

The transcriptome landscape of developing barley seeds

Martin Kovacik ^{1,2} Anna Nowicka ^{1,3} Jana Zwyrteková ¹ Beáta Strejčková ¹ Isaia Vardanega ⁴
Eddi Esteban ⁵ Asher Pasha ⁵ Kateřina Kaduchová ¹ Maryna Krautsova ¹ Marie Červenková ¹
Jan Šafář ¹ Nicholas J. Provart ⁵ Rüdiger Simon ⁴ Ales Pecinka ^{1,2,*}

- 1 Institute of Experimental Botany, Czech Acad Sci, Centre of Plant Structural and Functional Genomics, Šlechtitelů 31, 779 00 Olomouc, Czech Republic
- 2 Department of Cell Biology and Genetics, Faculty of Science, Palacký University, Šlechtitelů 27, 779 00 Olomouc, Czech Republic
- 3 Franciszek Górski Institute of Plant Physiology Polish Academy of Sciences, Niezapominajek 21, 30 239 Kraków, Poland
- 4 Institute for Developmental Genetics, Heinrich-Heine-University, Universitätsstraße 1, 40225 Düsseldorf, Germany
- 5 Department of Cell and Systems Biology/Centre for the Analysis of Genome Evolution and Function, University of Toronto, 25 Willcocks St., Toronto, ON M5S 3B2, Canada

*Author for correspondence: pecinka@ueb.cas.cz (A.P.)

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plcell/pages/General-Instructions>) is: Ales Pecinka (pecinka@ueb.cas.cz).

Abstract

Cereal grains are an important source of food and feed. To provide comprehensive spatiotemporal information about biological processes in developing seeds of cultivated barley (*Hordeum vulgare* L. subsp. *vulgare*), we performed a transcriptomic study of the embryo, endosperm, and seed maternal tissues collected from grains 4–32 days after pollination. Weighted gene co-expression network and motif enrichment analyses identified specific groups of genes and transcription factors (TFs) potentially regulating barley seed tissue development. We defined a set of tissue-specific marker genes and families of TFs for functional studies of the pathways controlling barley grain development. Assessing selected groups of chromatin regulators revealed that epigenetic processes are highly dynamic and likely play a major role during barley endosperm development. The repressive H3K27me3 modification is globally reduced in endosperm tissues and at specific genes related to development and storage compounds. Altogether, this atlas uncovers the complexity of developmentally regulated gene expression in developing barley grains.

Introduction

Seeds are crucial structures in the life cycle of many plants that allow the survival of long periods of unfavorable conditions and colonization of new sites. High nutritive value makes seeds prime targets of plant breeding and cereal grains are one of the most valuable agronomic products (Carena 2009). Cultivated barley (*Hordeum vulgare* L. subsp. *vulgare*) is the fourth most important cereal worldwide and it is used as feed (70%), for the production of malt (21%), and as food

(9%; Food and Agriculture Organization of the United Nations 2023). Barley grains have three main compartments: embryo, endosperm, and seed maternal tissues (SMTs), each consisting of different parts and cell types (Gubatz et al. 2007). The diploid embryo originates from a fusion of the egg cell nucleus and sperm cell nucleus. A mature embryo consists of the embryonic axis (coleoptile, plumule, shoot apical meristem, radicle, and coleorhiza) and the scutellum. The triploid endosperm derives from a fusion of the diploid central cell nucleus and the second sperm cell nucleus. After an

Received August 18, 2023. Accepted March 05, 2024. Advance access publication April 18, 2024.

© The Author(s) 2024. Published by Oxford University Press on behalf of American Society of Plant Biologists. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

initial unicellular multinucleate coenocyte stage, cereal endosperm cellularizes and differentiates into the central starchy endosperm (CSE) that serves as the main storage of complex sugars, the aleurone layer (AL), the embryo surrounding region (ESR), and the endosperm transfer cells (ETCs; Olsen 2001). Finally, an outer layer of SMTs, consisting of a pericarp and two layers of seed coat, maintains stable conditions, transports assimilates, protects the embryo and endosperm, and serves as the first source of starch during early grain development (Weschke et al. 2003; Radchuk et al. 2009). The molecular mechanisms governing the development of individual seed tissues in barley remain poorly understood.

Spatiotemporal analysis of gene expression plays a crucial role in understanding developmental programs. Previous studies have explored transcriptional profiles of developing barley grains and their parts. An early study using expressed sequence tags revealed transcriptional reprogramming at 0–7 days after flowering in seed maternal and filial tissues and early to late whole caryopses (Zhang et al. 2004). Genes encoding many protein- and lipid-activating enzymes were up-regulated while genes coding for reactive oxygen species-scavenging enzymes were down-regulated in SMTs, indicating mobilization of storage compounds and programmed cell death, respectively. The filial tissues contained highly expressed genes coding for factors involved in cell growth and cell wall biosynthesis. Expression microarray-based barley genome-wide transcriptomic studies used samples from dissected embryos and endosperm 16–25 days after pollination (DAP) and whole caryopses 5–16 DAP (Druka et al. 2006; Sreenivasulu et al. 2008). This revealed dynamic changes in gene expression related to metabolic and phytohormonal pathways during grain development and germination. RNA sequencing (RNA-seq) of whole caryopses explored the extent of RNA editing in barley grain development (Bian et al. 2019). Most recently, a spatially resolved cellular map for germinating barley seeds was provided and showed enrichment of different processes in specific tissues (Peirats-Llobet et al. 2023).

Here, we performed a comprehensive transcriptome profiling of barley seed tissues at different DAP timepoints with an aim to create an atlas of gene expression and thus provide detailed information on the temporal and spatial distribution of the key molecular processes acting during barley grain development. Our focus on nucleus-driven processes will be fundamental for future functional studies of the role of the epigenome in barley grain development.

Results

Generating tissue-specific transcriptome profiles of developing barley seeds

To identify transcriptional signatures of the major tissues of developing barley seeds, we performed RNA-seq of the manually dissected embryo, endosperm, and SMTs of barley cultivar (cv.) Morex at 4, 8, 16, 24, and 32 DAP (Fig. 1, A and B).

Sample hierarchical clustering revealed a strict grouping of the tissues and time points, except for the 4 DAP endosperm that clustered separately, indicating its unique transcriptome (Supplementary Fig. S1). This pattern was further corroborated by principal component analysis (PCA), which revealed a predominant grouping by DAP (PC1) and tissue (PC2), explaining 70% of the total variability (Fig. 1C). The 4 and 8 DAP endosperm samples were notably distant from the later time points, indicating a massive transcriptional reprogramming during endosperm proliferation and cellularization. The closer distance between the 24 and 32 DAP samples for both embryo and endosperm suggested relatively minor transcriptional changes toward the end of grain development. No major changes in clustering were found by inspecting PC3 (18% of the variability; Supplementary Fig. S2, A and B). When compared with the transcriptomes of eight different tissues (The International Barley Genome Sequencing Consortium 2012), the overall distribution was defined by the seed samples. The vegetative tissues (root, shoot, nodule, and inflorescences) clustered together with the germinating embryo (Fig. 1D and Supplementary Fig. S2, C and D). On the contrary, caryopsis 5 and 15 days after anthesis grouped with the 4 and 16 DAP endosperm, respectively, suggesting that the endosperm tissues dominate the caryopsis transcriptome.

To allow easy visualization of the transcriptomic data in a user-friendly format, we integrated our dataset into the Barley ePlant on the Bio-Analytic Resource (BAR) for Plant Biology (Fig. 1E; https://bar.utoronto.ca/eplant_barley/). The expression levels for individual genes (shown as transcripts per million, TPM; Supplementary Data Set 1) revealed striking differences between the tissues. The endosperm median expression was 0.96–1.35 TPM compared to 1.80–2.03 TPM in the embryo and 1.81–2.11 TPM in SMTs (Fig. 1F). A substantial portion of genes with low expression (TPM 0–1; $n = 31,571$) were found in different parts of the seeds, specifically 67.8% in the endosperm ($n = 21,400$), 50.7% in the embryo ($n = 16,009$), and 55.0% in SMT ($n = 17,364$; Fig. 1G). Despite having a 2-fold lower median expression, endosperm contained the majority of the highly expressed genes (Fig. 1H and Supplementary Fig. S3). These genes were significantly enriched in the Gene Ontology (GO) molecular functions of negative regulation of proteolysis, defense response, development, response to wounding, lipid transport, and cell wall macromolecule catabolic processes (Supplementary Table S1).

Grain development is associated with extensive transcriptional changes

Spatial and temporal changes in the seed transcriptome were assessed by plotting differentially expressed genes (DEGs) at one or more time points for each tissue (Fig. 2A). During embryo development, the major transcriptional changes occurred from 8 to 24 DAP. Most DEGs were detected from 8 to 16 DAP ($n = 8,952$) and 16 to 24 DAP ($n = 10,340$;

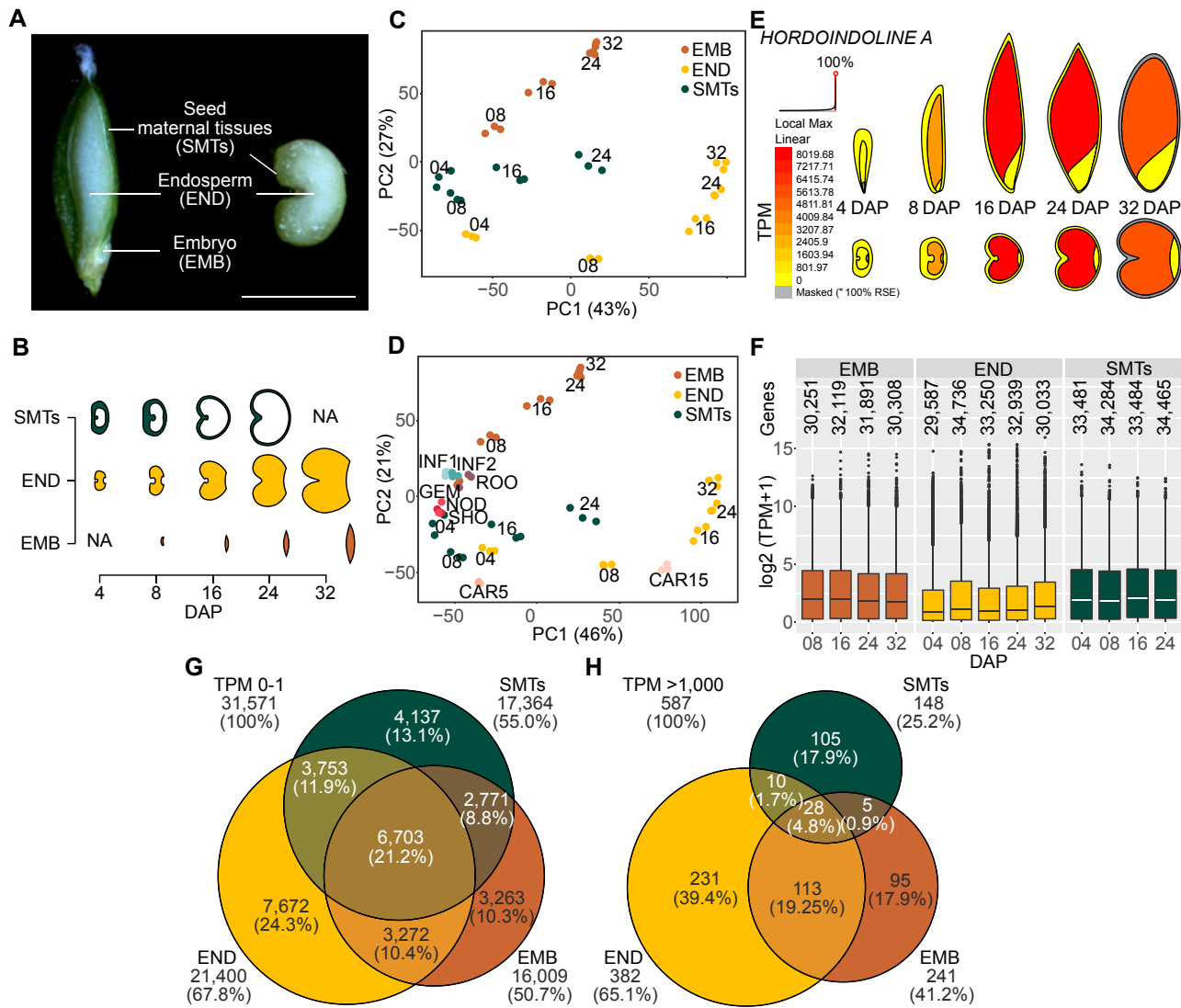


Figure 1. Transcriptomes of developing barley grain. **A and B**) Overview of the analyzed tissues and time points used for transcriptomic analysis. Scale bar, 5 mm. DAP, days after pollination; NA, not analyzed. **C**) The variance of the 39 samples represented by PCA for the embryo, endosperm, and SMTs. The numbers within the graph indicate DAP and three close spots represent biological triplicates. The number next to PC indicates variance. **D**) PCA of seed tissues used in **C**) in combination with other barley tissues (The International Barley Genome Sequencing Consortium 2012). ROO, root; GEM, germinating embryo; NOD, nodule; SHO, shoot; INF1 and INF2, developing inflorescence of 5- and 10-mm length; CAR5 and CAR15; caryopsis 5 and 15 DAP. **E**) Visualization of interactive heatmap in the Barley ePlant. Heatmap in transverse (upper) and sagittal (lower) sections shows expression levels of *HORDOINDOLINE A* in different tissues of developing barley grain. Gene expression in transcript per million (TPM) is indicated by the color scale. **F**) Boxplots of expression for genes with non-zero expression (source values in Supplementary Data Set 1). The lower and upper hinges of the boxes correspond to the first and third quartiles, respectively. The line within the boxes marks the median, whiskers extend to 1.5 times the interquartile range, and points represent the outliers. **G**) Venn diagram showing numbers of lowly expressed genes (0–1 TPM during entire grain development). **H**) Venn diagram of the highest expressed genes (>1,000 TPM).

false discovery rate (FDR)-adjusted $P < 0.05$; Fig. 2A, Supplementary Fig. S4A, and Supplementary Data Set 2). The major changes in embryo occurred between 8 and 16 DAP (8,952 DEGs) and 16 and 24 DAP (10,340 DEGs), while in SMTs more DEGs were observed from 4 to 8 DAP (5,604 DEGs) and 8 to 16 DAP (8,264 DEGs). The endosperm showed massive transcriptional changes, including a subset of unique genes from 4 to 8 DAP (15,990 DEGs and 9,869 unique). The number of endosperm DEGs gradually decreased toward 32

DAP (Fig. 2A and Supplementary Fig. S4B). There was a trend similar to embryos also in SMTs, with more DEGs between the early and middle time points (5,604 between 4 and 8 DAP and 8,264 between 8 and 16 DAP; FDR-adjusted $P < 0.05$; Fig. 2A, Supplementary Fig. S4C, and Supplementary Data Set 2) and less at the late stage of grain development (3,953 DEGs at 16–24 DAP).

We associated transitions between individual DAPs with over-/under-represented GO terms (Fig. 2B and

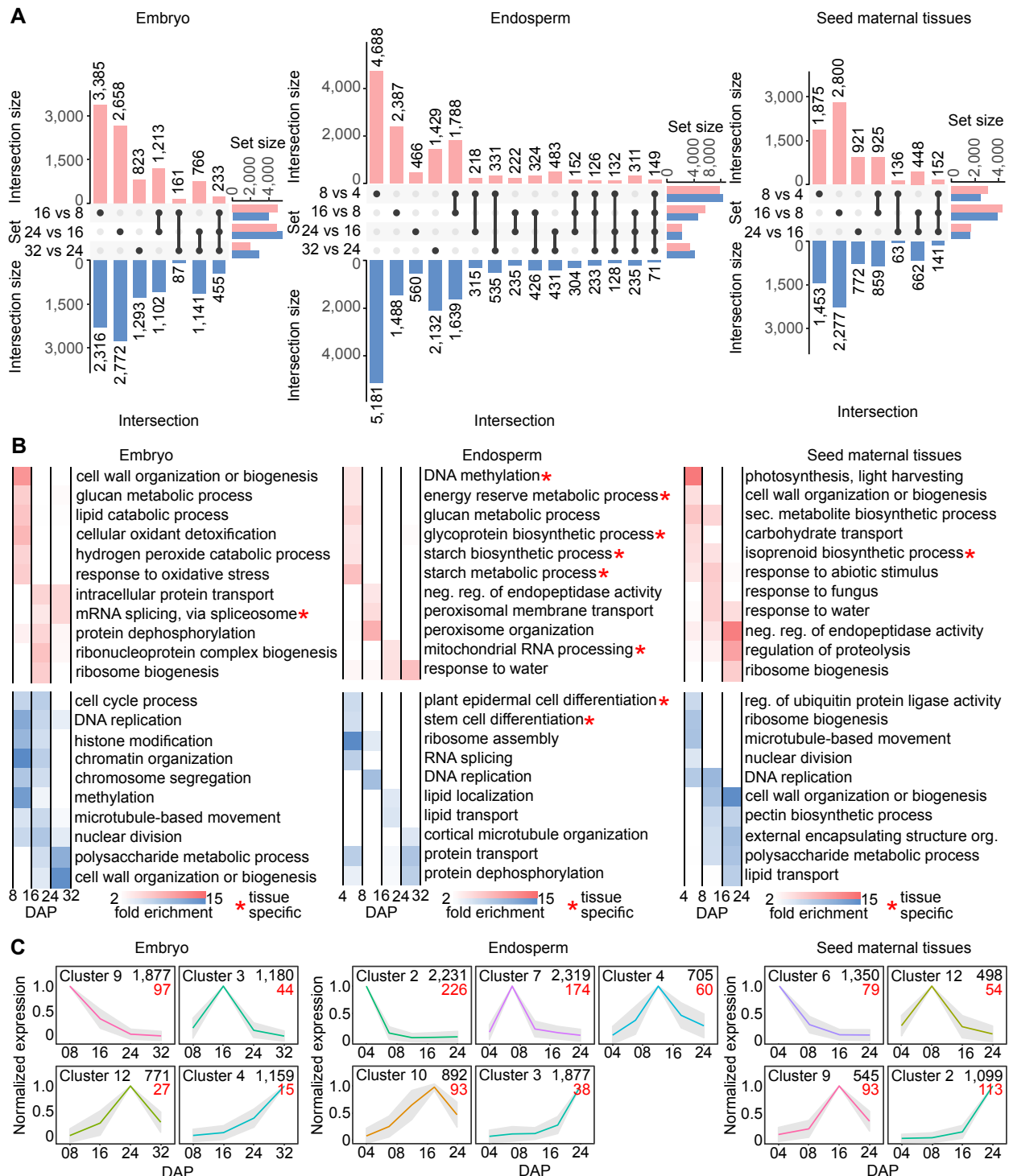


Figure 2. Tissue and timepoint specificity of gene expression during barley grain development. **A**) The UpSet plots show up-regulated (upper part) and down-regulated (lower part) DEGs in embryo, endosperm, and SMTs. The set size on the x-axis defines the total number of DEGs between two subsequent experimental time points. The y-axis shows the number of DEGs (intersection size) in single stages (single dots) and their combinations (connected dots; source data in [Supplementary Data Set 2](#)). **B**) GO term enrichment of DEGs (FDR-adjusted $P < 0.05$) between developmental time points in embryo, endosperm, and SMTs. The top representative enriched GO categories among up-regulated (upper part) or down-regulated (lower part) genes are shown. Color saturation corresponds to fold enrichment (full list in [Supplementary Tables S2 and S3](#)). Tissue-specific GO categories are indicated by an asterisk. neg., negative; reg., regulation; sec., secondary; org., organization. **C**) K-means co-expression clusters of DEGs, peaking in single consecutive developmental time points in embryo, endosperm, and SMTs. The gray area indicates standard deviation. The numbers in the upper right corner indicate gene count in individual clusters (upper) and a subset of tissue-specific marker genes (lower). DAP, days after pollination (the full list in [Supplementary Fig. S5](#) and [Supplementary Data Set 3](#)).

Supplementary Data Set 3). Embryos at 8 DAP showed enrichment in terms linked to cell division (GO terms cell cycle, DNA replication, histone, and chromatin modification) and their reduction before 16 DAP. The 16 DAP embryo transcriptome indicated intense ribosome biogenesis and cell wall biosynthesis. However, these enriched GO terms are typical for many actively growing tissues and were also shared to a large extent with endosperm and SMTs (Supplementary Data Sets 3 and 4). The GO term “mRNA splicing via spliceosome” was found to be significantly enriched at 16 DAP and even more so after 24 DAP (Fig. 2B). This is consistent with the accumulation of mRNAs in developing seeds and subsequent translation during germination (Sano et al. 2020). In endosperm, many tissue-specific enriched GO terms were associated with storage compounds (GO terms: lipid-, starch-, glucan-, glycoprotein-biosynthesis and metabolism; Fig. 2B and Supplementary Data Sets 3 and 4), which underlie the role of endosperm as the main nutritive tissue in grains. Furthermore, there was a tissue-specific expression of genes coding for DNA methylation factors from 4 to 8 DAP. The most prominent enriched GO terms in SMTs included photosynthesis and cell wall development (both peaking at 8 DAP), and a wave of expression from fungi and abiotic stress-responsive factors (Fig. 2B and Supplementary Data Sets 3 and 4). The unique enriched GO term of SMTs was the upregulation of isoprenoid biosynthesis from 4 to 16 DAP. Accumulation of isoprenoids is important for seed nutritional and physiological quality (Vishal and Kumar 2018).

Using k-means clustering, we defined molecular marker genes for individual tissues and DAP timepoints (Fig. 2C, Supplementary Fig. S5, and Supplementary Data Set 5). The embryo ($n = 17,214$), endosperm ($n = 21,889$), and SMTs ($n = 15,034$) DEGs were divided into 12, 13, and 14 co-expression clusters, respectively, where four to five clusters showed expression peaking at single consecutive experimental points. We searched the clusters for tissue-specific genes defined as having less than 5% TPM in other seed tissues (Fig. 2C and Supplementary Fig. S5—red numbers). Most of the stage- and tissue-specific genes in embryo and endosperm were found in 4 and 8 DAP-specific clusters.

Specific promoter motifs are associated with transcriptional regulation of seed development

To identify expression correlations among genes, we performed weighted gene co-expression network analysis (WGCNA) for each tissue. The identified WGCNA modules were organized temporally according to the early, middle, and late stages of development (Fig. 3A, Supplementary Fig. S6, and Supplementary Data Set 6). The majority of the genes were found to accumulate either at early (EMB_M2, $n = 1,883$; END_M2/M7/M6, $n = 1,962$; and SMT_M1, $n = 1,572$) or late (EMB_M1, $n = 1,934$; END_M1/M2, $n = 1,838$; and SMT_M3, $n = 2,497$) modules of individual tissues, suggesting roles for these genes in transcriptional reprogramming.

Next, we analyzed transcription factors (TFs) and regulatory motifs important during barley grain development by performing promoter motif enrichment analysis. The regions $-1,500$ to -1 bp from the transcription start site of barley genes included in WGCNA modules were analyzed for the presence of known TF binding motifs using the Arabidopsis (*Arabidopsis thaliana*) HOMER database (Heinz et al. 2010). This resulted in collections of significantly enriched motifs (P -value < 0.05 ; Supplementary Data Set 7) that grouped into clusters based on similarity (Supplementary Data Sets 8 and 9). The proportion of major motif clusters (MCs) in each collection was quantified (Fig. 3B). The MC2 and MC9 were enriched across many WGCNA modules of all tissues and their consensus motifs corresponded to the G-box (CACGTG) or Prolamin-box (P-box; AWAAAG), respectively (Fig. 3B).

The common TF families predicted to bind MC2 DNA motifs were the basic/helix-loop-helix family and basic region/leucine zipper family members known to bind the G-box and TKACGT motif variants (Supplementary Fig. S7A). To test whether the predicted TFs in those families might be further supported by our data, we identified barley putative orthologs of Arabidopsis TFs (Supplementary Data Set 10), and retrieved their expression profiles from our data (Supplementary Fig. S7B and Supplementary Data Set 10). The expression profiles of many TFs appeared early or late during grain development, e.g. the barley homolog of ABA INSENSITIVE 5 (*ABIS*) exhibited an expression profile peaking later in embryo and endosperm development. In Arabidopsis, *ABIS* is involved in signaling during seed maturation and regulates a subset of LATE EMBRYOGENESIS ABUNDANT (*LEA*) genes (Goyal et al. 2005). Many *LEA* genes were highly expressed in the late embryo module, and some were also found to be expressed in the late endosperm module (Supplementary Data Set 6).

The MC9 was enriched particularly with genes expressed during middle embryo and endosperm development and these contained the characteristic DNA-binding with one finger (Dof) TF family binding P-box motif (A/T)AAAG (Fig. 3, B and C and Supplementary Data Set 11). Dof type TFs PROLAMIN BINDING FACTOR (PBF) and DOF ZINC FINGER PROTEIN 1 (DOF1) have been shown to regulate seed storage protein synthesis in barley and maize (*Zea mays*), respectively (Mena et al. 1998; Yanagisawa and Sheen 1998). Barley PBF presumably activates the expression of genes related to storage compound biosynthesis in endosperm by binding to the P-box motif present in the promoters of *Hordein* genes (Mena et al. 1998). Our in situ hybridization data extended the knowledge about spatial and temporal expression patterns of this gene in barley endosperm (Fig. 3, D and E and Supplementary Fig. S8). The PBF transcript was detected around the embryonic pole starting at 8 DAP and its pattern of expression expanded to the endosperm periphery at the ventral side toward 16 DAP. Observed expression in embryos at 16 DAP further confirms its suggested role in embryo development (Cook et al. 2018). As the P-box motif was enriched in the middle endosperm module, we investigated

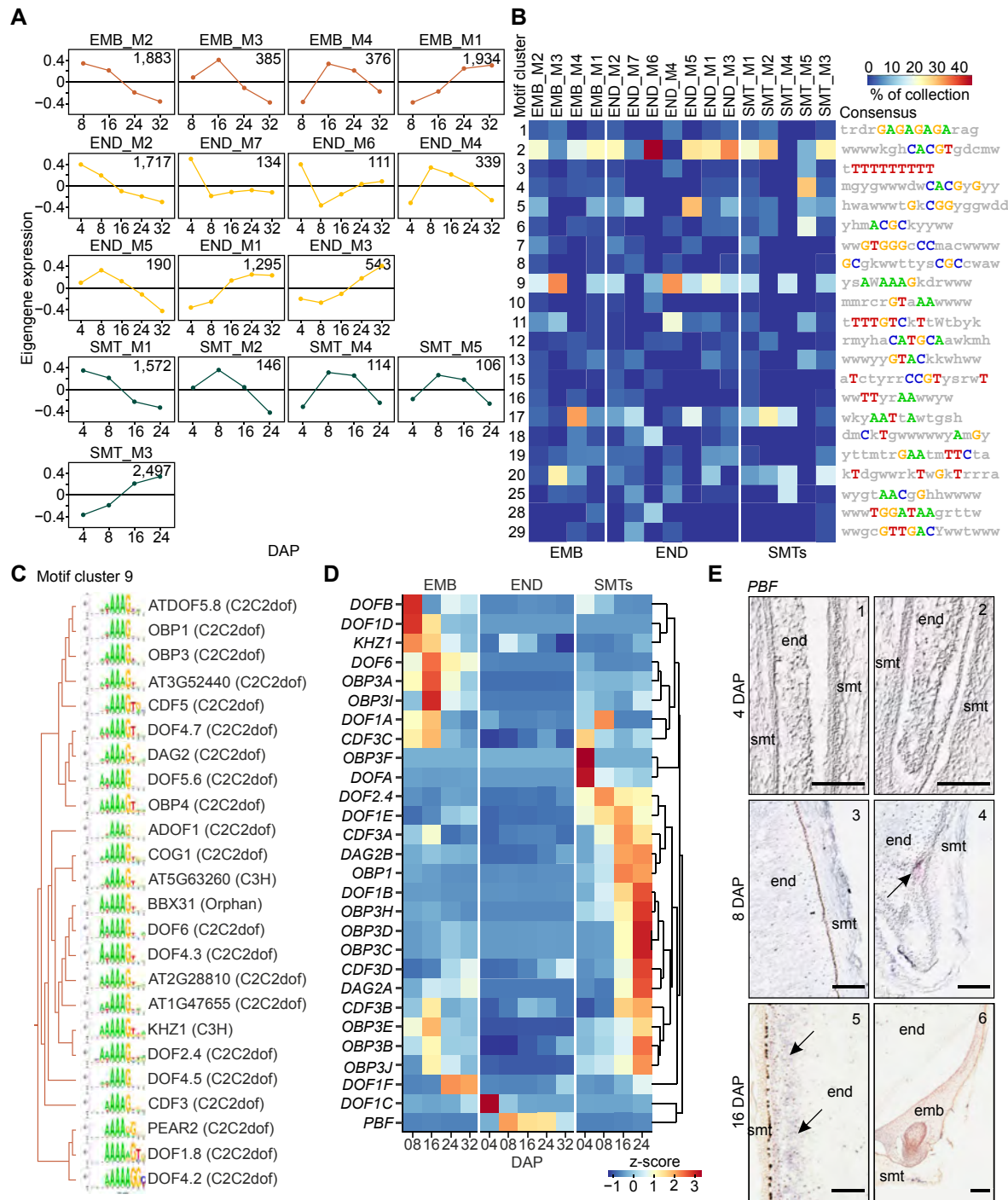


Figure 3. Gene co-expression network analysis and promoter motif enrichment. **A**) Display of selected WGCNA modules (M1–M7; the full list of modules is provided in [Supplementary Data Set 4](#) and [Supplementary Fig. S6](#)). The graphs show eigengene expression in each module. The numbers in the upper right corner indicate gene count in individual clusters. EMB, embryo; END, endosperm; SMTs, seed maternal tissues; DAP, days after pollination. **B**) Analysis of TF binding motifs in individual WGCNA modules shown in (A) by screening promoters of barley genes against the Arabidopsis TF binding motif database and clustering. The consensus motif in each MC (rows) is shown on the right side and its representation (%) in the WGCNA module is indicated by the color scale. The heatmap depicts some motifs from individual clusters identified in each WGCNA module containing at least 10 motifs (full list in [Supplementary Table S5](#)). **C**) Hierarchical clustering of TF binding sequence motifs in MC9 with the in silico-predicted binding Arabidopsis TFs and their families (in parentheses). **D**) Heatmap of hierarchically clustered expression for barley putative orthologs of Arabidopsis TFs from (C) in EMB, END, and SMTs. Gene expression z-score is indicated by the color scale (source data in [Supplementary Table S7](#)). **E**) RNA in situ hybridization of PBF with antisense probe in barley grains at 4, 8, and 16 DAP. emb, embryo; end, endosperm; SMT, seed maternal tissues. Arrows indicate signal deposition in the endosperm. The full-scale images are presented in [Supplementary Fig. S8](#) and the numbers in the top right corners correspond to the inset numbers. Scale bars, 200 μ m.

the genes most strongly expressed in the END_M4 module. We found increased transcripts of many genes encoding enzymes involved in oligo- and poly-saccharide biosynthesis (sucrose synthase, alpha-glucan-branching enzymes, and starch synthase) and major endosperm proteins, represented by low-molecular-weight glutenin subunit (Supplementary Data Set 6).

To further extend the understanding of transcriptional regulation of seed development, we performed de novo motif enrichment analysis within WGCNA modules. Collections of de novo motifs from each WGCNA module were curated for false positive, low complexity, and simple repeat motifs, resulting in a list of 168 motifs identified in embryo, endosperm, and SMTs (Supplementary Data Set 12). The major portion of de novo motifs were observed in early and late modules, providing insight into regulatory motifs driving the process of seed development.

Barley endosperm differentiation is initiated before cellularization

Several gene markers for individual endosperm domains have been identified across cereals (Kalla et al. 1994; Opsahl-Ferstad et al. 1997; Hueros et al. 1999; Gómez et al. 2002; Magnard et al. 2003; Bate et al. 2004), including a few markers for ETCs and AL (Leah et al. 1991; Kalla et al. 1994; Doan et al. 1996; Hertig et al. 2020, 2023). Here, we aimed to extend the list of markers for barley by performing a comparative analysis of 12 endosperm marker genes described in maize and rice (*Oryza sativa* L.). By reciprocal BLAST, we identified in total 29 homologs in barley (Supplementary Data Set 13). Interestingly, these known marker genes typically reached a maximum expression at younger (4 or 8 DAP) stages (Fig. 4, Supplementary Fig. S9, and Supplementary Data Set 13), suggesting a biased selection. Nevertheless, several candidates might be used for identifying later (16 or 24 DAP) stages of endosperm development.

The timing of starch accumulation in endosperm differs across cereals. It begins around 6 DAP in barley and only later, around 10 DAP, in maize (BenNET et al. 1975; Charlton et al. 1995). The CSE marker *STARCH BRANCHING ENZYME 1* has three gene copies in barley (*SBE1A-C*), where *SBE1A* and *1B* were first detected at 8 DAP and then decreased their expression, while *SBE1C* started low at 8 DAP and dramatically increased its expression (2,745 TPM) at 16 and 24 DAP (Fig. 4A). Rice *SUCROSE NON-FERMENTING-1 RELATED PROTEIN KINASE 1b* (*SnRK1b*) has four homologs in barley. Expression of *SnRK1b* in rice coincided with the emergence of starch granules (Kanegae et al. 2005). However, specific expression in barley endosperm was observed only for the *SnRK1bA* (Supplementary Fig. S9). Almost no expression was observed for the homolog of maize *SHRUNKEN 1* (*SH1*) encoding a sucrose synthase (Supplementary Fig. S9; Chourey and Nelson 1976). As to the aleurone, the barley *AL9A* and *AL9B* (homologs of maize *ALEURONE9*) were expressed exclusively 8 DAP. Other AL markers, such as barley homologs of rice subunit *B1A* of the *NUCLEAR FACTOR Y*

(*NF-YB1A*) and of maize *COLORED ALEURONE1* (*C1A*), reached their expression maxima at 16 and 24 DAP, respectively (Fig. 4, A and B and Supplementary Fig. S9).

We further analyzed the expression of *NF-YB1A* at 4–16 DAP using RNA in situ hybridization and found strong signals in the AL (Fig. 4C and Supplementary Fig. S9). Findings in maize suggest that AL differentiates from the outer layers of endosperm cells between 6 and 10 DAP soon after alveolation and the first periclinal division of the cellularized endosperm (Olsen 2001). Surprisingly, *NF-YB1A* was also expressed in 4 DAP endosperm, where the signals accumulated at the embryonic pole of the seed (Fig. 4C and Supplementary Fig. S9). This provided additional evidence supporting the recent observation that the aleurone identity is already defined in specific endosperm nuclei before cellularization begins (Hertig et al. 2023) and suggests that barley endosperm differentiation starts at the embryonic pole. The homolog of maize *OUTER CELL LAYER 4* (*OCL4*), encoding a homeo domain-leucine zipper IV TF expressed in AL, did not show endosperm-specific expression in barley (Sosso et al. 2010).

ETC markers, such as antimicrobial protein-coding genes *EMBRYO SAC/BASAL ENDOSPERM TRANSFER LAYER/ EMBRYO SURROUNDING REGION 1* putative paralogs A to D and 2 (*EBE1A-D* and *EBE2*) and TF *MYB-RELATED PROTEIN 1* putative paralogs A to C (*MRP1A-C*), were expressed at 4 DAP and reached expression maxima at 8 DAP (Fig. 4A and Supplementary Fig. S9). This indicates that transfer cell fate specification occurs during a narrow temporal window of coenocyte development in barley, which is similar to the situation described in maize (Costa et al. 2003; Hertig et al. 2023). ESR markers, such as sugar management coding gene *INVERTASE INHIBITOR 1A* (*INVINH1A*) and defensin-type gene *EMBRYO SURROUNDING REGION 6* (*ESR6*), reached their peak of expression at 4 DAP (Fig. 4A). Another class of ESR markers was represented by genes coding for putative antifungal proteins *ANDROGENIC EMBRYO 1A* to *B* (*ANE1A-B*) and *ANDROGENIC EMBRYO 3A* and *3B* (*ANE3A-B*) that were highly expressed 4 and 8 DAP, respectively (Fig. 4A and Supplementary Fig. S9). This corresponds to findings in maize, where ESR cells differentiate upon completion of the endosperm cellularization phase (Kiesselbach and Walker 1952; Serna et al. 2001).

Although the endosperm markers from various cereals also exhibited tissue-specificity in barley, their temporal expression dynamics differed in several cases. The AL, ETC, and ESR markers were already expressed at 4 and 8 DAP. It was previously reported that the specificity of AL and ETCs is already defined in their founder nuclei before the cellularization (Hertig et al. 2023). Here we extend this observation to the ESR.

Multiple pathways are controlled by H3K27me3 in barley endosperm

Our GO term enrichment analysis indicated chromatin-based regulation of barley seed development (Fig. 2B and

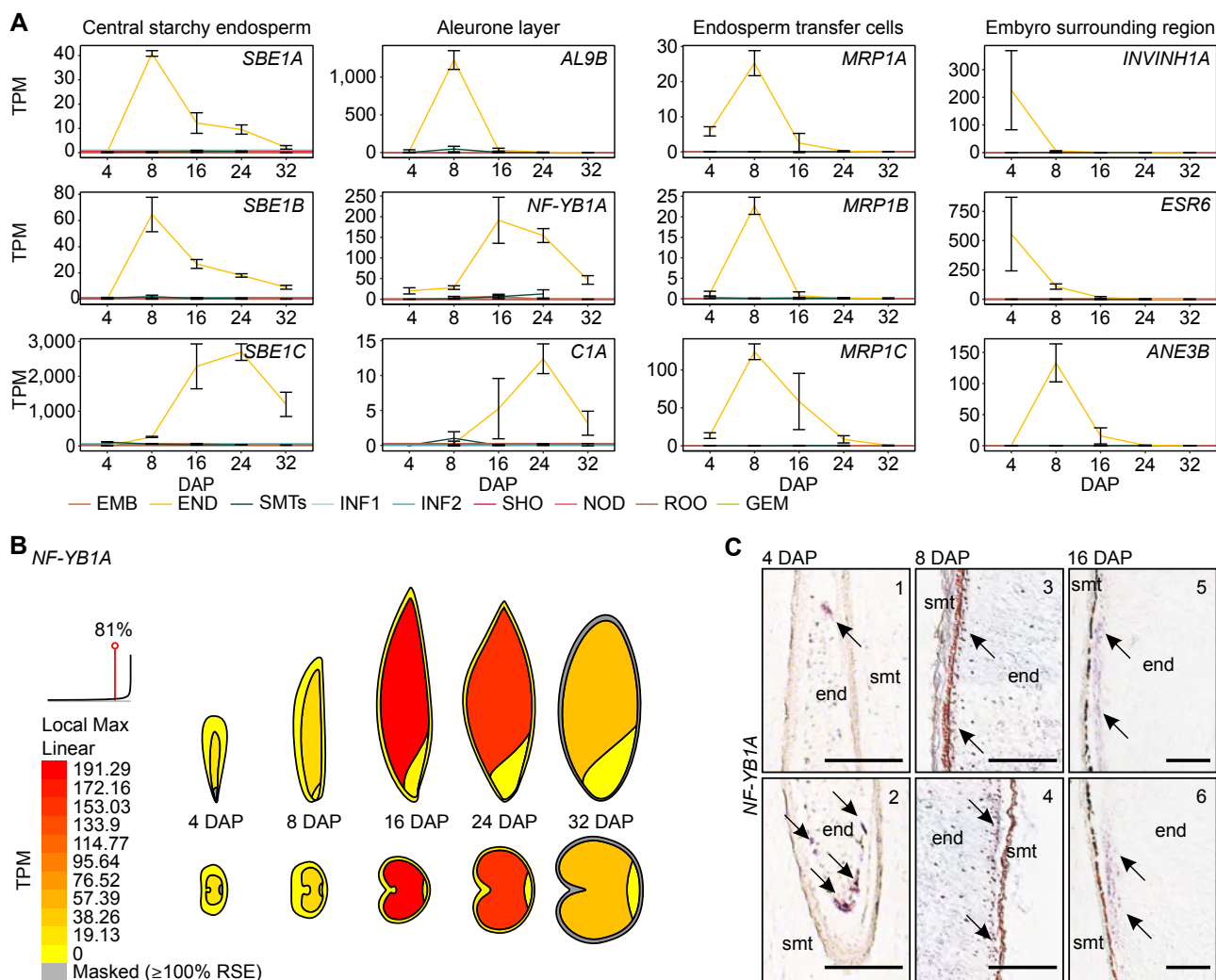


Figure 4. Expression of marker genes in endosperm domains in barley. **A**) Expression profiles of barley putative orthologs of selected endosperm marker genes in other cereals grouped according to the domain of expression—CSE, AL, ETCs, and ESR across different barley tissues (The International Barley Genome Sequencing Consortium 2012). TPM, transcript per million; EMB, embryo; END, endosperm; SMTs, seed maternal tissues; ROO, root; GEM, germinating embryo; NOD, nodule; SHO, shoot; INF1 and INF2, developing inflorescence 5 and 10 mm; CAR5 and CAR15, caryopsis 5 and 15 DAP. Error bars indicate standard deviation (the full list in [Supplementary Table S8](#) and [Supplementary Fig. S9](#)). **B**) Visualization *NF-YB1A* expression in different tissues of developing barley grain in the Barley ePlant. Gene expression in transcript per million (TPM) is indicated by the color scale. **C**) RNA in situ hybridization of *NF-YB1A* in barley grains at 4, 8, and 16 DAP using an antisense probe. emb, embryo; end, endosperm; SMT, seed maternal tissues. The full-scale images are presented in [Supplementary Fig. S9](#) and the numbers in the top right corners correspond to the inset numbers. Arrows indicate signal deposition in the aleurone region. Scale bars, 200 μ m.

[Supplementary Data Sets 3 and 4](#)). To explore this further, we studied the expression patterns of barley putative orthologs of histone and POLYCOMB REPRESSIVE COMPLEX 2 (PRC2) genes ([Supplementary Data Set 14](#)).

We found in total 188 barley histone genes corresponding to all canonical forms and common plant variants ([Supplementary Fig. S10](#) and [Supplementary Data Set 15](#)). In total, 152 out of 175 expressed histone genes (87%, TPM ≥ 1) were part of the endosperm k-means CL2 (including 30 copies of *H3*; 31 copies of *H2A.W*; and one *H2A.Z*), with predominant expression levels during the coenocyte stage ([Fig. 2C](#) and [Supplementary Data Set 5](#)). The peak of expression in CL2 coincided with the period of DNA replication and nuclei division during coenocyte

development. After cellularization, the expression of histone genes decreased, and only a few copies, mostly putative paralogs of non-canonical variants, remained expressed ([Supplementary Fig. S10](#) and [Supplementary Data Set 15](#)). The initial stages of embryo and SMTs were also marked by the peak of histone expression, but the overall maxima were lower compared to endosperm ([Supplementary Fig. S10](#)). A single *H2B* copy showed a prominent expression in 16 DAP and later embryo stages. Closer inspection of this copy revealed that it is a recently described seed-specific histone *H2B.S* variant ([Jiang et al. 2020](#)). The *H2B.S* was not expressed in endosperm and SMTs. Altogether, this indicates dynamic epigenetic control and changes in nucleosome composition during endosperm and embryo development.

The PRC2 complex catalyzes tri-methylation of lysine 27 at histone H3 (H3K27me3). This modification transcriptionally represses its target genes and thus contributes to developmental transitions, including endosperm cellularization (Mozgova and Hennig 2015). Arabidopsis PRC2 consists of four evolutionary conserved subunits: FERTILIZATION INDEPENDENT ENDOSPERM (FIE), MULTICOPY SUPPRESSOR OF IRA 1 (MSI1), EMBRYONIC FLOWER 2 (EMF2), and the catalytic subunit represented by the three homologs SET DOMAIN GROUP 1, 5 and 10 (SDG1, 5, and 10) alias CURLY LEAF (CLF), MEDEA (MEA), and SWINGER (SWN), respectively. All four core PRC2 subunits are present in cereals (Strejčková et al. 2020). We found that the barley genome lacks an *SDG5/MEA* homolog and contains single copies of *FIE* and *SDG10* (*SWN*) and multiple copies of *MSI1*, *EMF2*, and *SDG1/CLF*. At least one copy of each PRC2 subunit was well expressed (TPM > 10) throughout the whole embryo and SMTs development (Fig. 5A and Supplementary Data Set 15). The pattern in the endosperm was more complex. Genes coding for *EMF2A/B*, *SDG1A/B*, and *SWN* were silent or only weakly expressed at 4 DAP, suggesting no or little function of the PRC2 complex in barley early endosperm development. From 8 DAP onwards, *EMF2A*, *FIE*, *MSI1B*, and *SWN* maintained moderate expression levels, while both *SDG1* copies remained silent or weakly expressed.

To estimate whether these expression patterns have any impact on the H3K27me3 levels, we performed immunofluorescence in situ hybridization (immunofISH) staining on 8 DAP and 24 DAP endosperm and 24 DAP embryo nuclei isolated by flow-cytometry-based on their C-values (Fig. 5B). There was an intense H3K27me3 signal at the telomeric poles of all analyzed types of embryo nuclei. Such signal distribution is caused by the accumulation of H3K27me3 at the gene-rich chromosome ends and the presence of Rab1 chromosome organization in barley (Baker et al. 2015; Nowicka et al. 2023). On the contrary, the H3K27me3 signals were weaker at 8 DAP endosperm nuclei and almost absent at 24 DAP (Fig. 5B and Supplementary Fig. S11). The H3K27me3 loss was also progressive over increasing C-values, suggesting a developmental stage and ploidy-dependent global loss of H3K27me3 in endosperm nuclei.

To investigate H3K27me3 levels in endosperm at greater resolution, we performed ChIP-seq on 16 DAP endosperm samples and compared the signals to H3K27me3 ChIP-seq data from seedlings (Baker et al. 2015; Supplementary Figs. S12 and S13A). The H3K27me3 signals concentrated on gene-rich chromosome termini in seedlings but were strongly reduced in endosperm. Although the overall trend was biased toward H3K27me3 reduction in endosperm, the changes were more complex when looking at individual H3K27me3 peaks (Supplementary Fig. S13, B–D). The total number of H3K27me3 peaks was higher in the endosperm, but they were generally shorter and smaller compared to seedlings. We focused on the most prominent peaks (fold enrichment ≥ 5) and performed differential analysis.

We found 17,194 regions with a significant H3K27me3 loss and 13,845 regions with gain in endosperm relative to seedling ($\log_2\text{FoldChange} < 0$ or > 0 , respectively, Benjamini–Hochberg FDR-adjusted P -value < 0.05 ; Fig. 5D and Supplementary Data Set 16). The genomic regions that lost or gained H3K27me3 peaks included 1,856 and 1,118 genes, respectively (Fig. 5E and Supplementary Data Set 17). To assess the direct role of H3K27me3 in transcriptional regulation, we searched which of these genes were significantly up-regulated or down-regulated ($\text{padj} < 0.05$) in at least one analyzed endosperm time point (Fig. 5F, Supplementary Fig. S13E, and Supplementary Data Sets 18 and 19). For 60 genes, H3K27me3 –depletion in endosperm coincided with their increased expression. This included several storage compound-associated genes, such as *LOW-MOLECULAR-WEIGHT GLUTENIN SUBUNITs*, *C-HORDEIN*, and *OMEGA SECALIN* (Fig. 5F and Supplementary Data Set 18). Interestingly, other such genes included inhibitors of sugar and protein degradation and were expressed in a seed stage-dependent manner (Supplementary Data Set 18). The INVERTASE INHIBITORS block hydrolysis of sucrose to fructose and glucose, and were highly expressed at 8 DAP but not during subsequent stages, possibly allowing feeding of the growing embryo or endosperm. The inhibitor of protein degradation *SERPIN* was expressed from 16 DAP (Supplementary Data Set 18). This suggests that the accumulation of storage carbohydrates and proteins is accompanied by simultaneous inhibition of their degradation in the endosperm tissues in an H3K27me3-dependent manner. This was supported also by the enrichment of the storage protein-related GO terms for these 60 genes (Fig. 5G and Supplementary Table S2).

There were 238 genes that showed a significant decrease in expression of at least one analyzed endosperm time point, coinciding with enrichment of H3K27me3 (Supplementary Fig. S13E and Supplementary Data Set 19). This cluster was dominated by the two prominent groups. The first included putative *SENESCENCE-ASSOCIATED PROTEINs* ($n = 72$) which might be regulating tissue maturation by inhibiting specific proteases (Roberts et al. 2012). The second cluster included 27 copies of the RNA Pol II subunit *MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 12* (*MED12*; Supplementary Data Set 19). *MED12* could be linked with transcriptional regulation of specific genes in barley. In Arabidopsis, *MED12* contributes to the regulation of flowering genes and the mutants show late-flowering phenotype (Stewart Gillmor et al. 2014). Among enriched GO categories for these genes, we observed terms related to respiration and generation of energy (Supplementary Fig. S13F and Supplementary Table S3).

This suggests that grain filling, senescence, and several other biological pathways are controlled by the H3K27me3 modification in barley seeds.

Identification of conserved imprinted genes

The H3K27me3 plays an important role in the epigenetic regulation of uniparental gene expression by genomic

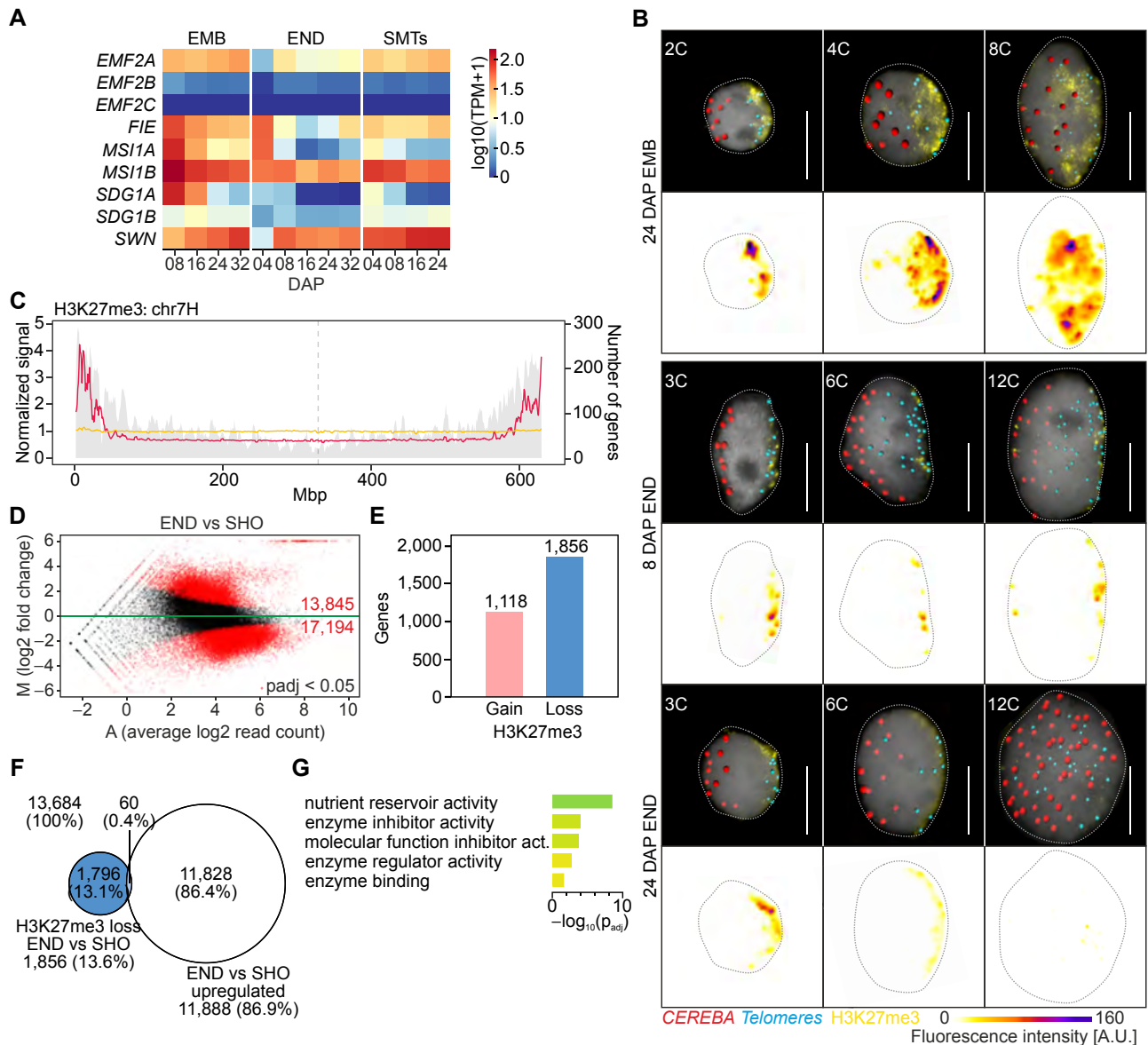


Figure 5. Dynamics of expression from PRC2 genes in barley grain tissues. **A)** Heatmap of expression from genes coding PRC2 complex subunits in the embryo (EMB), endosperm, and SMTs at different DAP. TPM, transcript per million (source data in [Supplementary Table S10](#)). **B)** Black background images show representative embryo and endosperm nuclei of different C-values collected at 8 and 24 DAP. DNA was stained with DAPI (gray), H3K27me3 was immunostained (yellow), and CEREB centromeric (red) and telomeric (blue) repeats were visualized by fluorescence in situ hybridization and signal segmentation. White background images show H3K27me3 immunostaining fluorescence signal intensities in arbitrary units (A.U.). Scale bar = 10 μm . Raw images and 3D image segmentation pictures of the nucleus surface, and immunostaining and fluorescence in situ hybridization (FISH) signals are presented in [Supplementary Fig. S11](#). **C)** Normalized signal abundance of H3K27me3 in the endosperm (yellow) and 10-cm whole seedling ([Baker et al. 2015](#); magenta). The gray background signal is gene density (secondary y-axis; the full list is provided in [Supplementary Fig. S12](#)) across chromosome 7H. **D)** MA plot showing genomic intervals with differential signal intensities between endosperm and 10-cm seedling (SHO). The M axis shows the log-transformed fold change and the A axis shows the average log-transformed read count. Intervals passing the threshold (Benjamini–Hochberg FDR-adjusted P -value < 0.05) are in red. Black dots indicate intervals which did not pass the threshold. The red numbers indicate genomic intervals with significant gain ($M > 0$) or loss of H3K27me3 ($M < 0$). Source data are provided in [Supplementary Data Set 7](#). **E)** Number of genes corresponding to genomic regions with significant gain and loss of H3K27me3 (full list provided in [Supplementary Data Set 8](#)). **F)** Venn diagram showing the number of genes with loss of H3K27me3 in 10-cm seedlings (SHO; based on ChIP-seq data from [Baker et al. 2015](#)) and genes up-regulated at 8, 16, 24, or 32 DAP endosperm. **G)** GO term enrichment of genes with loss of H3K27me3 and significant upregulation in the endosperm (source data in [Supplementary Table S13](#)).

imprinting in Arabidopsis (Batista and Köhler 2020). Only eight imprinted genes have been identified in barley to date, based on the homology with imprinted genes in rice (Chen et al. 2018). We took an analogous approach with a broader dataset of 155, 156, and 697 imprinted genes from rice, maize, and wheat (*Triticum aestivum* L.), respectively (Luo et al. 2011; Waters et al. 2011; Zhang et al. 2011; Chen et al. 2018; Yang et al. 2018), and performed a comparative search for evolutionarily conserved imprinted genes in barley endosperm tissues. We identified 449 barley putative orthologs (Fig. 6A and Supplementary Data Set 20) with 19 genes shared across two species and two genes shared across three species (together 4.3% of all candidates). Such a low number could indicate a relatively fast evolution of imprinted genes in grasses.

To provide experimental validation of the 21 conserved candidates, we checked their expression patterns in our seed transcriptomic data and identified four main expression pattern groups (Fig. 6B and Supplementary Data Set 21). Group 1 contained nine genes that were expressed across all seed tissues. Group 2 had four genes expressed in embryo and SMTs but weakly expressed or silent in the endosperm. Group 3 consisted of seven genes with expression restricted to endosperm, and group 4 contained a single gene expressed at early endosperm, 8 DAP SMTs, and silent in the embryo. Next, we analyzed parent-of-origin specific expression pattern of these candidates using 8 DAP endosperms of reciprocal hybrids between Morex × HOR 12560 (HS) genotypes. Three genes were excluded due to the lack of diagnostic single nucleotide polymorphisms (SNPs) in this parental combination (Fig. 6, C and D, and Supplementary Fig. S14). Amplification of cDNA from three genes—*DEFECTIVE ENDOSPERM 18* (*DE18*), *PALADIN* (*PALD*), and *PROTEIN PHOSPHATASE HOMOLOG 48* (*PRH48*)—was not successful and therefore these candidates remain unclassified.

The *FRIGIDA-ESSENTIAL 1* (*FES1*) gene showed an HS genotype-specific expression. Four genes *CYTOCHROME P450 REDUCTASE*, *PROTEIN KINASE FAMILY PROTEIN* (*PKP*), and *PROBABLE PROTEIN PHOSPHATASE 2C 27* (*PP2C27*) from group 1 were expressed from both parents. Three genes were classified as potentially imprinted because they showed either maternally (*MICRORCHIDIA 6B*, *MORC6B*) or paternally (*RNA-BINDING FAMILY PROTEIN*, *RBP* and *TRANSLOCASE SUBUNIT SECA*, *SECA*) biased expression. Seven of the total 18 tested genes (38.9%) were confirmed as imprinted in barley (Fig. 6D). Three maternally expressed genes (MEGs) included *CA-RESPONSIVE PROTEIN* (*CARP*), *RNA-BINDING PROTEIN* (*ARP1*), and *PROLINE-RICH PROTEIN* (*PRP3*). Wheat and rice MEG *CARP* was part of barley group 1 and was expressed only weakly in endosperm compared to moderate expression in embryos. *ARP1* and *PRP3* were both MEGs in rice and maize and their expression was relatively specific to barley endosperm.

We also confirmed four paternally expressed genes (PEGs) *REGULATION OF NUCLEAR PRE-MRNA DOMAIN-CONTAINING PROTEIN 1B* (*RPRD1B*), *DA1-RELATED 1* (*DAR1*),

AT-RICH INTERACTION DOMAIN 1B (*ARID1B*), and *UBIQUITIN-SPECIFIC PROTEASE 16* (*UBP16*). The *RPRD1B* and *ARID1B* genes were found as PEGs in rice and maize. *RPRD1Bs* encode epsin N-terminal homology domain-containing proteins with roles in endocytosis and cytoskeletal regulation (De Camilli et al. 2002). Arabidopsis *ARID1* is a transcriptional activator expressed in pollen development, which could be consistent with its role as a PEG in barley. *DAR1* was previously found as a PEG in wheat and maize (Fig. 6B). It is related to Arabidopsis *DA1*, a ubiquitin-activated peptidase that plays a role in the regulation of endoreduplication, determination of plant architecture, and possibly maternal control of seed weight (Peng et al. 2015; Gu et al. 2022). *DA1* functions antagonistically with its direct substrate *UBIQUITIN-SPECIFIC PROTEASE 15* (*UBP15*) in Arabidopsis (Du et al. 2014). Interestingly, we confirmed another member of the *UBIQUITIN CONJUGATING ENZYME* family, *UBP16* as an evolutionary conserved PEG in maize, rice, wheat (Luo et al. 2011; Waters et al. 2011; Chen et al. 2018; Yang et al. 2018), and barley (this study).

These findings encourage future genome-wide searches of imprinted genes in barley and provide an initial set of evolutionarily conserved candidates for functional studies.

Discussion

We generated a comprehensive gene expression atlas of developing barley seeds. This resource offers a higher resolution to the tissues and time points compared to previous studies (Zhang et al. 2004; Druka et al. 2006; Sreenivasulu et al. 2008). The stage- and tissue-specific marker genes, especially those for the later developmental stages, can serve as a basis for the follow-up functional investigations of crucial players involved in barley grain development. While we focused mostly on the analysis of endosperm tissues, the datasets have an equal resolution for embryos that give rise to the next generation and often neglected SMTs that play a critical role in protecting seeds and defining their parameters, e.g. size (Radchuk et al. 2011). It has to be emphasized that the embryo, endosperm, and SMTs were manually dissected and each tissue consisted of multiple cell types. Hence, our data show a tissue rather than cell type-specific transcriptional profiles. Manual dissection of spatially complex grain tissues can lead to potential tissue cross-contamination. We routinely monitored tissue purity by flow-cytometry-based ploidy measurements, and samples with signs of nuclei from other tissue were discarded. In addition, multiple wash steps were included due to the same ploidy level of embryo and SMTs. However, the embryo tissue was relatively solid and easy to separate and expected contamination between these two tissues should be minor ($\leq 5\%$). The transcriptomic data are easily accessible through the Barley ePlant under the V2 and V3 genome releases.

TFs play crucial roles in developmental transitions and control major maturation events, including storage reserve

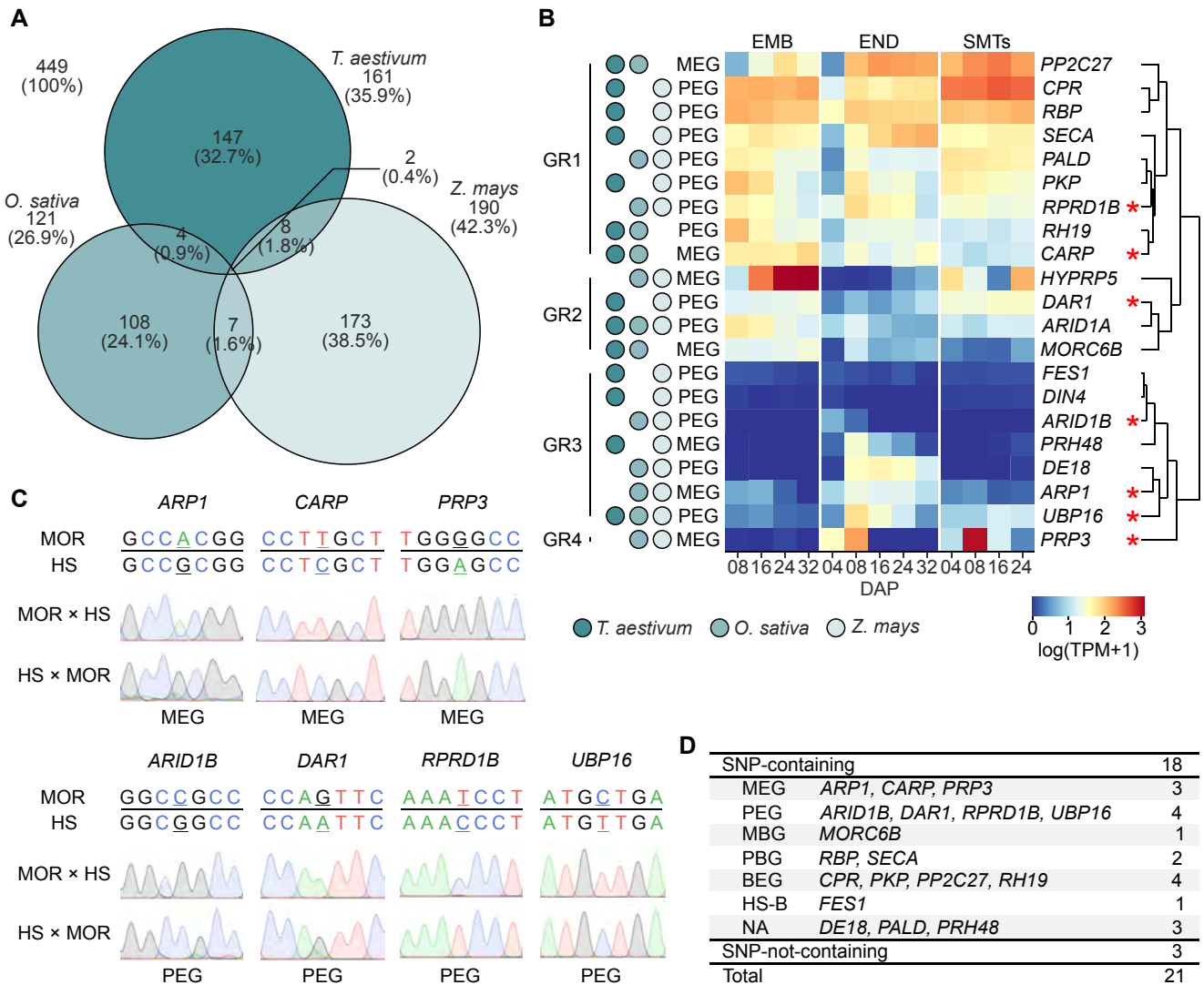


Figure 6. Identification and validation of imprinted genes in barley endosperm. **A**) Venn diagram showing the number of genes imprinted in single and multiple cereal species (source data in [Supplementary Table S15](#)). **B**) Heatmap of expression of barley putative orthologs of 21 genes imprinted in multiple cereal species in the embryo (EMB), endosperm, and SMTs at different DAP. Gene expression is indicated by the color scale. The heatmap is partitioned into groups based on different expression profiles in seed tissues. TPM, transcript per million; MEG, maternally expressed gene; PEG, paternally expressed gene; *, imprinted in barley (source data in [Supplementary Table S16](#)). **C**) Validation of SNPs for selected imprinted genes in 8 DAP endosperm tissues by Sanger sequencing. F1 reciprocal hybrids were obtained by crossing genetically distant cultivars Morex (MOR) and wild barley (HS). The maternal genotype is mentioned first in the crosses. The informative SNPs are underlined. **D**) Summary of validation of imprinted genes in barley. MBG, maternally biased gene; PBG, paternally biased gene; BEG, biallelic expressed gene; HS-B, wild barley biased; NA, information not available.

accumulation, chlorophyll degradation, and dormancy ([Alizadeh et al. 2021](#)). A detailed description of the transcriptional regulation exists for the embryo. We analyzed expression of TFs facilitating whole seed development, considering both TF expression and the presence of its binding motifs in promoter regions of the expressed genes. These two views often support and complement each other. We highlighted the case of DOF TFs and the corresponding expression of seed storage proteins. However, understanding the regulatory network comprising endosperm development will require further investigation of individual genes.

Using information from other cereals, we defined putative orthologs for individual markers of endosperm compartments in barley ([Olsen 2001](#)). Most of them showed endosperm-specific expression, suggesting that these markers will be useful also in barley. Many of the marker genes had expression maxima at 4 or 8 DAP, suggesting an early endosperm differentiation. Observations in maize suggested endosperm differentiation upon completion of endosperm cellularization. However, a recent study from barley ([Hertig et al. 2023](#)) and our results already exhibit the expression of the aleurone marker before cellularization. This is reminiscent

of the situation in *Arabidopsis*, where endosperm differentiation has already started at the 16-nuclei syncytial stage, and when cellularization is initiated around the embryo (Brown et al. 2003).

Studies in *Arabidopsis* revealed important roles of several epigenetic processes and molecular factors in embryo and endosperm development (Le et al. 2010). The degree to which these processes are conserved in cereals, including barley, remains unknown. We performed several pilot experiments that defined expression patterns of two important groups of epigenetic regulators—histones as the basic constituents of nucleosomes and members of the PRC2 complex, the key repressor of developmentally regulated genes (Zhang et al. 2002; Probst et al. 2020). The histone genes were expressed mostly at early stages of embryo and endosperm development, which coincides with a rapid round of DNA replication and cell division. Notably different was the *H2B.S* variant (Jiang et al. 2020) that was expressed in the late stages of embryo development. This indicates conserved function of *H2B.S* across angiosperms. Low expression of several PRC2 subunit genes at the early stages of barley endosperm development was surprising and suggested possible dynamics in the H3K27me3.

By a combination of transcriptomics and chromatin profiling, we also defined a set of endosperm genes that are directly controlled by H3K27me3. These include genes encoding storage components and their modifying enzymes such as *LOW-MOLECULAR-WEIGHT GLUTENIN SUBUNITS*, *C-HORDEIN* and *OMEGA SECALIN*, *SERPIN* and *INVERTASE INHIBITORS*. We hypothesize that the global reduction in H3K27me3 might not be sufficient for these transcriptional changes and it is likely that other mechanisms, possibly involving specific histone demethylases, are involved. However, our data point toward the pivotal role of chromatin dynamics in barley grain development suggests the need for further functional studies.

Finally, we laid down a fundament for future analysis of genomic imprinting in barley. A comparative approach identified 4.3% ($n = 21$) conserved (shared by at least two species) imprinted genes in maize, rice, and wheat. This is a surprisingly low number that points to the potential high speed of evolution of imprinted genes among cereal species. Experimental validation in barley confirmed the imprinted status of almost 40% of the conserved candidates. When combined with the previous study (Chen et al. 2018), we conclude that there are currently 15 confirmed imprinted genes in barley. Interestingly, the confirmed candidates point to the role of ubiquitin-based regulation in genomic imprinting. Proteasomal degradation of specific targets of imprinted genes could be a fine balancing mechanism between the maternal and paternal genome contributions. However, whole genome studies will be needed to uncover the full spectrum of barley imprinted genes.

In conclusion, our study generated valuable data for functional research on barley grain development and provided numerous unique resources that will enhance the capacity

of barley genomic research. Altogether, this will help in understanding the role of nucleus-governed processes during cereal grain development

Materials and methods

Plant materials and growth conditions

Six-rowed spring barley (*H. vulgare* subsp. *vulgare*) cv. Morex was used throughout the study. For analysis of imprinted genes, also wild barley (*H. vulgare* subsp. *spontaneum*) strain HOR 12560 (HS) was used. For plant growth, the seeds were germinated on damp cellulose tissue paper covered with one layer of filter paper for 3 days at 25 °C in the dark. Germinating grains were moved into 5 × 5 cm peat pots containing a mixture of soil and sand (3/1; v/v) and were grown in a climatic chamber under a long-day regime (16 h day 20 °C, 8 h night 16 °C; light intensity 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$; humidity 60%). After 10 days, the plants were transferred into 15 × 15 cm pots containing the same soil and cultivated in the same conditions until flowering. The number of DAPs was defined by determining the day of self-pollination as described (Kovacik et al. 2020). The developing seeds were collected at 4, 8, 16, 24, and 32 DAP in at least three biological replicates. Only the central row seeds from the middle of the spikelets were used. Corresponding tissues were manually dissected and checked for its purity by flow-cytometric ploidy measurement as described (Kovacik et al. 2020). The 4 DAP embryo was omitted due to its small size. For sampling 8 DAP embryos, at least 10 embryos were pooled per biological replicate. For embryo samples at later stages, five embryos were pooled per biological replicate. Young (4 and 8 DAP) and older endosperm (>8 DAP) tissues were isolated from 20 or three seeds, respectively, per biological replicate as described (Kovacik et al. 2020). SMTs were isolated from at least five grains per biological replicate, irrespective of the stage of development. SMTs at 32 DAP were excluded due to their dry nature. Dissected tissues were immediately frozen in liquid nitrogen and stored at −80 °C until use.

RNA extraction, sequencing, and analysis

Total RNA was isolated using an RNeasy Plant Mini Kit (QIAGEN) and Spectrum Plant Total RNA Kit (Sigma-Aldrich) according to the manufacturer's instructions including on-column DNase I (Sigma-Aldrich) treatment. The RNA quality was checked using Bioanalyzer 2100 with RNA 6000 Pico Chips (all Agilent). Samples with RNA integrity number >6.8 were processed into RNA-Seq mRNA libraries using NEBNext Ultra RNA Library Prep Kit for Illumina with poly-A selection. The mRNA-enriched libraries were sequenced as single-end 100 bp RNA-seq reads on a HiSeq2500 instrument (Illumina). The raw reads were trimmed for adaptors by Trim Galore v.0.4.1 (Martin 2011) and aligned to the *H. vulgare* cv. Morex reference genome v3 (Mascher 2021) using HiSat2 v.2.1.0 genomic mapper

(Kim et al. 2019). Aligned reads were assigned to features and meta-features using Subread v.1.5.2 software (Liao et al. 2013) according to the genome annotation. Differential expression analysis was performed using DESeq2 v.1.24.0 package (Love et al. 2014) in R v.3.6.3 software (R Core Team 2020). A gene was declared to be significantly differentially expressed if the Benjamini–Hochberg FDR-adjusted *P*-value was <0.05. Published barley RNA-Seq supplemental data sets were retrieved at the National Center for Biotechnology Information sequence read archive from Bioproject PRJEB3149 and analyzed as described above. The PCAs were done after applying the variance stabilization transformation (Anders and Huber 2010). Venn diagrams were drawn using the eulerr v.6.1.0 package in R (Larsson 2020).

Clustering analyses

For k-means clustering, unique DEGs from all tested combinations within the tissue were clustered using the k-means algorithm in R (R Core Team 2020). K-means clustering was performed using the Hartigan–Wong algorithm with 1,000 iterations. An optimal number of clusters was determined by statistical testing using a gap statistics method (Tibshirani et al. 2001). For WGCNA a gene co-expression network was constructed for each tissue with the raw read counts after the *rlog* transformation using the WGCNA library in R (Langfelder and Horvath 2008, 2012). An adjacency matrix was made using the soft thresholded Pearson correlation (power = 18) among 5,000 genes with the most varying expression among experimental points. Hierarchical clustering was performed for grouping the genes with highly similar co-expression patterns. The modules were identified using the dynamic hybrid tree cut algorithm (Langfelder et al. 2008) keeping the minimum size of the module to 15 and DeepSplit set to 4 to produce fine clusters. Each module was represented by color coding, with 12–15 modules detected depending on the tissue. The expression profile of each module was summarized by module eigengene defined as its first principal component. The probes that did not fit any of the main modules were placed into the “unspecific” module that was removed from further analysis.

A seed view in barley ePlant

The barley ePlant framework (Thiel et al. 2021) was modified to accept V3 barley gene identifiers (Mascher 2021). To create a new seed view in the barley ePlant, the data described in this publication in TPM were databased on the BAR for Plant Biology's server (Toufighi et al. 2005). An Scalable Vector Graphics image depicting the parts of the seed that were sampled in this work was generated and an Extensible Markup Language file linking specific parts of the image with database sample names was manually created to configure this new view.

GO term enrichment analyses and annotation of TFs

GO terms provided in Morex reference genome v3 annotation (Mascher 2021) were used to perform an enrichment

analysis by the topGO v.2.44.0 package in R (Alexa and Rahnenfuhrer 2021). Redundant GO terms were filtered using the web-based tool Reduce + Visualize Gene Ontology (Supek et al. 2011) with default settings and general terms were filtered using size selection (Yon Rhee et al. 2008). Terms with a size sufficient for robust statistical analysis ($n > 100$) and fold enrichment >3 were investigated. TFs were classified into families (Supplementary Data Set 22) based on the presence of specific domains according to PlantTFDB (Jin et al. 2017).

Cis-motif identification and clustering

For cis-motif identification and enrichment analysis, 1,500 bp upstream sequences from the predicted start codon (ATG) of all WGCNA module genes were used. The analysis was carried out using the “findMotif.pl” program from the HOMER suite (Heinz et al. 2010) that performs known and de novo motif identification and enrichment analysis with default parameters. The enrichment of identified motifs was calculated respective to all 1,500 bp background sequences. Collections of identified motifs in each WGCNA module were post-filtered for plant motifs and clustered using the Regulatory Sequence Analysis Tools (RSAT) (Castro-Mondragon et al. 2017) with default parameters.

RNA in situ hybridization

Barley seeds were harvested at 4, 8, and 16 DAP in at least three biological replicates, fixed with 4% freshly prepared paraformaldehyde (w/v), with 2% Tween-20 (v/v), and 2% Triton X-100 (pH 7, adjusted by HCl, v/v) for 1 h under vacuum. For increased fixation, efficiency vacuum was broken every 10 min and applied again. Subsequently, seeds were transferred into fresh fixatives and stored overnight at 4 °C, dehydrated using ethanol series, cleared by ROTIHistol series, and embedded into Paraplast. Longitudinal dorsoventral sections of 10 µm were cut with a Reichert-Jung 2030 microtome and attached to Adhesion Slides Superfrost Ultra Plus (Thermo Fisher Scientific). DNA templates for synthesis of RNA probes were amplified from cDNA (reverse transcribed by RevertAid H Minus First Strand cDNA Synthesis Kit; Thermo Fisher Scientific) by PCR. Sense and antisense digoxigenin (DIG)-labeled RNA probes were amplified using gene-specific primers containing T7 promoter sequences (Supplementary Table S4) and DIG-UTP (Thermo Fisher Scientific) according to the manufacturer's instructions. After purification, the efficiency of DIG labeling was verified by a modified dot blotting hybridization (Zöllner et al. 2021). For hybridization, slides with tissue sections were washed in ROTIHistol, rehydrated, and treated with 0.2 M HCl for 10 min, pronase (0.125 mg mL⁻¹) for 10 min, 0.2% glycine (v/v) for 2 min, 4% formaldehyde (v/v) for 10 min, and acetic anhydride (1% in 0.1 M Triethanolamine, pH 8.0, v/v). Hybridizations with denatured probes were carried out at 50 °C using the hybridization buffer containing 100 µL 10× salts, 400 µL deionized formamide, 200 µL 50% dextran sulfate

(w/v), 10 μ L of yeast (*Saccharomyces cerevisiae* Meyen ex E.C. Hansen) tRNA (100 mg mL⁻¹), 20 μ L 50 $\times$ Denhardt's Solution (Thermo Fisher Scientific), and 70 μ L dH₂O. After washing, unbound RNA was digested for 30 min at 37 °C using RNase A (20 μ g mL⁻¹). Immunological detection using DIG antibodies (1:3,000 in blocking solution) coupled with alkaline phosphatase and staining procedure with 4-nitro blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl-phosphate was done for 24–36 h at room temperature in dark. Hybridization signal analysis was performed using a light microscope BX60 (Olympus).

Identification of chromatin genes in barley

Identification was done using a subset of 47 Arabidopsis (*A. thaliana* (L.) Heynh.) genes encoding histones and 12 encoding PRC2 complex subunits. Homology searches were performed using BLAST+ (Camacho et al. 2009). The resulting hits were confirmed with reciprocal homology searching using the whole genome of 48,359 *A. thaliana* genes. The candidates were further filtered by standard BLAST+ *E*-value (≤ 0.01) and additional parameters counting with the comparison of (1) barley and *A. thaliana* gene length ($\geq 40\%$) and (2) alignment coverage of the hit ($\geq 40\%$).

ImmunoFISH, microscopy, and image analysis

ImmunoFISH was performed on flow-sorted nuclei from 24 DAP embryos and 8 and 24 DAP endosperm as described (Nowicka et al. 2023). Preparations were incubated with the rabbit anti-H3K27me3 primary antibody (1:200; Abcam, 195,477) at 4 °C overnight and the secondary goat anti-rabbit-Alexa Fluor 488 antibody (1:300, Molecular Probes, A11008) at 37 °C for 90 min. Barley centromeres were detected with a synthetic 28-mer oligonucleotide (5'-AGGGAGA-3')₄ CEREBA probe labeled at the 5' end with Cy3 (Eurofins). Telomeres were visualized with a synthetic 28-mer oligonucleotide probe (5'-CCCTAAA-3')₄ labeled at the 5 end with Cy5 (Eurofins).

The images were acquired with an AxioImager Z2 microscope (Zeiss, Oberkochen, Germany) equipped with a pE-4000 LED illuminator light source (CoolLED), laser-free confocal spinning disk device (DSD2, Andor, Belfast, UK), and with $\times 100/1.4$ NA Oil M27 Plan-Apochromat (Zeiss) objective. Illumination LED intensities (385 nm = 50%; 460, 565, and 635 nm = 75%), scanning speed, and pinhole (1.5 airy units) were kept identical for the image series in an experiment. Image stacks of 40–80 slides depending from the C-value of the nucleus, on average, with 0.2 μ m z-step were taken separately for each fluorochrome using the appropriate excitation (DAPI λ = 390/40 nm, GFP λ = 482/18 nm, RED FLUORESCENT PROTEIN (RFP) λ = 561/14 nm, Cy5 = 640/14 nm) and emission (DAPI λ = 452/45 nm, GFP λ = 525/45 nm, RFP λ = 609/54 nm, Cy5 = 676/29 nm) filters. The 4.2 MPx sCMOS camera (Zyla 4.2) and the iQ 3.6.1 acquisition software (both Andor) were used to drive the microscope and for fluorescence detection. Exposure times and gain were kept identical throughout all images

(exposure time: DAPI = 30 ms, GFP = 90 ms, RFP = 80 ms, Cy5 = 220 ms; gain all channels = 20). The images were saved as maximum intensity projection files with Imaris File Converter 9.2.1 (Bitplane, Zurich, Switzerland). Furthermore, Imaris 9.7 functions "Surface" and "Spots" were used for the nucleus surface, immunosignals, and centromere and telomere 3D modeling. Fluorescence intensity measurements were performed in Fiji using the "Interactive 3D Surface plot" plugin.

Analysis of imprinted genes

The lists of cereal imprinted genes were extracted from published works (Luo et al. 2011; Waters et al. 2011; Zhang et al. 2011; Chen et al. 2018; Yang et al. 2018) and their overlaps were visualized using Venn diagrams in R package eulerr (Larsson 2020). *H. vulgare* subsp. *spontaneum* accession HOR 12560 was grown as described (Nowicka et al. 2021) and synchronized for flowering with cv. Morex. The strains were reciprocally crossed, 8 DAP endosperm was manually dissected as described (Kovacik et al. 2020), total RNA was isolated using Spectrum Plant Total RNA Kit (Sigma-Aldrich), and reverse transcribed into cDNA using RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Primers were designed to amplify 200–1,100 bp fragments of the candidate genes harboring informative SNP(s) (Supplementary Table S4) using a standard PCR. The amplicons were gel purified using GeneJet Gel Extraction Kit (Thermo Fisher Scientific) and subjected to Sanger sequencing followed by in silico analysis using SnapGene v6.2 software (GSL Biotech LLC).

H3K27me3 ChIP-seq and data analysis

Approximately 2 g of 16 DAP endosperm tissue from cv. Morex were isolated in three biological replicates and cross-linked under vacuum in 1% (w/v) formaldehyde for 15 min at room temperature. The crosslinking was quenched by adding glycine to a final concentration of 0.1 M, and the vacuum was applied for 5 min. Endosperm tissue was rinsed with water and frozen in liquid nitrogen. ChIP was performed as described previously (Gendrel et al. 2005), with the following modifications. Briefly, isolated nuclei were resuspended in Nuclei lysis buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5% SDS, cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche)) and incubated at 4 °C under gentle agitation for 20 min. The chromatin solution was sonicated using a Biorupter Plus (Diagenode) with 10 cycles of 30 s pulse/90 s cooling at high power at 4 °C. The resulting sheared chromatin was pooled and diluted 1:4 with ChIP dilution buffer (16.7 mM Tris-HCl pH 8.0, 167 mM NaCl, 1.25 mM EDTA, 1.25% Triton X-100, 1 $\times$ cOmplete, EDTA-free Protease Inhibitor Cocktail). Six hundred microliter aliquots of diluted chromatin were incubated with 7 μ L of the rabbit anti-trimethyl-Histone H3 (Lys27) antibody (Millipore, 07-449) in a rotator at 4 °C overnight. Samples without antibody were used as negative controls. The next day, 40 μ L of the Dynabeads Protein A (Invitrogen) were added to each

tube, and the samples were further incubated for 2 h. Afterward, beads were washed with a sequence of buffers, and immune complexes were eluted as described. Control chromatin aliquots (“input DNA”) were taken prior to immunoprecipitation. Reverse crosslinking was performed for all samples, and DNA was extracted and purified using the ChIP DNA Clean & Concentrator kit (Zymo Research). Sequencing libraries were prepared and 150 bp pair-end reads were sequenced using Illumina NovaSeq 6000 platform (Illumina) by Novogene.

The raw reads of samples sequenced in our study and also publicly available H3K27me3 ChIP-seq data from shoot (Baker et al. 2015) were trimmed for adaptors and aligned to the Morex reference genome, duplicates were removed using MarkDuplicates in genome analysis toolkit (Van der Auwera and O'Connor 2020), and the peak calling was performed using MACS2 (Zhang et al. 2008). The peaks were filtered (fold enrichment ≥ 5) and tested for differential signal intensity between endosperm and shoot samples using R package MAnorm2 (Tu et al. 2021). Testing was performed in genomic intervals of size 2,000 bp and intervals with differential signal intensities localized in coding regions or 2,000 bp upstream were related to genes. GO term enrichment analysis was performed by web-based tool g:Profiler (Raudvere et al. 2019) using barley GO annotation for Morex V3 available at Ensembl Plants.

Accession numbers

Unique identifiers for all genes and their products mentioned in the text are provided in [Supplementary Data Set 23](#). All data supporting this study are included in the article and its [Supplementary material](#). Sequence data from this article can be found in the Gene Expression Omnibus under accession numbers GSE233316 and GSE238237.

Acknowledgments

We thank H. Tvardíková for technical assistance, Z. Bursová for plant care, and P. Navrátil for IT support.

Author contributions

M.Ko. performed sample preparation and RNA extraction, data analysis, and interpretation. M.Ko., I.V., and M.Č. performed RNA in situ hybridization. A.N. performed immunostaining, fluorescence in situ hybridization, microscopy, image analysis, and barley crosses. J.Z. performed the identification of chromatin genes. B.S. performed ChIP and contributed to data analysis. N.J.P., E.E., and A.Pa. prepared a seed view in barley ePlant. K.K. and M.Kr. performed validation of imprinted genes. R.S. and J.Š. supervised parts of the project. A.Pe. conceived, designed, and supervised the project. A.Pe. and M.Ko. wrote the manuscript. All authors have seen and agreed upon the final version of the manuscript.

Supplementary data

The following materials are available in the online version of this article.

Supplementary Figure S1. Sample distance heatmap.

Supplementary Figure S2. PCA of seed transcriptomes.

Supplementary Figure S3. Numbers of tissue-specific and shared genes among seed tissues, expressed at different levels.

Supplementary Figure S4. Volcano plots of DEGs in consecutive developmental transitions.

Supplementary Figure S5. K-means co-expression clusters of DEGs.

Supplementary Figure S6. WGCNA corresponding modules and their hierarchical clustering.

Supplementary Figure S7. TF MC 2 analysis.

Supplementary Figure S8. RNA in situ hybridization of PBF gene in barley grains at 4, 8, and 16 DAP.

Supplementary Figure S9. Expression profiles of selected putative barley orthologs of endosperm marker genes.

Supplementary Figure S10. Heatmap showing expression patterns of 175 histone genes in embryo (EMB), endosperm, and SMTs at different DAP.

Supplementary Figure S11. Cytology analysis of H3K27me3.

Supplementary Figure S12. Abundance of H3K27me3.

Supplementary Figure S13. Comparison of H3K27me3 profiles between endosperm and shoot.

Supplementary Figure S14. Validation of candidates for imprinted genes.

Supplementary Table S1. GO term category enrichment based on the subset of genes with TPM > 1,000 in endosperm tissues.

Supplementary Table S2. GO term enrichment for up-regulated and H3K27me3 significantly reduced genes in 16 DAP endosperm compared to shoot.

Supplementary Table S3. GO term enrichment for down-regulated and H3K27me3 significantly reduced genes in 16 DAP endosperm compared to shoot.

Supplementary Table S4. Primers used in this study.

Supplementary Data Set 1. RNA-seq raw read counts for embryo (EMB), endosperm, and SMTs.

Supplementary Data Set 2. Expression changes between individual subsequent DAP in embryo (EMB), endosperm, and SMTs.

Supplementary Data Set 3. GO term enrichment for down-regulated DEGs.

Supplementary Data Set 4. GO term enrichment for up-regulated DEGs.

Supplementary Data Set 5. Assignment of DEGs from embryo (EMB), endosperm, and SMTs to individual k-means clusters (K).

Supplementary Data Set 6. Gene assignment to WGCNA modules in embryo (EMB), endosperm, and SMTs.

Supplementary Data Set 7. Identified plant motifs within WGCNA modules in embryo, endosperm, and SMTs.

Supplementary Data Set 8. List of motifs assigned to different MCs.

Supplementary Data Set 9. Contribution of each motif collection to the MCs identified by RSAT.

Supplementary Data Set 10. Barley genes homologous to Arabidopsis genes with binding motifs assigned to MC 2.

Supplementary Data Set 11. Barley genes homologous to Arabidopsis genes with binding motifs assigned to MC 9.

Supplementary Data Set 12. Identified de novo plant motifs within WGCNA modules in embryo, endosperm, and SMTs.

Supplementary Data Set 13. Homologies of known rice and maize endosperm domains marker genes with barley genes.

Supplementary Data Set 14. Barley histone and PRC2 genes.

Supplementary Data Set 15. Seed transcriptomic data for barley histone and PRC2 genes.

Supplementary Data Set 16. List of genomic intervals with significantly different H3K27me3 signal intensities in endosperm compared to shoot (SHO) in ChIP-seq.

Supplementary Data Set 17. List of genes with the differential signal intensity of H3K27me3 in endosperm compared to shoot (SHO).

Supplementary Data Set 18. Up-regulated genes with significantly reduced H3K27me3 in 16 DAP endosperm compared to shoot (SHO).

Supplementary Data Set 19. Down-regulated genes with significantly reduced H3K27me3 in 16 DAP endosperm compared to shoot (SHO).

Supplementary Data Set 20. Homology of barley genes to imprinted genes identified in maize, rice, and wheat.

Supplementary Data Set 21. Seed transcriptomic data for barley homologs of imprinted genes identified in at least two other cereals.

Supplementary Data Set 22. List of TFs and their assignments to families.

Supplementary Data Set 23. Gene names and IDs used in the text.

Funding

Computational resources were supplied by the project “e-Infrastruktura CZ” (e-INFRA CZ LM2018140) supported by the Ministry of Education Youth and Sports of the Czech Republic. This work was supported by the Czech Science Foundation grant 21-02929S (A.P.) and German Research Foundation grants EXC2048 (R.S.) and IRTG2466 (I.V.).

Conflict of interest statement. The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

Alexa A, Rahnenfuhrer J. topGO: enrichment analysis for gene ontology. R package version 2.54.0. <https://doi.org/10.18129/B9.bioc.topGO>

- Alizadeh M, Hoy R, Lu B, Song L. Team effort: combinatorial control of seed maturation by transcription factors. *Curr Opin Plant Biol.* 2021;**63**:102091. <https://doi.org/10.1016/j.pbi.2021.102091>
- Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010;**11**(10):1–12. <https://doi.org/10.1186/gb-2010-11-10-r106>
- Baker K, Dhillon T, Colas I, Cook N, Milne I, Milne L, Bayer M, Flavell AJ. Chromatin state analysis of the barley epigenome reveals a higher-order structure defined by H3K27me1 and H3K27me3 abundance. *Plant J.* 2015;**84**(1):111–124. <https://doi.org/10.1111/tpj.12963>
- Bate NJ, Niu X, Wang Y, Reimann KS, Helentjaris TG. An invertase inhibitor from maize localizes to the embryo surrounding region during early kernel development. *Plant Physiol.* 2004;**134**(1):246–254. <https://doi.org/10.1104/pp.103.027466>
- Batista RA, Köhler C. Genomic imprinting in plants-revisiting existing models. *Genes Dev.* 2020;**34**(1–2):24–36. <https://doi.org/10.1101/gad.332924.119>
- Bennet MD, Smith JB, Barclay I. Early seed development in the triticeae. *Philos Trans R Soc Lond B Biol Sci.* 1975;**272**(916):199–227. <https://doi.org/10.1098/rstb.1975.0083>
- Bian J, Deng P, Zhan H, Wu X, Nishantha M, Yan Z, Du X, Nie X, Song W. Transcriptional dynamics of grain development in barley (*Hordeum vulgare* L.). *Int J Mol Sci.* 2019;**20**(4):962. <https://doi.org/10.3390/ijms20040962>
- Brown RC, Lemmon BE, Nguyen H. Events during the first four rounds of mitosis establish three developmental domains in the syncytial endosperm of *Arabidopsis thaliana*. *Protoplasma.* 2003;**222**(3–4):167–174. <https://doi.org/10.1007/s00709-003-0010-x>
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;**10**(1):1–9. <https://doi.org/10.1186/1471-2105-10-421>
- Carena MJ, ed. Cereals. New York (NY): Springer US; 2009.
- Castro-Mondragon JA, Jaeger S, Thieffry D, Thomas-Chollier M, van Helden J. RSAT matrix-clustering: dynamic exploration and redundancy reduction of transcription factor binding motif collections. *Nucleic Acids Res.* 2017;**45**(13):e119–e119. <https://doi.org/10.1093/nar/gkx314>
- Charlton WL, Keen CL, Merriman C, Lynch P, Greenland AJ, Dickinson HG. Endosperm development in Zea mays; implication of gametic imprinting and paternal excess in regulation of transfer layer development. *Development.* 1995;**121**(9):3089–3097. <https://doi.org/10.1242/dev.121.9.3089>
- Chen C, Li T, Zhu S, Liu Z, Shi Z, Zheng X, Chen R, Huang J, Shen Y, Luo S, et al. Characterization of imprinted genes in rice reveals conservation of regulation and imprinting with other plant species. *Plant Physiol.* 2018;**177**(4):1754–1771. <https://doi.org/10.1104/pp.17.01621>
- Chourey PS, Nelson OE. The enzymatic deficiency conditioned by the shrunken-1 mutations in maize. *Biochem Genet.* 1976;**14**(11–12):1041–1055. <https://doi.org/10.1007/BF00485135>
- Cook F, Hughes A, Nibau C, Orman-Ligeza B, Schatrowski N, Uauy C, Trafford K. Barley lys3 mutants are unique amongst shrunken-endosperm mutants in having abnormally large embryos. *J Cereal Sci.* 2018;**82**:16–24. <https://doi.org/10.1016/j.jcs.2018.04.013>
- Costa LM, Gutierrez-Marcos JF, Brutnell TP, Greenland AJ, Dickinson HG. The globby1-1 (glo1-1) mutation disrupts nuclear and cell division in the developing maize seed causing alterations in endosperm cell fate and tissue differentiation. *Development.* 2003;**130**(20):5009–5017. <https://doi.org/10.1242/dev.00692>
- De Camilli P, Chen H, Hyman J, Panepucci E, Bateman A, Brunger AT. The ENTH domain. *FEBS Lett.* 2002;**513**(1):11–18. [https://doi.org/10.1016/S0014-5793\(01\)03306-3](https://doi.org/10.1016/S0014-5793(01)03306-3)
- Doan DNR, Linnestad C, Olsen A. Isolation of molecular markers from the barley endosperm coenocyte and the surrounding nucellus cell layers. *Plant Mol Biol.* 1996;**31**(4):877–886. <https://doi.org/10.1007/BF00019474>

- Druka A, Muehlbauer G, Druka I, Caldo R, Baumann U, Rostoks N, Schreiber A, Wise R, Close T, Kleinhofs A, et al.** An atlas of gene expression from seed to seed through barley development. *Funct Integr Genomics*. 2006;**6**(3):202–211. <https://doi.org/10.1007/s10142-006-0025-4>
- Du L, Li N, Chen L, Xu Y, Li Y, Zhang Y, Li C, Li Y.** The ubiquitin receptor DA1 regulates seed and organ size by modulating the stability of the ubiquitin-specific protease UBP15/SOD2 in Arabidopsis. *Plant Cell*. 2014;**26**(2):665–677. <https://doi.org/10.1105/tpc.114.122663>
- Food and Agriculture Organization of the United Nations.** <https://www.fao.org/home/en/2023>
- Gendrel AV, Lippman Z, Martienssen R, Colot V.** Profiling histone modification patterns in plants using genomic tiling microarrays. *Nat Methods*. 2005;**2**(3):213–218. <https://doi.org/10.1038/nmeth0305-213>
- Gómez E, Royo J, Guo Y, Thompson R, Hueros G.** Establishment of cereal endosperm expression domains identification and properties of a maize transfer cell-specific transcription factor, ZmMRP-1. *Plant Cell*. 2002;**14**(3):599–610. <https://doi.org/10.1105/tpc.010365>
- Goyal K, Walton LJ, Tunnacliffe A.** LEA proteins prevent protein aggregation due to water stress. *Biochem J*. 2005;**388**(1):151–157. <https://doi.org/10.1042/BJ20041931>
- Gu B, Dong H, Smith C, Cui G, Li Y, Bevan MW.** Modulation of receptor-like transmembrane kinase 1 nuclear localization by DA1 peptidases in Arabidopsis. *Proc Natl Acad Sci U S A*. 2022;**119**(40):e2205757119. <https://doi.org/10.1073/pnas.2205757119>
- Gubatz S, Dercksen VJ, Brütz C, Weschke W, Wobus U.** Analysis of barley (*Hordeum vulgare*) grain development using three-dimensional digital models. *Plant J*. 2007;**52**(4):779–790. <https://doi.org/10.1111/j.1365-3113X.2007.03260.x>
- Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK.** Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;**38**(4):576. <https://doi.org/10.1016/j.molcel.2010.05.004>
- Hertig C, Melzer M, Rutten T, Erbe S, Hensel G, Kumlehn J, Weschke W, Weber H, Thiel J.** Barley HISTIDINE KINASE 1 (HvHK1) coordinates transfer cell specification in the young endosperm. *Plant J*. 2020;**103**(5):1869–1884. <https://doi.org/10.1111/tpj.14875>
- Hertig C, Rutten T, Melzer M, Schippers JHM, Thiel J.** Dissection of developmental programs and regulatory modules directing endosperm transfer cell and aleurone identity in the syncytial endosperm of barley. *Plants*. 2023;**12**(8):1594. <https://doi.org/10.3390/plants12081594>
- Hueros G, Royo J, Maitz M, Salamini F, Thompson RD.** Evidence for factors regulating transfer cell-specific expression in maize endosperm. *Plant Mol Biol*. 1999;**41**(3):403–414. <https://doi.org/10.1023/A:1006331707605>
- Jiang D, Borg M, Lorković ZJ, Montgomery SA, Osakabe A, Yelagandula R, Axelsson E, Berger F.** The evolution and functional divergence of the histone H2B family in plants. *PLoS Genet*. 2020;**16**(7):e1008964. <https://doi.org/10.1371/journal.pgen.1008964>
- Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G.** PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res*. 2017;**45**(D1):D1040–D1045. <https://doi.org/10.1093/nar/gkw982>
- Kalla R, Shimamoto K, Potter R, Nielsen PS, Linnestad C, Olsen O-A.** The promoter of the barley aleurone-specific gene encoding a putative 7 kDa lipid transfer protein confers aleurone cell-specific expression in transgenic rice. *Plant J*. 1994;**6**(6):849–860. <https://doi.org/10.1046/j.1365-3113X.1994.6060849.x>
- Kanegae H, Miyoshi K, Hirose T, Tsuchimoto S, Mori M, Nagato Y, Takano M.** Expressions of rice sucrose non-fermenting-1 related protein kinase 1 genes are differently regulated during the caryopsis development. *Plant Physiol Biochem*. 2005;**43**(7):669–679. <https://doi.org/10.1016/j.plaphy.2005.06.004>
- Kiesselbach TA, Walker ER.** Structure of certain specialized tissue in the kernel of corn. *Am J Bot*. 1952;**39**(8):561–569. <https://doi.org/10.1002/j.1537-2197.1952.tb13069.x>
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL.** Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;**37**(8):907–915. <https://doi.org/10.1038/s41587-019-0201-4>
- Kovacik M, Nowicka A, Pecinka A.** Isolation of high purity tissues from developing barley seeds. *J Vis Exp*. 2020(164):e61681. <https://doi.org/10.3791/61681>
- Langfelder P, Horvath S.** WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;**9**(1):559. <https://doi.org/10.1186/1471-2105-9-559>
- Langfelder P, Horvath S.** Fast {R} functions for robust correlations and hierarchical clustering. *J Stat Softw*. 2012;**46**(11):1–17. <https://doi.org/10.18637/jss.v046.i11>
- Langfelder P, Zhang B, Horvath S.** Defining clusters from a hierarchical cluster tree: the dynamic tree cut package for R. *Bioinformatics*. 2008;**24**(5):719–720. <https://doi.org/10.1093/bioinformatics/btm563>
- Larsson, J.** (2020). eulerr: area-proportional Euler and Venn diagrams with ellipses.
- Le BH, Cheng C, Bui AQ, Wagmaster JA, Henry KF, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Horvath S, et al.** Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proc Natl Acad Sci U S A*. 2010;**107**(18):8063–8070. <https://doi.org/10.1073/pnas.1003530107>
- Leah R, Tommerup H, Svendsen I, Mundy J.** Biochemical and molecular characterization of three barley seed proteins with antifungal properties. *J Biol Chem*. 1991;**266**(3):1564–1573. [https://doi.org/10.1016/S0021-9258\(18\)52331-0](https://doi.org/10.1016/S0021-9258(18)52331-0)
- Liao Y, Smyth GK, Shi W.** The subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. 2013;**41**(10):e108–e108. <https://doi.org/10.1093/nar/gkt214>
- Love MI, Huber W, Anders S.** Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;**15**(12):550. <https://doi.org/10.1186/s13059-014-0550-8>
- Luo M, Taylor JM, Spriggs A, Zhang H, Wu X, Russell S, Singh M, Koltunow A.** A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. *PLoS Genet*. 2011;**7**(6):e1002125. <https://doi.org/10.1371/journal.pgen.1002125>
- Magnard JL, Lehouque G, Massonneau A, Frangne N, Heckel T, Gutierrez-Marcos JF, Perez P, Dumas C, Rogowsky PM.** ZmEBE genes show a novel, continuous expression pattern in the central cell before fertilization and in specific domains of the resulting endosperm after fertilization. *Plant Mol Biol*. 2003;**53**(6):821–836. <https://doi.org/10.1023/B:PLAN.0000023672.37089.00>
- Martin M.** Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;**17**(1):10–12. <https://doi.org/10.14806/ej.17.1.200>
- Mascher M.** Pseudomolecules and annotation of the third version of the reference genome sequence assembly of barley cv. Morex [Morex V3], 2021. <http://doi:10.5447/ipk/2021/3>
- Mena M, Vicente-Carbajosa J, Schmidt RJ, Carbonero P.** An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. *Plant J*. 1998;**16**(1):53–62. <https://doi.org/10.1046/j.1365-3113x.1998.00275.x>
- Mozgova I, Hennig L.** The polycarb group protein regulatory network. *Annu Rev Plant Biol*. 2015;**66**(1):269–296. <https://doi.org/10.1146/annurev-arplant-043014-115627>
- Nowicka A, Ferková I, Said M, Kovacik M, Zwyrtková J, Baroux C, Pecinka A.** Non-Rabl chromosome organization in endoreduplicated nuclei of barley embryo and endosperm tissues. *J Exp Bot*. 2023;**74**(8):2527–2541. <https://doi.org/10.1093/jxb/erad036>
- Nowicka A, Sahu PP, Kovacik M, Weigt D, Tokarz B, Krugman T, Pecinka A.** Endopolyploidy variation in wild barley seeds across environmental gradients in Israel. *Genes (Basel)*. 2021;**12**(5):711. <https://doi.org/10.3390/genes12050711>

- Olsen O-A. ENDOSPERM DEVELOPMENT: cellularization and cell fate specification. *Annu Rev Plant Biol.* 2001;**52**(1):233–267. <https://doi.org/10.1146/annurev.arplant.52.1.233>
- Opsahl-Ferstad HG, Le Deunff E, Dumas C, Rogowsky PM. Zmesr, a novel endosperm-specific gene expressed in a restricted region around the maize embryo. *Plant J.* 1997;**12**(1):235–246. <https://doi.org/10.1046/j.1365-313X.1997.12010235.x>
- Peirats-Llobet M, Yi C, Liew LC, Berkowitz O, Narsai R, Lewsey MG, Whelan J. Spatially resolved transcriptomic analysis of the germinating barley grain. *Nucleic Acids Res.* 2023;**51**(15):7798–7819. <https://doi.org/10.1093/nar/gkad521>
- Peng Y, Chen L, Lu Y, Wu Y, Dumenil J, Zhu Z, Bevan MW, Lia Y. The ubiquitin receptors DA1, DAR1, and DAR2 redundantly regulate endoreduplication by modulating the stability of TCP14/15 in Arabidopsis. *Plant Cell.* 2015;**27**(3):649–662. <https://doi.org/10.1105/tpc.114.132274>
- Probst AV, Desvoves B, Gutierrez C. Similar yet critically different: the distribution, dynamics and function of histone variants. *J Exp Bot.* 2020;**71**(17):5191–5204. <https://doi.org/10.1093/jxb/eraa230>
- R Core Team. <https://www.r-project.org/2020>
- Radchuk VV, Borisjuk L, Sreenivasulu N, Merx K, Mock HP, Rolletschek H, Wobus U, Weschke W. Spatiotemporal profiling of starch biosynthesis and degradation in the developing barley grain. *Plant Physiol.* 2009;**150**(1):190–204. <https://doi.org/10.1104/pp.108.133520>
- Radchuk V, Weier D, Radchuk R, Weschke W, Weber H. Development of maternal seed tissue in barley is mediated by regulated cell expansion and cell disintegration and coordinated with endosperm growth. *J Exp Bot.* 2011;**62**(3):1217–1227. <https://doi.org/10.1093/jxb/erq348>
- Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. G:profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 2019;**47**(W1):W191–W198. <https://doi.org/10.1093/nar/gkz369>
- Roberts IN, Caputo C, Criado MV, Funk C. Senescence-associated proteases in plants. *Physiol Plant.* 2012;**145**(1):130–139. <https://doi.org/10.1111/j.1399-3054.2012.01574.x>
- Sano N, Rajjou L, North HM. Lost in translation: physiological roles of stored mRNAs in seed germination. *Plants (Basel).* 2020;**9**(3):347–361. <https://doi.org/10.3390/plants9030347>
- Serna A, Maitz M, O'Connell T, Santandrea G, Thevissen K, Tienens K, Hueros G, Faleri C, Cai G, Lottspeich F, et al. Maize endosperm secretes a novel antifungal protein into adjacent maternal tissue. *Plant J.* 2001;**25**(6):687–698. <https://doi.org/10.1046/j.1365-313x.2001.01004.x>
- Sosso D, Wisniewski JP, Khaled AS, Hueros G, Gerentes D, Paul W, Rogowsky PM. The vpp1, Esr6a, Esr6b and OCL4 promoters are active in distinct domains of maize endosperm. *Plant Sci.* 2010;**179**(1–2):86–96. <https://doi.org/10.1016/j.plantsci.2010.04.006>
- Sreenivasulu N, Usadel B, Winter A, Radchuk V, Scholz U, Stein N, Weschke W, Strickert M, Close TJ, Stitt M, et al. Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant Physiol.* 2008;**146**(4):1738–1758. <https://doi.org/10.1104/pp.107.111781>
- Stewart Gillmor C, Silva-Ortega CO, Willmann MR, Buendía-Monreal M, Poethig RS. The Arabidopsis mediator CDK8 module genes CCT (MED12) and GCT (MED13) are global regulators of developmental phase transitions. *Development.* 2014;**141**(23):4580–4589. <https://doi.org/10.1242/dev.111229>
- Strejčková B, Čegan R, Pecinka A, Milec Z, Šafář J. Identification of polycomb repressive complex 1 and 2 core components in hexaploid bread wheat. *BMC Plant Biol.* 2020;**20**(S1):1–13. <https://doi.org/10.1186/s12870-020-02384-6>
- Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One.* 2011;**6**(7):e21800. <https://doi.org/10.1371/journal.pone.0021800>
- The International Barley Genome Sequencing Consortium. A physical, genetic and functional sequence assembly of the barley genome. *Nature.* 2012;**491**(7426):711–716. <https://doi.org/10.1038/nature11543>
- Thiel J, Koppolu R, Trautewig C, Hertig C, Kale SM, Erbe S, Mascher M, Himmelbach A, Rutten T, Esteban E, et al. Transcriptional landscapes of floral meristems in barley. *Sci Adv.* 2021;**7**(18):eabf0832. <https://doi.org/10.1126/sciadv.abf0832>
- Tibshirani R, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J R Stat Soc Ser B Stat Methodol.* 2001;**63**(2):411–423. <https://doi.org/10.1111/1467-9868.00293>
- Toufighi K, Brady SM, Austin R, Ly E, Provart NJ. The botany array resource: e-northern, expression angling, and promoter analyses. *Plant J.* 2005;**43**(1):153–163. <https://doi.org/10.1111/j.1365-313X.2005.02437.x>
- Tu S, Li M, Chen H, Tan F, Xu J, Waxman DJ, Zhang Y, Shao Z. MAAnorm2 for quantitatively comparing groups of ChIP-seq samples. *Genome Res.* 2021;**31**(1):131–145. <https://doi.org/10.1101/gr.262675.120>
- Van der Auwera GA, O'Connor BD. Genomics in the cloud: using Docker, GATK, and WDL in Terra. Sebastopol (CA): O'Reilly Media; 2020.
- Vishal B, Kumar PP. Regulation of seed germination and abiotic stresses by gibberellins and abscisic acid. *Front Plant Sci.* 2018;**9**:838. <https://doi.org/10.3389/fpls.2018.00838>
- Waters AJ, Makarevitch I, Eichten SR, Swanson-Wagner RA, Yeh CT, Xu W, Schnable PS, Vaughn MW, Gehring M, Springer NM. Parent-of-origin effects on gene expression and DNA methylation in the maize endosperm. *Plant Cell.* 2011;**23**(12):4221–4233. <https://doi.org/10.1105/tpc.111.092668>
- Weschke W, Panitz R, Gubatz S, Wang Q, Radchuk R, Weber H, Wobus U. The role of invertases and hexose transporters in controlling sugar ratios in maternal and filial tissues of barley caryopses during early development. *Plant J.* 2003;**33**(2):395–411. <https://doi.org/10.1046/j.1365-313X.2003.01633.x>
- Yanagisawa S, Sheen J. Involvement of maize dof zinc finger proteins in tissue-specific and light-regulated gene expression. *Plant Cell.* 1998;**10**(1):75–89. <https://doi.org/10.1105/tpc.10.1.75>
- Yang G, Liu Z, Gao L, Yu K, Feng M, Yao Y, Peng H, Hu Z, Sun Q, Ni Z, et al. Genomic imprinting was evolutionarily conserved during wheat polyploidization. *Plant Cell.* 2018;**30**(1):37–47. <https://doi.org/10.1105/tpc.17.00837>
- Yon Rhee S, Wood V, Dolinski K, Draghici S. Use and misuse of the gene ontology annotations. *Nat Rev Genet.* 2008;**9**(7):509–515. <https://doi.org/10.1038/nrg2363>
- Zhang S, Laurie AE, Ae W, Meng L, Lemaux PG. Similarity of expression patterns of knotted1 and ZmLEC1 during somatic and zygotic embryogenesis in maize (*Zea mays* L.). *Planta.* 2002;**215**(2):191–194. <https://doi.org/10.1007/s00425-002-0735-3>
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 2008;**9**(9):1–9. <https://doi.org/10.1186/gb-2008-9-9-r137>
- Zhang H, Sreenivasulu N, Weschke W, Stein N, Rudd S, Radchuk V, Potokina E, Scholz U, Schweizer P, Zierold U, et al. Large-scale analysis of the barley transcriptome based on expressed sequence tags. *Plant J.* 2004;**40**(2):276–290. <https://doi.org/10.1111/j.1365-313X.2004.02209.x>
- Zhang M, Zhao H, Xie S, Chen J, Xu Y, Wang K, Zhao H, Guan H, Hu X, Jiao Y, et al. Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. *Proc Natl Acad Sci U S A.* 2011;**108**(50):20042–20047. <https://doi.org/10.1073/pnas.1112186108>
- Zöllner NR, Bezruczyk M, Laureyns R, Nelissen H, Simon R, Frommer WB. An RNA in situ hybridization protocol optimized for monocot tissue. *STAR Protoc.* 2021;**2**(2):100398. <https://doi.org/10.1016/j.xpro.2021.100398>