) (Supplementary Data 2). Among all the hDEGs between 7 DPA and 14 DPA, 84.7% and 70.4% of them showed positive correlation with changed H3K4me3 and H3K9ac levels, respectively and more than 66.7% of them showed negative correlation with

changed H3K27me3 (Fig. 2f). We found that DEGs between 7 DPA and 14 DPA marked by different histone modification types showed distinct functions in different biological process. DEGs marked by H3K4me3 were mainly involved in mitotic cell cycle phase transition, and DEGs with H3K9ac were mainly involved in microtubule-based movement, core promoter sequence-specific DNA binding and chromatin silencing while DEGs with H3K27me3 modification were mainly involved in nutrient reservoir activity, alpha-amylase inhibitor activity and response to oxidative stress (Fig. 2g). Of course, there are DEGs marked by combinatorial effects between active and repressive histone marks (Supplementary Data 3). For instance, the expression of two Alpha-gliadin *Gli-2* (*TraesCS6A02G049400* and *TraesCS6A02G049800*) were sharply increased at 14 DPA and this change was accompanied by an increase in H3K4me3 and H3K9ac but a decrease in H3K27me3 (Supplementary Fig. 4b).

To see how histone modifications correlate with the asymmetric gene expression of homoeologous triads, we mapped the histone modifications in the seven clusters including six unbalanced groups of triads and a balanced one at different developmental stages of endosperm (Supplementary Fig. 5). In general, dominantly expressed homoeologs were coupled with higher H3K4me3 or H3K9ac and lower level of H3K27me3 and for the suppressed genes, lower level of active modifications but high level of H3K27me3 could be detected. The histone modifications in the genes with balanced expression mode tended to be of equal intensity among the three homoeologs. One interesting phenomenon is that when one sub-genome member is preferentially expressed, the histone modifications of the remaining two tend to be of the same intensity. In addition, we found that histone modifications in the A and D sub-genomes were strongly antagonistic in maintaining the asymmetric expression pattern, especially at the early stages (4 DPA and 7 DPA) (Supplementary Fig. 5).

Expression of starch and protein synthesis related genes is closely related to histone modifications

Starch biosynthesis consists of two consecutive processes. First, photosynthetic products and sucrose are hydrolyzed into glucose 1 phosphate (G1P) and then G1P is catalyzed by a series of enzymes to produce amylose and amylopectin⁴⁰ (Fig. 3a and Supplementary Data 4). Genes that function to generate G1P were mostly expressed at the early stages (4 DPA and 7 DPA), while amylose and amylopectin synthesis genes to utilize G1P were preferentially expressed at the later stages (14 DPA and 18 DPA) (Fig. 3a, b and Supplementary Data 4). Dynamic changes of H3K27me3, H3K4me3 and H3K9ac were closely related to the altered expression of starch synthesis related genes along the endosperm development (Fig. 3a), e.g., *BTI* was hardly expressed at early stages but started to be expressed at later stages. Meanwhile, compared with early stages, H3K27me3 level on *BTI* was decreased while H3K4me3 and H3K9ac were increased at later stages. The expression of the *Waxy*, *SBE* and *DBE* was elevated at later stages and these genes were hardly marked by H3K27me3 but H3K4me3 and



H3K9ac were highly enriched on these genes at 14 DPA and 18 DPA (Fig. 3a).

Glutenins and gliadins are two major types of SSPs in bread wheat and 86 genes were reported to be involved in the synthesis of glutenins and gliadins¹⁷. We found that 77.91% of them were expressed (TPM > 0.5) at least at one of the four stages and all these genes were preferentially expressed at 14 DPA and 18 DPA (Fig. 3b, c, Supplementary

Fig. 6 and Supplementary Data 5). Most SSPs related genes were highly enriched with H3K4me3 and H3K9ac but with decreased H3K27me3 at 14 DPA and 18 DPA (Fig. 3c). Specifically, the expression of the high molecular weight glutenins (HMW-GSs) encoded genes were solely activated by H3K4me3 at 14 DPA and 18 DPA, while the genes for the low molecular weight glutenins (LMW-GSs) were regulated by H3K27me3 in concert with H3K4me3 and H3K9ac (Fig. 3c). Regarding

Fig. 2 | Dynamic changes of histone modifications during the endosperm development of wheat. **a** The correlations between the intensity of histone modifications and gene expression at 14 DPA. **b** The expression level (in TPM) and enriched signal of H3K27me3, H3K4me3 and H3K9ac on the three *Waxy* genes. **c** The intensity variation of H3K27me3, H3K4me3 and H3K9ac on upstream, gene body and distal regions of genomic genes across different endosperm development stages. **d** The intensity changes of H3K27me3, H3K4me3 and H3K9ac on stage specifically expressed genes. Each sample consists of 12593, 3626, 8151, and 4799 observations. The box plots display the

median (horizontal line), second to third quartiles (box), with the maxima and minima as whiskers, and outliers as points beyond 1.5 times the interquartile range. **e** The correlation between variation fold changes in gene expression and histone modification levels of genes at 7 DPA and 14 DPA. Number (**f**) and GO enrichment (**g**) of differentially expressed genes with similar (positive) or opposite (negative) variation pattern in expression level and modification level of H3K4me3, H3K9ac and H3K27me3, respectively. The *P* values were calculated using hypergeometric distribution and then adjusted using the FDR. Source data are provided as a Source Data file.

gliadin, the activation of most of genes encoding alpha alcoholic proteins at 14 DPA was affected by the elevation of H3K4me3 with reduced H3K27me3. H3K9ac in some cases stayed with H3K4me3 to mark the highly expressed alpha alcoholic proteins coding genes. Gamma-gliadin gene expression was associated with H3K4me3 and H3K9ac but the omega-gliadin genes were not (Fig. 3c). Overall, H3K4me3 was proven to be the more general modification that marks the active expression of SSP coding genes compared with H3K9ac.

We further looked into the expression asymmetry of starch synthesis related and SSP coded homoeologous gene triads across endosperm development and analyzed the histone modifications in each sub-genome homoeolog. A total of 43 homoeologous triads were found for starch synthesis but none for SSP coding genes (Fig. 3d). Most of these 43 starch synthesis-associated gene triads showed symmetry expression and higher expression of the sub-genome member was correlated with more enriched H3K4me3 and H3K9ac (Fig. 3d). H3K27me3 were found to be absent in the regulation of starch synthesis-related homoeologous gene triads (Fig. 3d).

Gene regulatory networks controlling endosperm development identified ERF5 as a hub regulator

Transcription factors act as upstream master or hub regulators to control downstream target gene activities mediating organ differentiation⁴¹. TF footprint analysis in open chromatin regions (OCRs) enables a comprehensive genome-wide prediction of TF binding. We conducted ATAC-seq to identify OCRs and an average of 85,850 high signal-to-noise OCRs across the four stages of endosperm development were detected (Supplementary Table 2 and Supplementary Fig. 7). By integrating gene co-expression data with *cis*-motifs and TF footprints identified in the OCRs, we constructed a TF regulatory network controlling wheat endosperm development (Fig. 4a). In this network, transcription factors were on the top of the hierarchy, serving as nodes. A total of 89 TFs were involved in the regulation of wheat endosperm development (Supplementary Data 6). The enriched GO of the TF target genes were shown in the network (Fig. 4b). The DAP-seq data for six hub TFs, including *SPL5* (*TraesCS3A02G432500*, *TraesCS3B02G468400* and *TraesCS3D02G425800*)⁴², *EFM* (*TraesCS5B02G075300*)⁴³, *SPL14* (*TraesCS7B02G142200*)⁴² and *ERF5* (*TraesCS6D02G225700*)⁴³, show that an average of 70.9% of regulatory interactions can be detected in vitro (Supplementary Fig. 8 and Supplementary Data 7), further validating the reliability of our network.

We identified *TaFUSCA3/TaABI3* (*TraesCS2A02G554300*), a B3 superfamily TF that had been shown to activate the expression of the high molecular weight glutenin subunit gene *Glu-1Bx7*⁴⁴, as a hub regulator at 14 DPA (Fig. 4a). Other TFs including *TaNAC019* (*TraesCS3A02G077900*, *TraesCS3B02G092800* and *TraesCS3D02G078500*)¹⁷, *PBF* (*TraesCSSB02G154100*)⁴⁵, *TabZIP28* (*TraesCS2B02G167900*)¹³, GA-dependent MYB TF *TaGAMyb* (*TraesCS6D02G173000*, *TraesCS5A02G159600*, *TraesCS5B02G157300*, *TraesCS1D02G283100*, *TraesCS4D02G176500* and *TraesCS1A02G219400*)⁴⁶, which had been reported to be involved in the accumulation of starch and protein during wheat endosperm development, were also present in the constructed regulatory network. All these results indicate the biological relevance of the regulatory network that we constructed.

Most of the expression of TFs identified at the early stages of endosperm development (4 DPA and 7 DPA) were marked by H3K4me3 (Supplementary Data 6). These TFs were involved in the process of cell differentiation, mitotic cell cycle and response to glucose (Fig. 4b). Hub TFs in the endosperm of 14 DPA were found to regulate the transcription of genes involved in seed maturation, regulation of cell shape, response to abscisic acid, circadian rhythm, vesicle-mediated transport and ethylene-activated signaling pathway (Fig. 4b). The hub TFs identified in the endosperm of 18 DPA were mainly marked by H3K4me3 or H3K9ac (Supplementary Data 6), and their target genes were mainly involved in the process of programmed cell death, nutrient reservoir activity, starch biosynthetic process, late endosome to vacuole transport, protein maturation and ubiquitin protein transferase activity in addition to the similar processes identified at 14 DPA (Fig. 4a, b).

We then analyzed the TFs that directly regulated the expression of genes related to starch synthesis and/or SSPs. We identified that an ethylene response factor *ERF5* (*AP2/ERF* transcription factor family, *TraesCS6D02G225700*), which was proposed to be involved in the ethylene-activated signaling pathway^{47,48}, showed largest number of regulated genes at 14 DPA (Fig. 5a), and may regulate not only SSPs-coding genes (*Alpha-gliadin* and *Omega-gliadin*) but also genes related to starch synthesis (*FBA* and *DBE*) (Fig. 5b). More importantly, *ERF5* functions as an upstream regulator of many hub TFs including *WRKY55* and *HSFA6B*, both of which are key hub genes in the regulation of SSP encoding genes (Fig. 5a–c). The expression of *ERF5* was highly correlated with the level of H3K9ac and H3K4me3 (Fig. 5d). To verify the regulatory role of *ERF5* at 14 DPA and 18 DPA, we analyzed *ERF5* binding using DAP-seq⁴³. The results revealed that 1419 predicted target genes in the network could be also bound by *ERF5* in vitro revealed by DAP-seq (Supplementary Fig. 9a and Supplementary Data 7). GO enrichment analysis of these target genes indicated that they are primarily enriched in the starch biosynthetic process, ethylene-activated signaling pathway, 1,4-alpha-glucan branching enzyme activity, and racemase and epimerase activity (Supplementary Fig. 9b). Some genes previously reported to be involved in starch synthesis and SSP coding were bound by *ERF5* revealed by the network and the DAP-seq data⁴³ (Supplementary Fig. 9a, b and Supplementary Data 4 and 5).

To further confirm the function of *ERF5*, we performed haplotype analysis and found significant phenotypic differences in grain length (GL), grain width (GW) and thousand-grain weight (TGW) between different haplotypes (Fig. 5e). In line with the pivotal role of *ERF5* in the gene regulatory network during grain filling, significantly ($P < 0.05$, two-tailed Student's *t* test) narrower grains were produced when *ERF5* was overexpressed although the knockout mutant produced seeds similar to wild type (Fig. 5f–h and Supplementary Fig. 10), possibly due to gene redundancy.

Natural variation of genes in the regulatory network contribute to grain traits in wheat

As endosperm is the main part of wheat grain, the development of endosperm largely affects the grain traits. To further estimate the genetic effect of the identified hub TFs and their target genes in the gene regulatory network on seed traits, we performed GWAS analysis

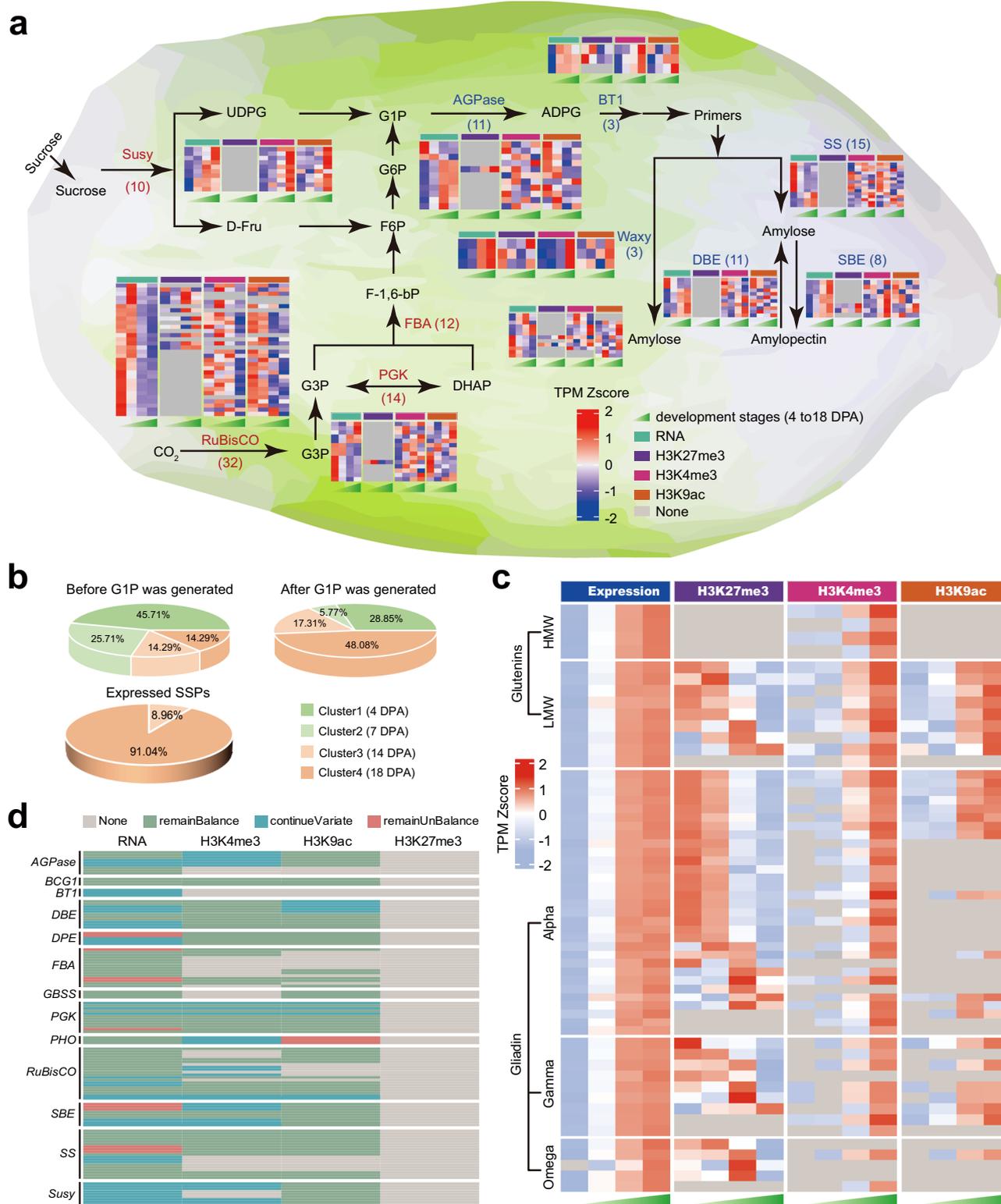


Fig. 3 | Dynamic changes of histone modifications and gene expression in starch and seed storage protein accumulation processes. **a** The process of starch synthesis and the changing gene expression and histone modifications at different developmental stages were presented. **b** Percentage of different stages preferentially expressed genes before and after G1P production and in expressed seed storage protein (SSP) encoded genes. Cluster 1–4 represent 4 DPA, 7 DPA, 14

DPA and 18 DPA specifically expressed genes in Fig. 1k, respectively. **c** The expression level and histone modification intensity of SSP encoded genes at different stages. **d** The varied gene expression and histone modification intensity of nutrient synthesis related homoeologous gene triads during different stages. Source data are provided as a Source Data file.

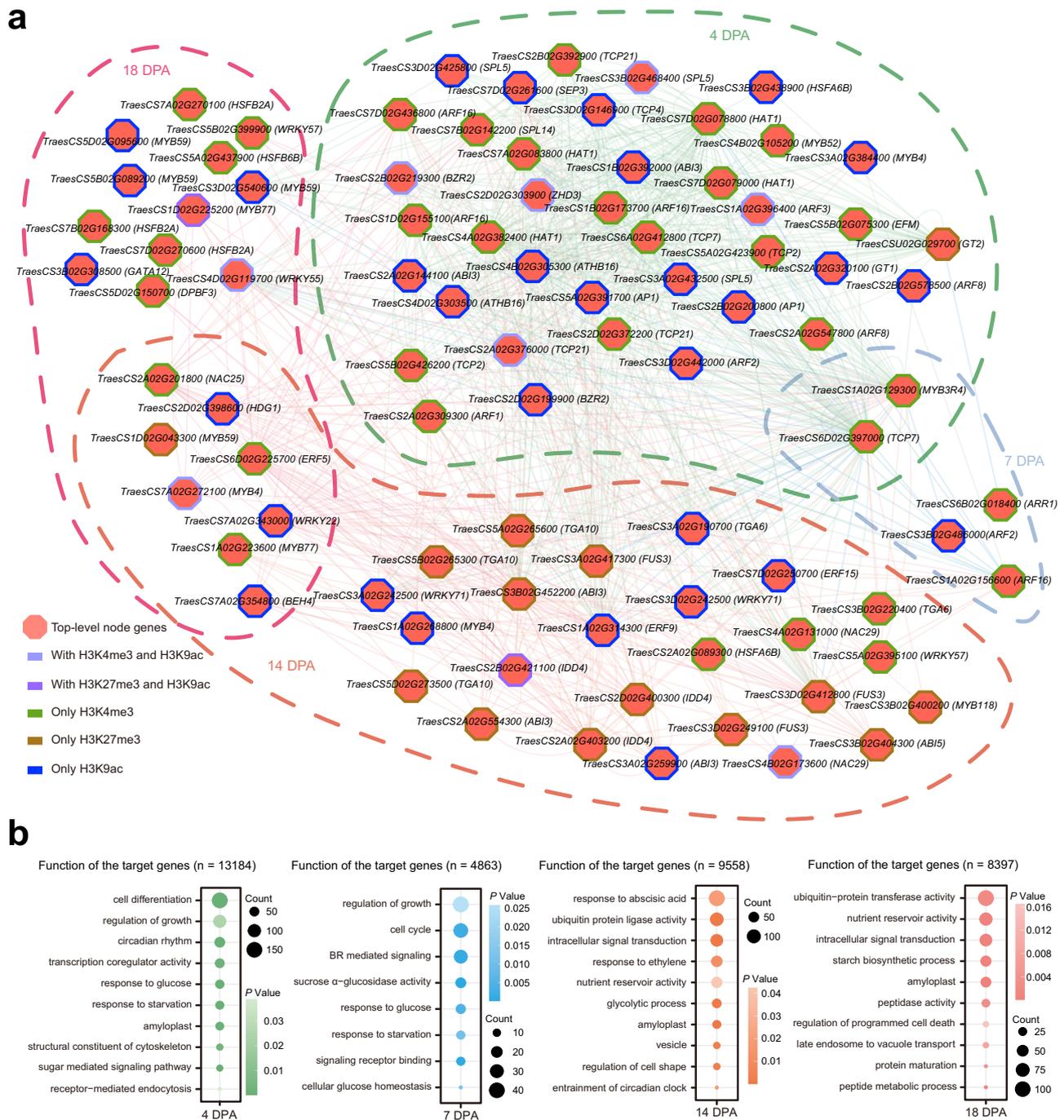


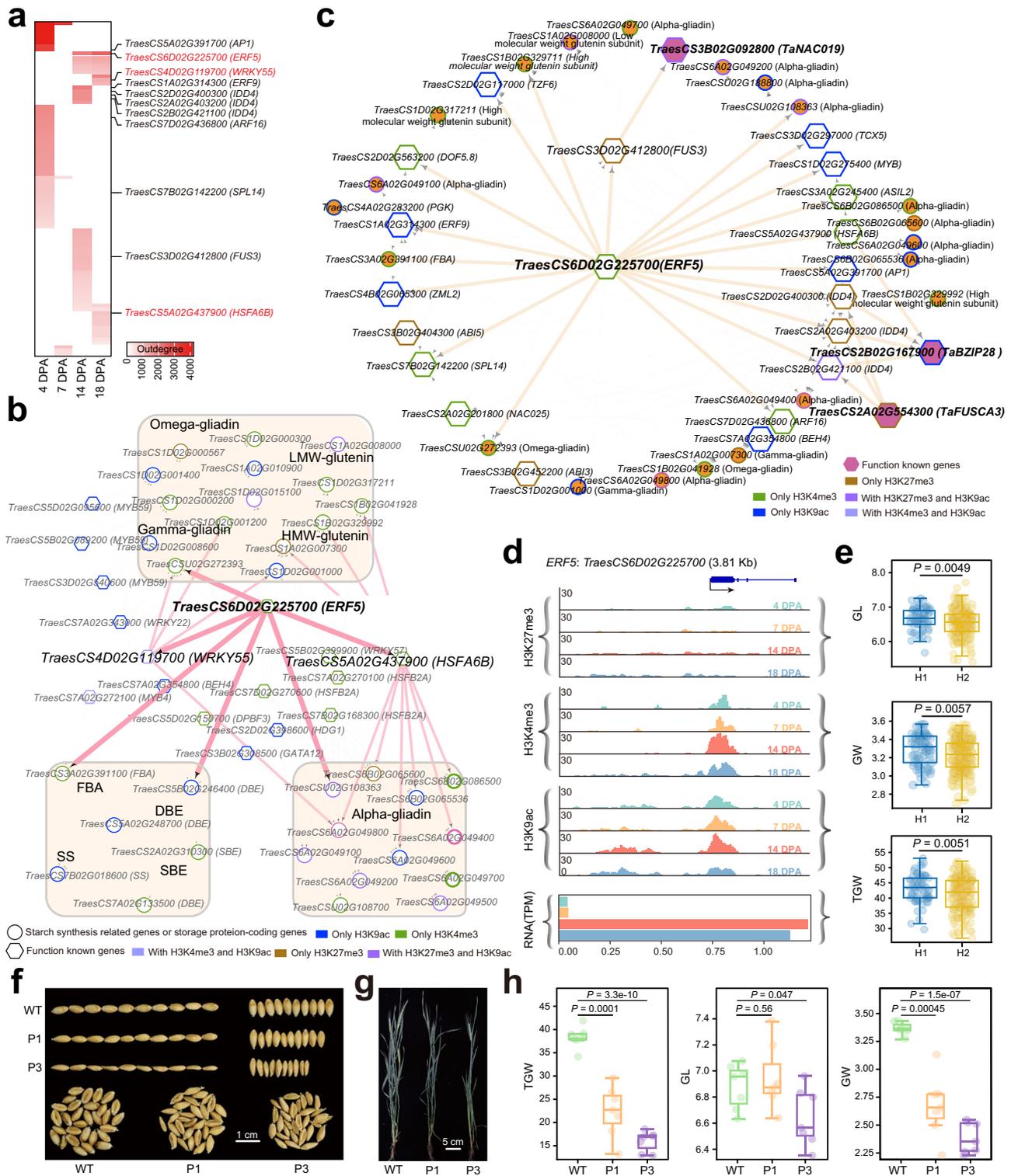
Fig. 4 | Gene regulatory network controlling dynamic endosperm development. **a** Stage-specific gene regulatory network. The hub regulators are indicated by red node. The octagon with brown, dark olive green and blue edges indicate that the genes were marked by H3K27me3, H3K4me3 and H3K9ac, respectively. Those

genes with two modifications were in octagons with either purple or light blue edges. **b** GO analysis of target genes in **(a)** at different stages. The *P* values were calculated using hypergeometric distribution and then adjusted using the Benjamini–Hochberg (BH) method. Source data are provided as a Source Data file.

using a germplasm population containing 516 accessions with broad genetic diversity and grain traits collected from different years and locations (Supplementary Data 8). As a result, we identified 1068 leading single nucleotide polymorphisms (SNPs) associated with grain traits, including TGW, GW, GL and ratio of grain length to grain width (GL/GW). These SNPs are loaded in 1068 quantitative trait locus (QTL) regions defined by LD blocks (Supplementary Fig. 11 and Supplementary Data 9). Several functionally known genes, including three previously reported transcription factors regulating starch synthesis, six starch synthesis-related genes, and twelve storage protein-coding

genes were found to be located in the QTL regions (Supplementary Fig. 11 and Supplementary Data 10).

We identified 969 genes in the regulatory network that are located in the QTL regions (Supplementary Data 11). Each of the identified 89 hub TFs in the network regulates at least one of these target genes with *MYB118* (*TraesCS3B02G400200*) and *GATA12* (*TraesCS3B02G308500*) located in the QTL regions (Supplementary Data 12). We then estimated the contribution of variations of the 89 hub TFs and the 1649 target genes, which include the 969 genes located in the QTL regions as well as the predicted targets of *MYB118* and *GATA12* (Supplementary



Data 13) to grain traits. In comparison with the effects of variation from randomly picked genomic regions, the ratio of genes, whose variation significantly ($P < 0.05$, two-tailed Student's t test) affected thousand TGW, GW, GL and GL/GW from both hub TFs and target genes was evidently higher than the average ratio of 100 times randomly selected genomic loci (Fig. 6a).

Genetic polymorphism of 71 hub TFs was found to significantly ($P < 0.05$, two-tailed Student's t test) contribute to variation on GW (Supplementary Data 14), GL (Supplementary Data 15), GL/GW (Supplementary Data 16), or/and TGW (Supplementary Data 17) of

wheat. Networks centered around these genes were constructed along with their 1622 target genes (Fig. 6b). Among these hub TFs, *MYB118* (*TraesCS3B02G400200*) was reported to be a transcription activator and played an important role during seed maturation⁴⁹. Overexpression of its homologous gene in *Arabidopsis* caused a remarkable accumulation of SSPs and other seed maturation proteins⁵⁰. Moreover, seeds of *myb118* mutants exhibited an endosperm-specific depression of maturation-related genes³⁴. *MYB118* was isolated as a key regulator for later endosperm development (14 DPA) and significant ($P < 0.05$, two-tailed Student's t test) differences in

Fig. 5 | *ERF5* is a hub regulator in starch and storage protein accumulation. **a** Number of genes regulated by each of the 89 TF regulators. **b** Regulation between *ERF5* and other TFs that directly regulated the starch synthesis related and storage protein encoded genes. The circles represent starch or SSPs-related genes. The octagon represents TFs. **c** The *ERF5* knitted regulatory network including known TFs (in deep pink), which were reported to regulate the accumulation of starch and storage protein. The octagon with brown, dark olive green and blue edges indicate that the genes were marked by H3K27me3, H3K4me3 and H3K9ac, respectively. Those genes with two modifications were in octagons with either purple or light blue edges. **d** IGV view of histone modifications and mRNA abundance of *ERF5*. The vertical blue line and arrow indicate the transcription site and direction, respectively. **e** Effects of different haplotypes of *ERF5* on grain length (GL), grain

width (GW) and thousand grain weight (TGW). The sample sizes of H1 and H2 are 73 and 213, respectively. The *P* values were calculated by two-tailed Student's *t* test. The box plots display the median (horizontal line), second to third quartiles (box), with the maxima and minima as whiskers, and outliers as points beyond 1.5 times the interquartile range. Grain phenotype (**f**), Plant appearance (**g**) and Statistical analysis (**h**) of seed traits in overexpression lines. WT indicates wild type, P1 and P3 indicate the two overexpression lines. Each sample consists of 7 observations, with each observation being the average value of 10 randomly selected seeds. The value above the line is *P* value calculated using a two-tailed Student's *t* test. The box plots display the median (horizontal line), second to third quartiles (box), with the maxima and minima as whiskers, and outliers as points beyond 1.5 times the interquartile range. Source data are provided as a Source Data file.

GW, GL and TGW were observed between the two different alleles (Fig. 6b, c). Furthermore, the grains with different genotypes of the *NAC029* (*TraesCS4B02G173600*) were found to exhibit significantly ($P < 0.05$, two-tailed Student's *t* test) difference in GW, GL and TGW (Fig. 6c). Interestingly, different genotypes of *MYB4* (*TraesCS7A02G272100*), which negatively regulated genes involved in anthocyanin biosynthesis⁵¹, showed a significant ($P < 0.05$, two-tailed Student's *t* test) difference in GW, GL, GL/GW and also TGW (Fig. 6c and Supplementary Data 13). Contribution of genetic variations in the genes identified by the constructed gene regulatory network on seed trait performance highlights the robustness of the approach.

Discussion

Seed trait is relying on the development of wheat endosperm and it is crucial for the final grain yield and flour quality. We constructed dynamic gene transcription network and identified a number of genes controlling early endosperm development by combining gene transcription and chromatin status.

14 DPA had the largest number of expressed genes with the most dramatic changes of gene expression when compared with other time-points (Fig. 1d, e). Enriched genes in GO analysis supported the role of 14 DPA as transition point from basic endosperm morphogenesis to nutrient accumulation^{6–8} (Supplementary Fig. 1c, d and Fig. 1g, i). 14 DPA was featured by sub-genome asymmetric gene expression pattern for SSPs and genes related to starch accumulation (Fig. 1g–j). These results suggest that 14 DPA is the critical stage for endosperm development.

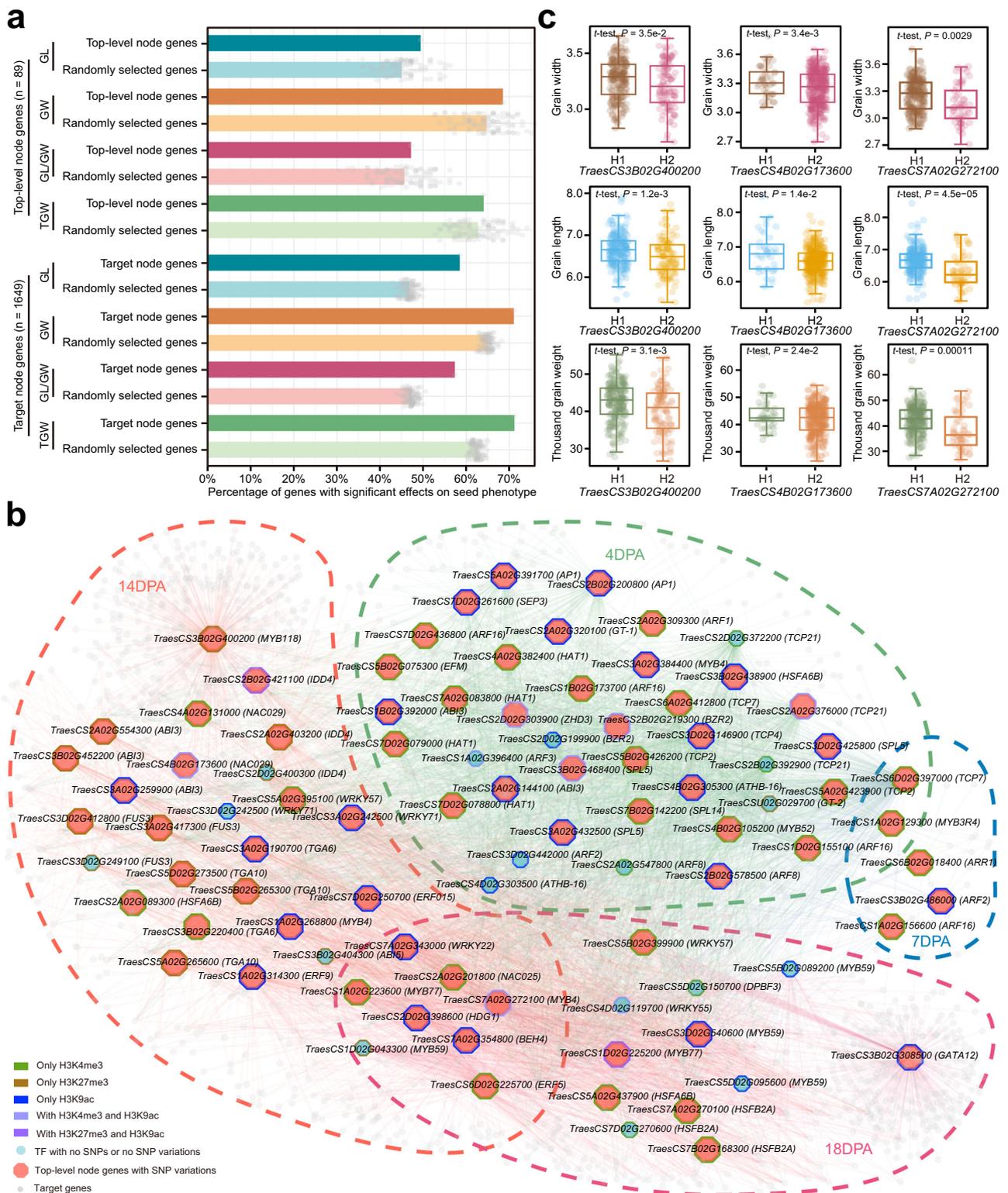
Status of histone modification is highly correlated with gene transcription⁵². Covalent modifications of histones can alter chromatin structure or recruit downstream effectors, leading to changes in transcriptional activity^{18,29–31}. A well-known example is the progressive repression of FLOWERING LOCUS C (FLC) by H3K27me3 during vernalization⁵³. Studies in rice and maize have reported the role of histone modifications in endosperm starch and protein accumulation^{24–28}. We mapped H3K27me3, H3K4me3 and H3K9ac in wheat endosperm to explore the role of histone modifications on gene transcription during endosperm development. We found that various types of histone modifications maintained their known correlation with gene transcription in endosperm (Fig. 2a and Supplementary Fig. 4a). H3K27me3 showed more dramatic change than H3K4me3 and H3K9ac did between early stage (4, 7 DPA) and late stages (14, 18 DPA) (Fig. 2c) and this pattern was also observed in a big number of genes related to starch and protein synthesis and accumulation (Fig. 3a, c). This is consistent with the function of H3K27me3 in mediating reprogramming^{54,55}. The change in H3K27me3 is critical in coordinating the transition of the developmental states of the endosperm. However, single type of modification is not sufficient to determine gene transcription and we found that the synergistic effects of active and repressive modifications controlled the temporal expression of genes (Fig. 2b and Supplementary Fig. 4b) and the asymmetric gene expression pattern (Supplementary Fig. 5), especially the genes related to starch and SSPs synthesis (Fig. 3).

Transcription factors (TFs) are key and central regulators of gene expression to control growth, organ development and response to environmental cues. Regulatory elements are supposed to be bound by upstream regulators, which are always located in OCRs. After scanned TF footprints within the OCRs, we constructed dynamic regulatory networks for early endosperm development. 89 hub TFs and 20,164 target genes were obtained (Fig. 4a). The AFL and VAL subclades of the plant specific B3 domain containing transcription factor family coordinate the spatiotemporal development of seed and they are known as the LAFL TF network⁵⁶. The LAFL mediated transcription network regulates a variety of seed-specific development processes, including starch, SSPs and lipids accumulation⁵⁷. Several LAFL related gene, such as *ABI3*, *ABI5*, *FUS3* and *MYB118* are present in our constructed network (Fig. 4a), and the putative functions of their target genes are corresponding to the characteristics of each developmental stage of wheat endosperm (Fig. 4b).

We identified an AP2/ERF gene, *ERF5* (Ethylene Response Factor, *TraesCS6D02G225700*) as one top-level node gene in the endosperm at both 14 DPA and 18 DPA (Fig. 5a). *ERF5* was shown to regulate both starch synthesis (FBA and DBE) related genes and SSPs (Alpha-gliadin and Omega-gliadin) encoding genes (Fig. 5b). More importantly, it is an upstream regulator of *WRKY55* and *HSFA6B*, which acted as two regulators of *SSP* (*Alpha gliadin*, *Omega-gliadin*, *Gamma-gliadin* and *HMW-GS*) encoding genes (Fig. 5b). Further analysis of the regulator network identified that a large number of hub TF regulators are regulated by *ERF5*. Specifically, three reported endosperm starch and protein synthesis-related transcription factors, *TaNAC019*¹⁷, *TaFUSCA3*⁴⁴ and *TaBZIP28*¹³, were presented in the network drove by *ERF5* (Fig. 5c). Furthermore, most of the target genes in the network were found to be bound by *ERF5* in vitro (Supplementary Fig. 9). Overexpression of *ERF5* resulted in significantly ($P < 0.05$, two-tailed Student's *t* test) narrower grains (Fig. 5f–h). These findings indicate the pivotal role of *ERF5* in the regulatory network during grain filling and its crucial function in regulating starch and protein accumulation.

In addition to *ERF5*, we also identified a large number of hormone-related transcription factors such as ARFs and MYBs in the network (Fig. 4a). This is consistent with the critical role of hormones in starch accumulation during endosperm development, especially the regulation of endosperm filling by auxin as reported by Li et al.¹². Several seed dormancy-related transcription factors, such as *GATA12*, *HSFA6B* and *WRKY57*, are also present in the network. It is believed that these genes mainly function in late seed maturation stage and germination through ABA and GA pathway^{58,59}. Here, they were identified to be important regulators at the 14 and 18 DPA when nutrients accumulate (Fig. 4a). They were also found to significantly ($P < 0.05$, two-tailed Student's *t* test) affect the seed traits (Supplementary Data 13). All these findings suggest that these genes may have pleiotropic roles throughout the whole process from endosperm development to germination.

When we looked into the effects of natural variations of the hub TFs in the networks on seed traits, strikingly but as expected, the variation of quite some node genes showed a significant ($P < 0.05$, two-tailed Student's *t* test) effect on seed traits (Fig. 6a). These results



strongly support the regulatory role of the genes we identified during wheat endosperm development. As expected, the dynamic expression of these regulators is highly correlated with the change of H3K4me3, H3K9ac and H3K27me3 (Supplementary Fig. 12), which further exemplified that orchestration of active and repressive histone modifications is an underlying driver of spatiotemporal gene transcription during endosperm development.

Wheat has narrow genetic diversity⁶⁰ but the trait performance of wheat grain traits varies a lot due to change in both environmental

conditions and position of the florets in a spike, which brings challenges to map genes controlling seed traits using traditional genetic approach⁶¹. In this study, we constructed dynamic gene transcription network and identified candidate genes that controls endosperm development. Recently, Zhao et al. presented a gene regulatory network in wheat⁶². However, the main objectives of the two studies differ, allowing for a complementary understanding of the developmental processes. Our work combines genetic analysis of seed traits with transcriptome and epigenome profiles across key developmental

Fig. 6 | Natural variation of hub genes in regulatory network contribute to grain traits in wheat. **a** Ratio of SNPs from hub TFs, target genes identified in the network and 89 & 1649 randomly selected genes (replicated 100 times) that caused significant ($P < 0.05$, two-tailed Student's t test) variation in wheat grain traits. TGW, GW, GL and GL/GW represent thousand grain weight, grain width, grain length and ratio of grain length to grain width, respectively. **b** Regulatory network of genes with SNPs that were significantly ($P < 0.05$, two-tailed Student's t test) associated with variation in GW, GL, GL/GW or/and TGW in wheat. The octagon with brown, dark olive green and blue edges indicate that the genes were marked by H3K27me3, H3K4me3 and H3K9ac, respectively. Those genes with two modifications were in

octagons with either purple or light blue edges. **c** Effects of different haplotypes of three function known hub genes on GL/GW, GL (mm), GW (mm) and TGW (g). For GW, GL and TGW, the sample sizes for the H1 and H2 haplotypes of *TraesCS3B02G400200* were 218 and 107. For *TraesCS4B02G173600*, the sample sizes were 39 and 348 for GL, and 45 and 357 for both GW and TGW. For *TraesCS7A02G272100*, the H1 and H2 sample sizes were 218 and 49 across all three traits. The P value above the line is calculated using a two-tailed Student's t test. The box plots display the median (horizontal line), second to third quartiles (box), with the maxima and minima as whiskers, and outliers as points beyond 1.5 times the interquartile range. Source data are provided as a Source Data file.

stages of wheat endosperm, uncovering the comprehensive genetic and regulatory mechanisms that govern endosperm development and seed trait formation. In contrast, Zhao et al. focused on genes involved in starch and storage protein synthesis, mapping the dynamic distribution of histone modifications associated with these genes to elucidate their roles in endosperm development. Additionally, Zhao et al. generated ATAC-seq data for early embryo sac and endosperm (0–8 DPA) and we offered ATAC-seq data for late endosperm development (14 DPA and 18 DPA). Zhao et al. concentrated on H3K27ac and we examined H3K9ac. The information provided by these two studies is complementary, leading to a more comprehensive understanding of wheat endosperm development. Interestingly, we identified a key regulator, *ERF5* (*TraesCS6D02G225700*), which is also listed among the core TFs reported by Zhao et al. Similarly, the key gene *TaABI3-A1* (*TraesCS3A02G417300*), experimentally validated by Zhao et al., is recognized as a hub TF in the current study. The findings from both studies can be synergistically applied to advance wheat breeding programs by manipulating the identified key genes that influence grain quality and yield. In combination with observed effects caused by natural variations of the candidates on seed traits, genes controlling seed traits could be efficiently identified thus to provide new genetic resources to dissect seed traits for quality and yield improvement in wheat.

Methods

Plant materials and growth conditions

Field-grown Aikang58 (AK58, a widely grown bread wheat in China) was used for harvesting endosperm. To ensure that the endosperm we harvested were from the same stage, we marked the flower opening time (anthesis) of each single floret and only the flowers from middle spikelet were harvested. The endosperms were sampled at 4, 7, 14 and 18 DPA, respectively. The embryo was removed using blade from the endosperm. The endosperms were either snap frozen in liquid nitrogen for RNA extraction, or fixed in formaldehyde for ChIP-seq or stored at -80°C for ATAC-seq.

RNA-seq, ChIP-seq and ATAC-seq experiment

Total RNA of endosperm was extracted using StarSpin Hipure Fast Plant RNA Mini Kit (#PI25-01; Genstar, Beijing, China) according to the manufacturer's instructions. Three biological replicates were conducted for each developmental stage. After checking the concentration and purity of RNA using Nanodrop (Thermo Fisher Scientific) and electrophoresis, the qualified RNA was used to construct the sequencing library using the VAHTS® Universal V8 RNA-seq Library Prep Kit for MGI (#NRM605; Vazyme, Nanjing, China). The libraries were paired-end sequenced (2×150 bp) on DNBSEQ-T7 (BGI, Shenzhen).

ChIP-seq was conducted by following the eChIP protocol⁶³ with minor modifications. In brief, wheat endosperm was fixed with 1% formaldehyde. 0.2 g fixed samples were grinded in liquid nitrogen and then the powder was added to 300 μl of Buffer S (50 mM HEPES-KOH (pH7.5), 150 mM NaCl, 1 mM Ethylene Diamine Tetraacetic Acid (EDTA), 1% Triton X-100, 0.1% sodium deoxycholate, 1% SDS) for 10 min at 4°C . The chromatin was fragmented by sonicator (Covaris S220, USA). 50 μl protein G magnetic beads pre-washed by PBST were added

to the sonicated chromatin to pre-clean the chromatin by rotating for 1 h at 4°C . A small proportion of the chromatin was saved as input and the rest was used for ChIP. Chromatin immunoprecipitation was performed using antibodies against H3K4me3 (cat #A2357, lot #5500002383; Abclonal, Wuhan, China), H3K27me3 (cat #9733S, lot #16; Cell Signaling, Danvers, MA, USA) and H3K9ac (cat #9649S, lot #13; Cell Signaling). After incubating overnight at 4°C , the chromatin-ab complex was captured by protein G magnetic beads at 4°C . The precipitated chromatin was washed with washing buffer and ChIP-ed DNA was eluted from beads. Thereafter, de-crosslinking was performed and QIAquick® PCR Purification Kit (#28106; Qiagen, Dusseldorf, NRW, Germany) was used to purify the ChIP-DNA. The ThruPLEX® DNA-seq Kit was used for library preparation (#R400675; Takara bio, Japan). The libraries were submitted to NovaSeq 6000 system for sequencing with paired-end mode.

The ATAC-seq library was constructed following the nuclei sorting based protocol⁶⁴. In brief, the nuclei were isolated by rapidly and repeatedly chopping in buffer using a blade on an ice magnet. Then, the crude nuclei were filtered with 30 μm cell strainer and were then sorted by Fluorescence activated Cell Sorter. The tagmentation reaction was performed at 37°C for 30 min. DNA was purified by using the QIAGEN MinElute PCR Purification Kit (#28006; Qiagen). After amplification using NEBNext® High-Fidelity 2 \times PCR Master Mix (#M0541L; NEB, Beijing, China), the libraries were purified with AMPure beads. The concentration and fragment distribution were detected by Qubit and Agilent 2100 Bioanalyzer, respectively. Qualified ATAC libraries were sequenced by NovaSeq 6000 system with paired-end mode.

RNA-seq data processing

Adapter trimming was performed using Trim Galore (v0.4.2) (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Transcripts per million (TPM) values were generated using Kallisto⁶⁵ (v0.48.0) with an index built on Chinese Spring cDNA sequences (IWGSC RefSeq v1.1). High-confidence genes from this gene model version were used for further analysis. Differential expression genes ($|\log_2$ fold change| > 1 and q value < 0.05) and genes with time-series changes in expression (q value < 0.01) were detected using Sleuth⁶⁶ (v0.30.0) in R. The Shannon entropy specificity index was calculated to identify the stage specifically expressed genes via Tspex (<https://tspex.lge.ibi.unicamp.br/>).

ChIP-seq and ATAC-seq data processing

The raw data of ChIP-seq and ATAC-seq were filtered using Trimmomatic (v0.32) (<https://github.com/usadellab/Trimmomatic>) to remove sequencing adapters and low-quality reads, with parameters set to "LEADING:20 TRAILING:20 SLIDINGWINDOW:4:15 MINLEN:36". High-quality clean reads were aligned to IWGSC v1.0 reference genome using Bowtie2 (v2.4.4) (<https://github.com/BenLangmead/bowtie2>) with the parameters of "--sensitive -k 3". Mapped reads were sorted with SAMtools (v1.9) (<https://github.com/samtools/samtools>), and reads with a MAPQ score below 5 were discarded. PCR duplicates were removed using the MarkDuplicates function of Picard (v2.23.9) (<https://github.com/broadinstitute/picard>). Peak calling was performed using the IDR (Irreproducible Discovery Rate) framework, with

files prepared by SAMtools (v1.9). The callpeak function of MACS2 (v2.2.7.1) (<https://github.com/macs3-project/MACS>) was used for peak calling, with parameters “-p 1e-2 --nomodel --shift -75 --extsize 150 --mfold 2 20 --to-large” for ATAC-seq and “-c input.bam -p 1e-2 --mfold 2 20 --nomodel --to-large” for ChIP-seq. For data visualization, merged bam files were converted to RPKM (Reads Per Kilobase per Million mapped reads) normalized bigwig files with 50 bp bin size using deepTools⁶⁷ (v3.5.1). The peak annotation was performed using method from wheatATAC (<https://github.com/hcph/wheatATAC>). The promoter region covering 3500 bp upstream to 1500 bp downstream of the TSS. The gene body region spans from 1500 bp downstream of the TSS to the TES. All other regions are defined as distal regions. For the quantification of ChIP-seq and ATAC-seq peaks across developmental stages, the peaks detected from the four differential stages were merged initially and then reads counting was performed using the merge and multicov function of BEDTools⁶⁸ (v2.27), respectively. We divided the counts by peak length for normalization. The normalized counts were summed and each normalized count was divided by the total sum, and the resulted read count was multiplied by one million to obtain the final normalized value, termed Fragment Per Million (FPM, similar to TPM in RNA-seq). Additionally, the shannon entropy specificity index was calculated to identify the stage specific regions that were chromatin accessible or modified by H3K4me3, H3K27me3 and H3K9ac. The DAP-seq peaks were obtained from Gene Expression Omnibus database under the accession code of GSE188699⁴² and GSE192815⁴³. After annotating these peaks using wheatATAC pipeline, the target genes were identified.

Classification of expression pattern of triad genes

The triad of homoeologs from three sub-genomes of wheat are provided in Supplementary Data 18. Here, we defined a triad as homoeolog with maximum TPM value > 0.5. The ternary plot-based method was applied for defining expression patterns of triads³⁸. In brief, the expression level of homoeologous gene from one sub-genome was normalized against the total expression level within the triad to represent the relative expression. The Euclidean distance of each gene along the three angles of the ternary plot was then calculated using R packages of ggtern⁶⁹ (v3.3.5), and seven categories were defined according to the Euclidean distance. Specifically, homoeologs with a higher level of expression in the A sub-genome than the homoeologs in the other two sub-genomes were defined as A-dominant, and similarly, those with higher expression in the B or D sub-genomes were defined as B-dominant or D-dominant, respectively. Homoeologs with a lower level of expression in the A sub-genome than in the other two sub-genomes were defined as A-suppressed, and the same classification applies to B-suppressed and D-suppressed homoeologs. Other regions with no expression differences among sub-genomes were defined as balance. The expression patterns of the triad genes were visualized using R packages of sankeyD3 (v0.3.2) (<https://github.com/fbreitwieser/sankeyD3>).

Gene regulatory network inference

TOBIAS was used for identifying TF footprints and establishing gene regulatory relationship inference⁷⁰. TF footprints with --motifvalue < 1e-4 were retained. Wheat TFs were predicted by PlantTFDB (<http://plantfdb.gao-lab.org/prediction.php>) and checked best hits in *Arabidopsis*. Plant non-redundant PFMs downloaded from JASPAR CORE 2018 were used to scan the TF binding sites⁷¹. As one motif might be corresponded to multiple TFs, we further filtered the upstream TFs according to their expression patterns at particular stages. In detail, for footprints located in OCRs at 4 DPA, we considered those TFs that are either highly expressed at 4 DPA or at both 4 DPA and 7 DPA. Similarly, for footprints in OCRs at 7 DPA, we focused on those that are

specifically expressed at 7 DPA or at both 4 DPA and 7 DPA. For footprints located in OCRs at 14 DPA, we included TFs highly expressed at 14 DPA or at both 14 DPA and 18 DPA. For footprints located in OCRs at 18 DPA, we considered TFs that are highly expressed at 18 DPA or at both 14 DPA and 18 DPA. In addition, to study the impact of histone modifications on the development of endosperm, we paid particular attention on upstream TFs whose expression levels are influenced by histone modifications. The GRN for different endosperm developmental stages were integrated and visualized using Cytoscape⁷² (v3.9.1). The connectivity for each gene defined as the number of connections it has with other genes within the gene regulatory network was calculated. Based on these connectivity scores, we identified the top-level genes with the highest connectivity, indicating their central role in the network.

Verification of GRN via natural genetic variation

Genotypes were derived from the resequencing of a germplasm population (CNP0003712, <https://db.cngb.org/>). Seed traits were obtained from plant grown in Xiangyang (Hubei, China) in 2019; Xiangyang and Luoyang (Henan, China) in 2020; and Xiangyang, Luoyang and Wuhan (Hubei, China) in 2021. Seed traits used lme4⁷³ to calculate Best Linear Unbiased Prediction (BLUP), which has been provided in Supplementary Data 8. For the association analysis, we utilized Fast-LMM (v0.2.32) with the covariate set to 6, which was generated by PCA using Plink⁷⁴ (v1.9). We calculated blocks of genetic variations using Plink with the parameters --show-tags, --list-all, and --tag-r2 0.5. SNPs located within the same block were merged to form a QTL interval for subsequent analysis. We used the R package CMplot (version 4.5.1) (<https://github.com/YinLiLin/CMplot>) to generate Manhattan and QQ plots. The BEDTools⁶⁸ (v2.27) was used to retrieve closest genes associated to the QTLs. And the QTL regions were used to filter the gene regulatory network constructed based on gene expression and TF footprint. For all test genes, SNPs located within the gene body and 3500 bp upstream of the TSS were extracted using Plink, and Student's *t* test was conducted for each SNP to obtain significant SNPs ($P < 0.05$, two-tailed Student's *t* test) for GL, GW, GL/GW and TGW. The hub genes with genetic polymorphisms associated to significant ($P < 0.05$, two-tailed Student's *t* test) seed trait variations were left and the corresponding regulatory relationships were visualized by Cytoscape⁷² (v3.9.1).

Vector construction and wheat transformation

First, we used CRISPR-Cereal⁷⁵ to identify target site for knocking out *ERF5* (*TraesCS6A02G243500*, *TraesCS6B02G280800*, *TraesCS6D02G225700*). The chosen guide sequence is AGTACATCCACGAGCAGCTG. The gRNA was then introduced into the CRISPR/Cas9 binary vector PBUE414. The *ERF5*-KO vector plasmid was subsequently transformed into *Agrobacterium tumefaciens* strain EHA105 via chemical transformation. We then transformed immature wheat embryos of the “Felder” variety using this *Agrobacterium* strain, with the addition of *BABY BOOM* and *WUSCHEL* (*BBM-WUS*)⁷⁶. Finally, genomic DNA from the transgenic plants was extracted using the CTAB method, and PCR was employed to confirm the presence of transgenes and to check for residual *BBM-WUS* and *Cas9*.

The coding sequence of *ERF5* (*TraesCS6D02G225700*) was inserted into the pUI301 vector for overexpression via homologous recombination. The resulting *ERF5*-OE vector was then introduced into *Agrobacterium tumefaciens* strain EHA105 through chemical transformation. We transformed immature embryos of the wheat variety “Felder” using both the strain containing *ERF5*-OE construct and the assistant strain containing *BBM-WUS* construct⁷⁶. Finally, genomic DNA from the transgenic plants was extracted using the CTAB method, and three positive transgenic plants were confirmed by PCR analysis. The trans-positive plants with *BBM-WUS* segregated out were used for further analysis.

Reporting summary

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

Data availability

The sequencing data generated in this study have been deposited in the Genome Sequence Archive at the Beijing Institute of Genomics (BIG) Data Center, Chinese Academy of Sciences, under accession code [CRA009416](#). The processed ChIP-seq data are available in the Open Archive for Miscellaneous Data (OMIX) under accession code [OMIX002851](#). The public DAP-seq data were downloaded from the Gene Expression Omnibus database under the accession code [GSE192815](#) and [GSE188699](#). Source data are provided with this paper.

Code availability

The scripts and codes for this study can be found at GitHub repository [<https://github.com/hcph/wheat-endosperm-development>]⁷⁷.

References

- Akter, N. & Rafiqul Islam, M. Heat stress effects and management in wheat. *Agron. Sustain. Dev.* **37**, 37 (2017).
- Boz, H., Gerçekaslan, K. E., Karaoğlu, M. M. & Kotancilar, H. G. Differences in some physical and chemical properties of wheat grains from different parts within the spike. *Turkish J. Agric. Forestry* **36**, 309–316 (2012).
- Gardiner, L. J. et al. Analysis of the recombination landscape of hexaploid bread wheat reveals genes controlling recombination and gene conversion frequency. *Genome Biol.* **20**, 69 (2019).
- Evers, T. & Millar, S. Cereal grain structure and development: Some implications for quality. *J. Cereal Sci.* **36**, 261–284 (2002).
- Gao, X., Francis, D., Ormrod, J. C. & Bennett, M. D. Changes in cell number and cell division activity during endosperm development in allohexaploid wheat, *Triticum aestivum* L. *J. Exp. Bot.* **43**, 1603–1609 (1992).
- Leroux, B. M. et al. Maize early endosperm growth and development: from fertilization through cell type differentiation. *Am. J. Bot.* **101**, 1259–1274 (2014).
- Olsen, O. A. ENDOSPERM DEVELOPMENT: Cellularization and cell fate specification. *Annu Rev. Plant Physiol. Plant Mol. Biol.* **52**, 233–267 (2001).
- Shewry, P. R. et al. An integrated study of grain development of wheat (cv. Hereward). *J. Cereal Sci.* **56**, 21–30 (2012).
- Huang, L., Tan, H., Zhang, C., Li, Q. & Liu, Q. Starch biosynthesis in cereal endosperms: An updated review over the last decade. *Plant Commun.* **2**, 100237 (2021).
- Dai, D., Ma, Z. & Song, R. Maize endosperm development. *J. Integr. Plant Biol.* **63**, 613–627 (2021).
- Pfeifer, M. et al. Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science* **345**, 1250091 (2014).
- Guan, J. et al. Transcriptome analysis of developing wheat grains at rapid expanding phase reveals dynamic gene expression patterns. *Biology-Basel*. **11**, 281 (2022).
- Song, Y. et al. *TubZIP28*, a novel bZIP family transcription factor from *Triticum urartu*, and *TabZIP28*, its homologue from *Triticum aestivum*, enhance starch synthesis in wheat. *N. Phytol.* **226**, 1384–1398 (2020).
- Liu, G. et al. Virus-induced gene silencing identifies an important role of the *TaRSR1* transcription factor in starch synthesis in bread wheat. *Int. J. Mol. Sci.* **17**, 1557 (2016).
- Guo, D. et al. Over-expressing *TaSPA-B* reduces prolamins and starch accumulation in wheat (*Triticum aestivum* L.) grains. *Int. J. Mol. Sci.* **21**, 3257 (2020).
- Liu, Y. et al. The NAC transcription factor NAC019-A1 is a negative regulator of starch synthesis in wheat developing endosperm. *J. Exp. Bot.* **71**, 5794–5807 (2020).
- Gao, Y. et al. The endosperm-specific transcription factor *TaNAC019* regulates glutenin and starch accumulation and its elite allele improves wheat grain quality. *Plant Cell* **33**, 603–622 (2021).
- Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412 (2007).
- Lauberth, S. M. et al. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021–1036 (2013).
- Kratz, A. et al. Core promoter structure and genomic context reflect histone 3 lysine 9 acetylation patterns. *Bmc Genomics* **11**, 257 (2010).
- Cao, R. et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
- Oka, R. et al. Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize. *Genome Biol.* **18**, 137 (2017).
- Du, Z. et al. Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in *Oryza sativa* L. Japonica. *Mol. Plant*. **6**, 1463–1472 (2013).
- Huang, Y. et al. Wide Grain 7 increases grain width by enhancing H3K4me3 enrichment in the *OsMADS1* promoter in rice (*Oryza sativa* L.). *Plant J.* **102**, 517–528 (2020).
- Nallamilli, B. R. et al. Polycomb group gene *OsFIE2* regulates rice (*Oryza sativa*) seed development and grain filling via a mechanism distinct from *Arabidopsis*. *Plos Genet* **9**, e1003322 (2013).
- Liu, X. et al. *SDG711* is involved in rice seed development through regulation of starch metabolism gene expression in coordination with other histone modifications. *Rice* **14**, 25 (2021).
- Kim, E. D. et al. Spatio-temporal analysis of coding and long non-coding transcripts during maize endosperm development. *Sci. Rep.* **7**, 3838 (2017).
- Tian, Q. et al. Riboflavin integrates cellular energetics and cell cycle to regulate maize seed development. *Plant Biotechnol. J.* **20**, 1487–1501 (2022).
- Weinhofer, I., Hehenberger, E., Roszak, P., Hennig, L. & Kohler, C. H3K27me3 profiling of the endosperm implies exclusion of polycomb group protein targeting by DNA methylation. *PLoS Genet.* **6**, e1001152 (2010).
- Dong, X. et al. Dynamic and antagonistic allele-specific epigenetic modifications controlling the expression of imprinted genes in maize endosperm. *Mol. Plant*. **10**, 442–455 (2017).
- Makarevitch, I. et al. Genomic distribution of maize facultative heterochromatin marked by trimethylation of H3K27. *Plant Cell* **25**, 780–793 (2013).
- Zhao, L. et al. Dynamic chromatin regulatory programs during embryogenesis of hexaploid wheat. *Genome Biol.* **24**, 7 (2023).
- Gillies, S. A., Futardo, A. & Henry, R. J. Gene expression in the developing aleurone and starchy endosperm of wheat. *Plant Biotechnol. J.* **10**, 668–679 (2012).
- Barthole, G. et al. MYB118 represses endosperm maturation in seeds of *Arabidopsis*. *Plant Cell* **26**, 3519–3537 (2014).
- Liu, H. et al. Pleiotropic *ZmICE1* is an important transcriptional regulator of maize endosperm starch biosynthesis. *Front. Plant Sci.* **13**, 895763 (2022).
- Chen, Z. et al. Expression of *AtAAP* gene family and endosperm-specific expression of *AtAAP1* gene promotes amino acid absorption in *Arabidopsis thaliana* and maize. *Agronomy* **11**, 1668 (2021).
- Gehring, M., Missirian, V. & Henikoff, S. Genomic analysis of parent-of-origin allelic expression in *Arabidopsis thaliana* seeds. *PLoS One* **6**, e23687 (2011).
- Ramirez-Gonzalez, R. H. et al. The transcriptional landscape of polyploid wheat. *Science*. **361**, eaar6089 (2018).
- Zhao, T., Zhan, Z. & Jiang, D. Histone modifications and their regulatory roles in plant development and environmental memory. *J. Genet. Genomics.* **46**, 467–476 (2019).

40. Ma, C. et al. iTRAQ-based quantitative proteome and phosphoprotein characterization reveals the central metabolism changes involved in wheat grain development. *BMC Genomics* **15**, 1029 (2014).
41. Kaufmann, K., Pajoro, A. & Angenent, G. C. Regulation of transcription in plants: mechanisms controlling developmental switches. *Nat. Rev. Genet.* **11**, 830–842 (2010).
42. Pei, H. et al. Low-affinity SPL binding sites contribute to subgenome expression divergence in allohexaploid wheat. *Sci. China Life Sci.* **66**, 819–834 (2023).
43. Zhang, Y. et al. Transposable elements orchestrate subgenome-convergent and -divergent transcription in common wheat. *Nat. Commun.* **13**, 6940 (2022).
44. Sun, F. et al. Functional characterization of *TaFUSCA3*, a B3-Superfamily transcription factor gene in the wheat. *Front. Plant Sci.* **8**, 1133 (2017).
45. Zhu, J. et al. 5-Azacytidine treatment and *TaPBF-D* over-expression increases glutenin accumulation within the wheat grain by hypomethylating the *Glu-1* promoters. *Theor. Appl. Genet.* **131**, 735–746 (2018).
46. Guo, W. et al. The wheat transcription factor *TaGAMyb* recruits histone acetyltransferase and activates the expression of a high-molecular-weight glutenin subunit gene. *Plant J.* **84**, 347–359 (2015).
47. Yoo, S. D., Cho, Y. H., Tena, G., Xiong, Y. & Sheen, J. Dual control of nuclear EIN3 by bifurcate MAPK cascades in C2H4 signalling. *Nature* **451**, 789–795 (2008).
48. Pan, Y. et al. An ethylene response factor (ERF5) promoting adaptation to drought and salt tolerance in tomato. *Plant Cell Rep.* **31**, 349–360 (2012).
49. Troncoso-Ponce, M. A. et al. Transcriptional activation of two delta-9 Palmitoyl-ACP desaturase genes by MYB115 and MYB118 is critical for biosynthesis of omega-7 monounsaturated fatty acids in the endosperm of *Arabidopsis* seeds. *Plant Cell* **28**, 2666–2682 (2016).
50. Zhang, Y., Cao, G., Qu, L. J. & Gu, H. Involvement of an R2R3-MYB transcription factor gene *AtMYB118* in embryogenesis in *Arabidopsis*. *Plant Cell Rep.* **28**, 337–346 (2009).
51. Franken, P., Schrell, S., Peterson, P. A., Saedler, H. & Wienand, U. Molecular analysis of protein domain function encoded by the myb-homologous maize genes *C1*, *Zm 1* and *Zm 38*. *Plant J.* **6**, 21–30 (1994).
52. Villagomez-Aranda, A. L. et al. Activating stress memory: eustressors as potential tools for plant breeding. *Plant Cell Rep.* **41**, 1481–1498 (2022).
53. Angel, A., Song, J., Dean, C. & Howard, M. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476**, 105–108 (2011).
54. He, C., Chen, X., Huang, H. & Xu, L. Reprogramming of H3K27me3 is critical for acquisition of pluripotency from cultured *Arabidopsis* tissues. *Plos Genet* **8**, e1002911 (2012).
55. Atlasi, Y. & Stunnenberg, H. G. The interplay of epigenetic marks during stem cell differentiation and development. *Nat. Rev. Genet.* **18**, 643–658 (2017).
56. Jia, H., Suzuki, M. & McCarty, D. R. Regulation of the seed to seedling developmental phase transition by the LAFL and VAL transcription factor networks. *Wiley Interdiscip. Rev. Dev. Biol.* **3**, 135–145 (2014).
57. Kagaya, Y. et al. LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of *FUSCA3* and *ABSCISIC ACID INSENSITIVE3*. *Plant Cell Physiol.* **46**, 399–406 (2005).
58. Ravindran, P., Verma, V., Stamm, P. & Kumar, P. P. A novel RGL2-DOF6 complex contributes to primary seed dormancy in *Arabidopsis thaliana* by regulating a GATA transcription factor. *Mol. Plant.* **10**, 1307–1320 (2017).
59. Zinsmeister, J. et al. The seed-specific heat shock factor A9 regulates the depth of dormancy in *Medicago truncatula* seeds via ABA signalling. *Plant Cell Environ.* **43**, 2508–2522 (2020).
60. van Ginkel, M. & Ogonnaya, F. Novel genetic diversity from synthetic wheats in breeding cultivars for changing production conditions. *Field Crop. Res.* **104**, 86–94 (2007).
61. Zhai, H. et al. QTL Analysis of spike morphological traits and plant height in winter wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map. *Front. Plant Sci.* **7**, 1617 (2016).
62. Zhao, L. et al. Deciphering the transcriptional regulatory network governing starch and storage protein biosynthesis in wheat for breeding improvement. *Adv. Sci.* **11**, e2401383 (2024).
63. Zhao, L. et al. Integrative analysis of reference epigenomes in 20 rice varieties. *Nat. Commun.* **11**, 2658 (2020).
64. Wang, X., Chen, C., He, C., Chen, D. & Yan, W. Mapping open chromatin by ATAC-seq in bread wheat. *Front. Plant Sci.* **13**, 1074873 (2022).
65. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
66. Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* **14**, 687–690 (2017).
67. 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).
68. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
69. Hamilton, N. E. & Ferry, M. ggtern: Ternary diagrams using ggplot2. *J. Stat. Softw.* **87**, 1–17 (2018).
70. Bentsen, M. et al. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat. Commun.* **11**, 4267 (2020).
71. Khan, A. et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **46**, D260–D266 (2018).
72. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
73. Bates, D. et al. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
74. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
75. He, C. et al. Cereal: a guide RNA design tool integrating regulome and genomic variation for wheat, maize and rice. *Plant Biotechnol. J.* **19**, 2141–2143 (2021).
76. Zhou, Z. et al. Overcoming genotypic dependency and bypassing immature embryos in wheat transformation by using morphogenic regulators. *Sci. China Life Sci.* **67**, 1535–1538 (2024).
77. He, C. et al. Dynamic atlas of histone modifications and gene regulatory networks in endosperm of bread wheat. Github. <https://doi.org/10.5281/zenodo.13767298> (2024).

Acknowledgements

This research was supported by the Biological Breeding-National Science and Technology Major Project (2023ZD0406802) to H.L.M and the Science and Technology Major Program of Hubei Province (NO.2022ABA001) to W.H.Y. We thank Prof. Dr. Kerstin Kaufmann from Humboldt-Universität zu Berlin for her critical reading and suggestions on the manuscript. We thank the China National GeneBank and the high-performance computing platform at the National Key Laboratory of Crop Genetic Improvement at Huazhong Agricultural University for data storage and server for data processing.

Author contributions

W.H.Y. and H.L.M. conceived and designed the study. Y.Q.L., C.H. and C.X.S. performed the experiment and collect the data with help from H.P.Z., X.T.X., S.S. and C.J.Z., C.H. and S.T.B. performed the data analyses. C.H. and S.T.B. draft the manuscript with input from H.D.S, C.X.L., Q.L. and W.C., W.H.Y. finalized the manuscript. All authors reviewed and approved 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/s41467-024-53300-7>.

Correspondence and requests for materials should be addressed to Hailiang Mao or Wenhao Yan.

Peer review information *Nature Communications* thanks Laura-Jayne Gardiner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

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

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

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

© The Author(s) 2024