



# Spatial and temporal regulation of parent-of-origin allelic expression in the endosperm

Yuri S. van Ekelburg <sup>1</sup>, Karina S. Hornslien <sup>1,\*</sup>, Tom Van Hautegeem <sup>2,3</sup>, Matyáš Fendrych <sup>2,3,†</sup>,  
Gert Van Isterdael <sup>4,5</sup>, Katrine N. Bjerkan <sup>1</sup>, Jason R. Miller <sup>6</sup>, Moritz K. Nowack <sup>2,3</sup>  
and Paul E. Grini <sup>1,\*</sup>

- 1 Section for Genetics and Evolutionary Biology, Department of Biosciences, University of Oslo, 0316 Oslo, Norway
- 2 Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium
- 3 VIB Center for Plant Systems Biology, 9052 Ghent, Belgium
- 4 VIB Flow Core, VIB Center for Inflammation Research, Ghent, Belgium
- 5 Department of Biomedical Molecular Biology, Ghent University, Ghent, Belgium
- 6 College of STEM, Shepherd University, Shepherdstown, West Virginia 25443-5000, USA

\*Authors for correspondence: E-mail: paul.grini@ibv.uio.no (P.E.G.), k.s.hornslien@ibv.uio.no (K.S.H.)

†Present address: Department of Experimental Plant Biology, Charles University, Prague, Czech Republic.

P.E.G., J.R.M., and M.K.N. designed the research; Y.S.v.E., T.v.H., M.F., K.N.B., G.v.I., M.N.K., and P.E.G. performed the experiments; Y.S.v.E., K.S.H., T.v.H., M.F., J.R.M., M.N.K., and P.E.G. analyzed and discussed the data; Y.S.v.E., K.S.H., and P.E.G. wrote the article. All authors revised and approved the article.

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/plphys/pages/General-Instructions>) is: Paul E. Grini.

## Abstract

Genomic imprinting promotes differential expression of parental alleles in the endosperm of flowering plants and is regulated by epigenetic modification such as DNA methylation and histone tail modifications in chromatin. After fertilization, the endosperm develops through a syncytial stage before it cellularizes and becomes a nutrient source for the growing embryo. Regional compartmentalization has been shown both in early and late endosperm development, and different transcriptional domains suggest divergent spatial and temporal regional functions. The analysis of the role of parent-of-origin allelic expression in the endosperm as a whole and the investigation of domain-specific functions have been hampered by the inaccessibility of the tissue for high-throughput transcriptome analyses and contamination from surrounding tissue. Here, we used fluorescence-activated nuclear sorting (FANS) of nuclear targeted GFP fluorescent genetic markers to capture parental-specific allelic expression from different developmental stages and specific endosperm domains. This approach allowed us to successfully identify differential genomic imprinting with temporal and spatial resolution. We used a systematic approach to report temporal regulation of imprinted genes in the endosperm, as well as region-specific imprinting in endosperm domains. Analysis of our data identified loci that are spatially differentially imprinted in one domain of the endosperm, while biparentally expressed in other domains. These findings suggest that the regulation of genomic imprinting is dynamic and challenge the canonical mechanisms for genomic imprinting.

Received March 8, 2022. Accepted October 17, 2022. Advance access publication November 28, 2022

© The Author(s) 2022. Published by Oxford University Press on behalf of American Society of Plant Biologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

Open Access

## Introduction

The double fertilization event in angiosperms forms the diploid zygote and triploid endosperm, respectively (Strasburger, 1900; Nowack et al., 2010). Endosperm development is characterized by three different stages, a syncytial stage, a cellular stage, and a maturation stage, having different functions (Berger et al., 2006; Li and Berger, 2012). In *Arabidopsis* (*Arabidopsis thaliana*) and other dicots, the endosperm does not persevere at maturation and the proliferating embryo accumulates reserves on the expense of the endosperm. After fertilization, the syncytial endosperm goes through rapid mitotic cycles without cytokinesis, and obtains nutrients from the mother plant. Upon cellularization in many dicots, hexose stored in the central vacuole becomes the main source of nutrients for the growing embryo, a process often referred to as the sink–source switch (Nowack et al., 2010; Lafon-Placette and Köhler, 2014). Following cellularization, cell types differentiate by mitosis, dividing the cellular endosperm into various subregions (Brown et al., 1999; Boissard-Lorig et al., 2001; Li and Berger, 2012). It has become evident that endosperm domains have distinct expression profiles (Belmonte et al., 2013; Del Toro-De León and Köhler, 2019; Picard et al., 2021) and results indicating designated functions for each domain are emerging (Ingram, 2010).

The anterior embryo surrounding region (ESR) (Opsahl-Ferstad et al., 1997) is important for molecular communication between the embryo and the endosperm (Yang et al., 2008; Doll et al., 2020). Furthermore, this region is required for cuticle formation, a hydrophobic barrier between the embryo and endosperm that protects the embryo/seedling from water loss (Tanaka et al., 2001). As the domain closest to the embryo, the ESR undergoes programmed cell death shortly after cellularization in order to provide stored nutrients to the embryo (Denay et al., 2014; Van Hautegeem et al., 2015).

The developing aleurone layer (DAL), also referred to as peripheral endosperm, is the last surviving endosperm domain that remains up until germination (Van Hautegeem et al., 2015). DAL has been shown to be crucial to maintain dormancy, and cell-wall breakdown in this region coincides with germination, starting with root tip proximal cells (Bethke et al., 2007). Lastly, endosperm cells positioned between the ESR and DAL gradually become part of the ESR upon expansion of the embryo (Van Hautegeem et al., 2015).

The posterior chalazal endosperm remains separated from the rest of the endosperm in a syncytial state and consists of two multinucleate regions: the chalazal cyst and chalazal nodule (Olsen, 2004; Picard et al., 2021). Transcriptome analysis showed enrichment for genes involved in sucrose unloading, consistent with a nutrient transfer function, and enrichment for genes involved in tetrahydrofolate and folic acid biosynthesis (Picard et al., 2021).

The importance of endosperm domain integrity is exemplified by mutation of the *FIS POLYCOMB REPRESSIVE*

*COMPLEX 2* (*PRC2*), where seed lethality is accompanied by delayed or the absence of endosperm cellularization, disrupted mitotic domain organization, and an enlarged chalazal cyst and ectopic chalazal nodules (Chaudhury et al., 1997; Kiyosue et al., 1999; Ingouff et al., 2005). A similar effect on endosperm compartmentalization can be observed in paternal excess *Arabidopsis thaliana* interploidy hybrid seeds, and can also be phenocopied by DNA hypomethylation of the maternal cross partner (Scott et al., 1998; Adams et al., 2000; Xiao et al., 2006). These observations link the development of endosperm compartmentalization and the function of endosperm domains to parent-of-origin effects. Due to the homodiploid nature of the embryo-sac central cell, fertilization by a haploid sperm cell gives rise to an unbalanced endosperm in regards to the allelic contribution of the parents (2:1 maternal:paternal), and therefore tight regulation is required for balanced parental gene expression (Birchler, 1993). The endosperm, analogous to the mammalian placenta, is the main site for the epigenetic phenomenon genomic imprinting, parent-of-origin-dependent expression of genes due to epigenetic marks (Feil and Berger, 2007). Traditionally, this process involves partial or full silencing of one of the parental alleles and is associated with both DNA methylation and histone modification (Jullien et al., 2006; Satyaki and Gehring, 2017). Furthermore, the involvement of small RNAs through the RNA-directed DNA methylation (RdDM) pathway resulting in *de novo* DNA methylation has been demonstrated (Vu et al., 2013; Hornslien et al., 2019; Satyaki and Gehring, 2019; Batista and Köhler, 2020).

The functional role of genomic imprinting in flowering plants has not been fully understood as most mutants of imprinted genes do not show an obvious seed phenotype (Shirzadi et al., 2011; Wolff et al., 2015; Zhang et al., 2018; Bjerkan et al., 2020). Moreover, the number of maternally and paternally imprinted genes previously identified remains controversial due to the detection of widespread contamination from the maternal seed coat and diploid embryo tissue (Schon and Nodine, 2017). Therefore, it is crucial for the study of genomic imprinting to avoid contamination from genetically distinct surrounding tissues. Furthermore, it has been shown that parent-of-origin-specific expression can be specific to endosperm regions, as suggested by differences in parental expression bias between domains (Picard et al., 2021). Although some examples have been identified where imprinted gene expression is completely inactivated (Kirkbride et al., 2019), becomes increasingly biparental (Ngo et al., 2012) or is stage-specific (Xin et al., 2013), dynamic regulation of imprinting patterns in a spatial or temporal manner has not been systematically investigated for the *Arabidopsis thaliana* endosperm.

Here, we aim to enhance our understanding of the role of genomic imprinting in flowering plants by investigating spatio-temporal dynamics of parent-of-origin allelic expression. Since different gene regulatory programs operate in different endosperm domains, specific mechanisms to modulate temporal and spatial-specific imprinting may

have evolved. We hypothesize that genomic imprinting in the *A. thaliana* endosperm is dynamically regulated in both a temporal and spatial manner. In order to avoid contamination issues that complicated previous imprinting studies, we have analyzed transcriptomes from isolated nuclei of endosperm-specific domains. To this end, we have generated nuclear targeted eGFP fluorescent genetic markers that report different temporal developmental stages or have a spatial resolution in endosperm domains. Using Fluorescence-Activated Nuclear Sorting (FANS) captured nuclei, we have successfully identified imprinted genes with an imprinting profile specific to developmental stages as well as subdomains of the endosperm. Our findings suggest that regulation of genomic imprinting is dynamic, both in a spatial and temporal perspective, challenging canonical mechanisms for genomic imprinting.

## Results and discussion

### Generation of endosperm-specific temporal and spatial expression marker lines

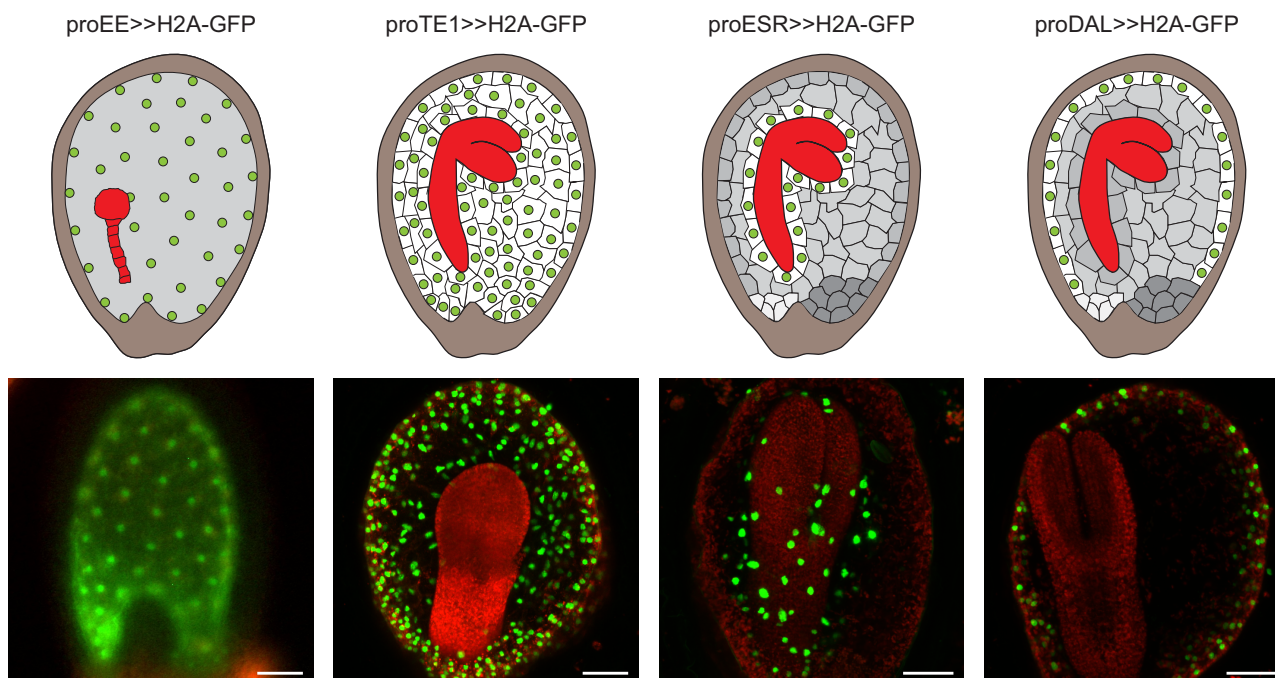
In order to investigate the temporal and spatial effects on parent-of-origin-specific allelic expression in *A. thaliana*, we developed endosperm-specific genetic reporters that had specific temporal and spatial profiles. A dual component system was used where nuclear-localized histone 2A-GFP fusion protein (H2A-GFP) is expressed under control of a domain-specific promoter (proMARKER>>H2A-GFP) (Weijers et al., 2003; Olvera-Carrillo et al., 2015). Based on seed tissue microarray data (Le et al., 2010) and on gene expression patterns (Winter et al., 2007) 39 candidate endosperm-specific gene promoters were selected, and cloned expression vectors were transformed into *A. thaliana* Col-0 (Supplemental Table S1). The spatial and temporal expression of these constructs was characterized during seed development in 48 independent T1 individuals per reporter line. For the purpose of FANS, four lines with domain or stage-specific markers were selected: AT5G09370 expressing in Early Endosperm (EE); AT3G45130 expressing in Embryo Surrounding Region (ESR); AT4G31060 expressing in DAL; and AT4G00220 expressing in Total Endosperm (TE1). For each line, the expression pattern was verified in subsequent generations by confocal imaging (Figure 1). No seed coat or embryo GFP signal was observed in any of these lines. To demonstrate the utilization of the marker lines to identify cellularized endosperm nuclei, we performed FANS using an optimized sorting protocol to separate GFP-positive nuclei from GFP-negative nuclei (Supplemental Figure S1A and B). The ploidy levels of GFP-positive nuclei corresponded to endosperm ploidies (3C,6C) compared with GFP-negative (2C, 4C and 8C) ploidies (Supplemental Figure S1C). Additionally, in order to verify the specificity of endosperm spatial markers, the level of GFP transcript from GFP-positive and GFP-negative ESR and DAL nuclei was assessed by reverse transcription quantitative PCR (RT-qPCR) (Supplemental Figure S1D), demonstrating a 50–

250-fold GFP transcript enrichment in GFP-positive nuclei compared with GFP-negative nuclei.

### FANS of endosperm nuclei

In order to isolate endosperm domain-specific transcript profiles for temporal and spatial genomic imprinting analysis, homozygous marker lines EE, ESR, DAL, and TE1 in the Col-0 accession were crossed as mothers to wild-type of the Tsu-1 accession. EE seeds were collected from dissected siliques at 4 days after pollination (DAP) and seeds from ESR, DAL, and TE1 were collected at 7 DAP (Supplemental Dataset S1). Seeds were homogenized to release nuclei and GFP-positive nuclei were sorted by FANS. Nuclear RNA was isolated and cDNA libraries were sequenced yielding 150 bp paired-end reads (Supplemental Dataset S1). Although nuclear RNA contains higher levels of pre-RNA and unspliced transcripts (Long et al., 2021), a high correlation between nuclear and total cellular mRNA has previously been demonstrated in plants, also including triploid endosperm tissue (Jacob et al., 2007; Deal and Henikoff, 2010; Del Toro-De León and Köhler, 2019; Picard et al., 2021). In order to increase mapping specificity of reads coming from Col-0 and Tsu-1, for all genes we empirically generated and polished target sequences of each ecotype with Pilon (Walker et al., 2014). Up to ten rounds of polishing, beyond which only minor, cyclic changes were observable, increased the Col-0/Tsu-1 separation. Interestingly, polishing also increased the overall map rate by about 1%, indicating that some reads had not initially mapped to Araport11 (Supplemental Dataset S1). Reads from sequenced marker lines showed an average of 25 million reads per replicate mapped to the Col-0 and Tsu-1 polished target sequences (Supplemental Dataset S1).

In order to verify that FANS generated transcriptomes were not contaminated with seed coat and embryo transcripts, we performed an analysis of transcripts previously identified to be more than eight-fold enriched in the seed coat or the embryo, compared with the endosperm (Schon and Nodine, 2017). We compared the expression of 137 seed coat- and 62 embryo-enriched genes in FANS expression profiles versus whole-seed Col-0 and Tsu-1 profiles generated, using five established reference genes as calibrators (Huang et al., 1997; Dekkers et al., 2011; Shirzadi et al., 2011). The whole-seed profiles were overall highly enriched for seed coat and embryo-specific gene transcripts compared with the endosperm-specific FANS generated transcriptomes (Supplemental Table S2A, Supplemental Dataset S2), demonstrating that seed coat and embryo transcript contamination can largely be excluded for the FANS generated transcriptomes. Depending on the calibrator used, all comparisons were enriched for seed coat and embryo transcripts in the whole-seed samples compared with FANS. In the case of using ACTIN11 as calibrator, more than 90% of the transcripts investigated were more than two-fold enriched and of these 55%–81% were more than eight-fold enriched for seed coat or embryo-specific transcripts in the whole-seed samples compared with FANS samples (Supplemental Table S2A).



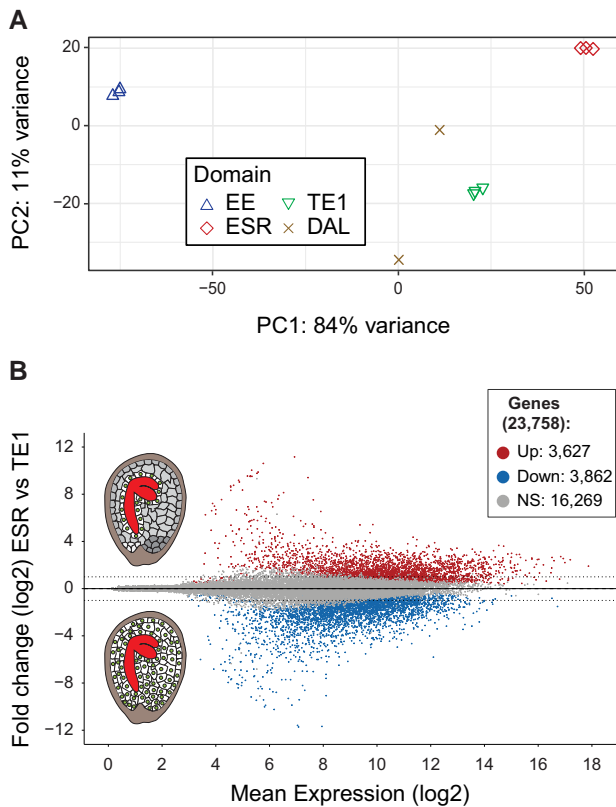
**Figure 1** Verification of endosperm domain-specific marker lines. Top panel, Diagrammatic representation of endosperm domains for the selected marker lines within the temporal and spatial plane. The GFP-expressing endosperm nuclei for the selected marker lines are highlighted in green. The embryo and seed coat are represented in red and brown, respectively. Endosperm that is not expressing GFP is depicted in gray. Bottom panel, Verification of domain-specific GFP expression in endosperm domain marker lines using fluorescence microscopy. Red indicates autofluorescence of chloroplasts in the embryo. Note that no embryo is visible in proEE>>H2A-GFP due to the use of a bandpass filter. proEE>>H2A-GFP = early endosperm; proTE1>>H2A-GFP = total endosperm; proESR>>H2A-GFP = embryo surrounding region; proDAL>>H2A-GFP = developing aleurone layer. Scale bar = 50  $\mu$ m.

To evaluate the variation between individual replicates of each marker line, reads were mapped to the Col-0 and Tsu-1 target sequences. Counts from the two mappings were subjected to differential expression analysis (Supplemental Dataset S3). Principal component analysis (PCA) of normalized read counts per gene explained 95% of the variances in the first two principal components. The PCA plot (Figure 2A) shows high between-group variation for all four lines, and low within-group variation for three lines. The analysis indicated high homogeneity of EE, TE1, and ESR replicates (Figure 2A), suggesting a high specificity of the sorting process. However, divergence in the DAL line (Figure 2A), combined with its insufficient replicates and read counts (Supplemental Dataset S1), indicated the DAL domain marker data should be omitted from further analyses. To further address the sensitivity of the marker-generated transcriptomes, the differential expression analysis was repeated excluding DAL and used to assess temporal (EE versus TE1), spatial (ESR versus TE1), and spatial–temporal (EE versus ESR) expression differences (Supplemental Dataset S3). As expected, due to comparison of different tissues or developmental time points, the differential expression was most pronounced when comparing both spatial and temporal expression (EE versus ESR—11,977 differentially expressed genes), followed by differential expression in 4 DAP versus 7 DAP stages (EE versus TE1—9,399 differentially expressed genes)

(Supplemental Figure S2A and B, respectively). Importantly, albeit the embryo surrounding region (ESR) is contained within the total endosperm (TE1), 7,489 genes were also differentially regulated comparing ESR with TE1 at 7 DAP (Figure 2B). This demonstrates that significant expression changes between total endosperm and the embryo surrounding region can be visualized using profiles from TE1 and ESR FANS. We next explored whether allelic parent-of-origin-specific expression could be addressed with these data.

### Informative read calling identifies imprinted genes

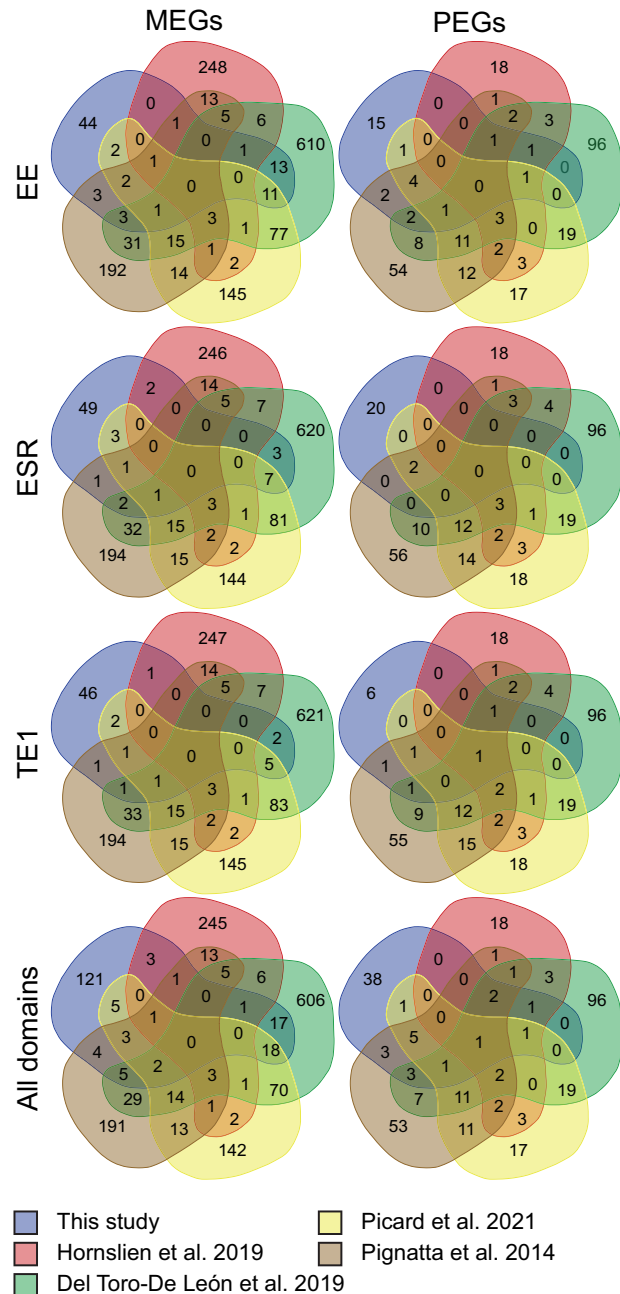
Prior to imprinting analysis various filters were applied (Supplemental Figure S3, Supplemental Table S2B, Supplemental Dataset S4) as described previously (Hornslien et al., 2019). To compensate for ecotype-specific expression bias in the lack of reciprocal crosses for our limited nuclei sorted libraries, a very stringent ecotype filter was applied to avoid the detection of false positives due to ecotype biased expression. Genes with any significant expression bias between the ecotype accessions were identified by differential expression analysis of homozygous Col-0 and Tsu-1 samples and 12,691 and 8,346 genes at early and late stage, respectively, with a significant expression bias between the two ecotypes were omitted from further imprinting analysis (Supplemental Figure S4, Supplemental Dataset S5).



**Figure 2** Homogeneity assessment of endosperm marker lines and verification of spatial differential gene expression. A, Principal component analysis (PCA) of gene expression of individual biological replicates for the respective marker lines; EE (blue triangle), ESR (red diamond), DAL (brown cross), and TE1 (green triangle) to assess the homogeneity within replicates and between each marker line. B, MA plot showing the fold change ( $\log_2$ ) differential expression between ESR and TE1 versus the mean expression ( $\log_2$ ). The analysis revealed 5,324 genes to be significantly differentially expressed (red = up, blue = down, adjusted  $P$ -value  $< 0.05$ ) within the spatial plane. Non-significant (NS) genes are depicted in gray. Dashed lines indicate  $\log_2$  of 1 and  $-1$ . EE = early endosperm; ESR = embryo surrounding region; DAL = developing aleurone layer; TE1 = total endosperm.  $P$ -values were calculated by DESeq2 (Love et al., 2014).

A read pair was considered informative if it mapped to the same gene in the polished transcripts from both parents, such that alignment features (InDels, SNPs) distinguished the parental allele of origin (Hornslie et al., 2019). Ultimately, 6,169, 7,847, and 8,024 genes with informative reads were selected for allele-specific differential expression analysis in EE, ESR, and TE1, respectively (Supplemental Table S2B, Supplemental Dataset S6).

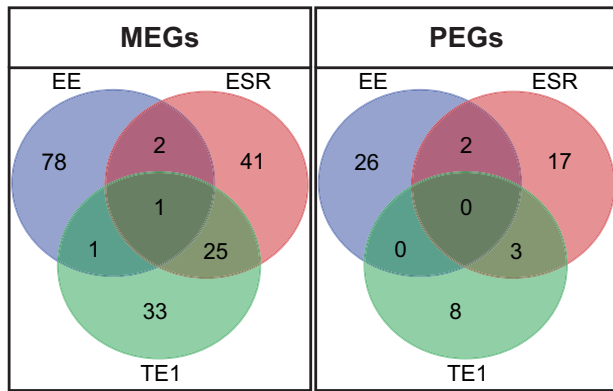
Genes were classified based on the total informative read count, parental bias, and statistical significance (Supplemental Figure S5). First, we identified genes with an insufficient number of informative read counts. Second, genes that did not show a parental bias were identified. Third, genes that showed nonsignificant parental bias were identified. Finally, genes showing a parental bias and a significant fold



**Figure 3** Overlap of identified imprinted genes with other studies. Maternally expressed genes (MEGs, left) and paternally expressed genes (PEGs, right) identified in early endosperm (EE), embryo surrounding region (ESR), total endosperm (TE1), and all domains combined (All domains) were compared with imprinted gene lists from (Pignatta et al., 2014) (brown), (Hornslie et al., 2019) (red), (Del Toro-De León and Köhler, 2019) (green), and (Picard et al., 2021) (yellow).

change (FC) were considered imprinted (Supplemental Figure S5).

In total, across all investigated marker lines, 181 maternally expressed genes (MEGs) and 56 paternally expressed genes (PEGs) were identified. The ratios of MEG:PEG per line was 82:28 in EE, 60:11 in TE1, 69:22 in ESR, and 181:56 overall (Supplemental Dataset S7). Gene ontology (GO) enrichment

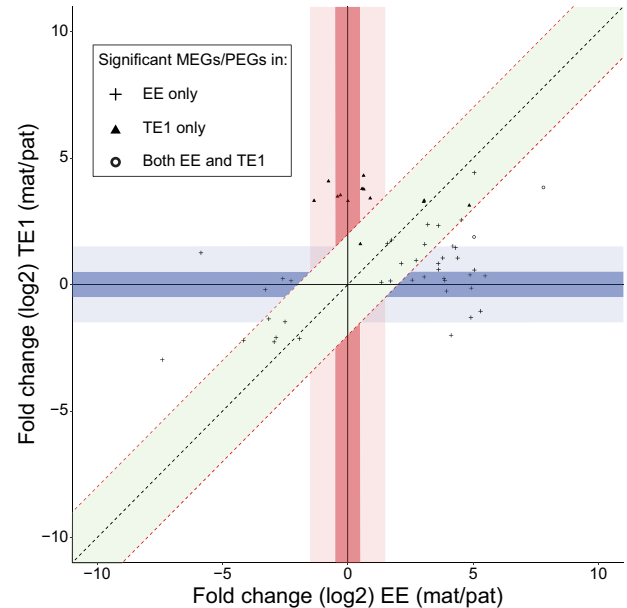


**Figure 4** Overlapping MEGs and PEGs among EE, ESR, and TE1. Venn diagrams display overlap between the identified imprinted genes in EE, ESR, and TE1 experiments. Identified MEGs (left) and PEGs (right) between EE (blue), ESR (red), and TE1 (green) show temporal (EE and TE1), spatial (ESR and TE1), and spatio-temporal (EE, ESR, and TE1) conservation of imprinting. Temporal (EE- or TE1-specific) and spatial (ESR- or TE1-specific) imprinting was observed for both MEGs and PEGs. EE = early endosperm; ESR = embryo surrounding region; TE1 = total endosperm. MEGs = maternally expressed gene; PEGs = paternally expressed genes.

analysis identified significantly enriched GO terms in the imprinted gene panels (Supplemental Table S3, Supplemental Dataset S7). Overall, and in line with previous reports (Pignatta et al. 2014; Picard et al. 2021), enrichment was found for genes encoding proteins involved in gene regulation. For all identified gene panels, out of 237 genes that could be tested, transcription factor activity was enriched (1.8-fold,  $n = 23$ ). More specifically, all MEGs taken together (2-fold,  $n = 20$ ) and also all single MEG panels were enriched for transcription factor activity (Supplemental Table S3, Supplemental Dataset S7), including proteins from several transcription factor (TF) families, such as Homeobox TFs (FWA, BEL1, KNAT3, RPL), MYB domain TFs (MYB7, MYB60), MADS-box (STK), WRKY (WRK50), heat shock (HSF4), zinc-finger (ZF3), EAR-containing (TIE4), and basic helix-loop-helix (TT8) proteins.

### Comparison of imprinting studies

In order to compare our data to previously reported imprinted gene sets, we estimated the overlap between previous recent studies and this study ((Pignatta et al., 2014; Del Toro-De León and Köhler, 2019; Hornslien et al., 2019; Picard et al., 2021); Supplemental Dataset S8). There is measurable overlap of identified imprinted gene identities between our study and other recent studies (Figure 3). All FANS marker lines (EE, TE1, ESR) identified imprinted genes that had been reported previously. The largest overlap to previous studies was found with EE (38 MEGs and 13 PEGs). The overlap to previously identified imprinted genes identified using the ESR marker was low for paternally expressed genes (20 MEGs and two PEGs) suggesting that the ESR experiment may report imprinted genes masked by total



**Figure 5** Identification of temporal regulation of imprinting. Scatter plot representation of the maternal:paternal fold change of significantly called imprinted genes at early developmental stage (EE), late developmental stage (TE1) or in both. Genes called as MEGs or PEGs in one domain; EE (plus) or TE1 (triangle) or both domains EE and TE1 (circle) were included. Genes in the green area show a similar parental bias in both EE and TE1 ( $\Delta FC$  ( $\log_2$ )  $< 2$ ) indicating that they retain their imprinted state throughout seed development. Genes in the dark blue area are only imprinted in EE and are biparentally expressed in TE1 ( $-0.5 < FC$  ( $\log_2$ )  $< 0.5$ ). Genes in the light blue area are MEGs or PEGs for EE, but the parental bias in TE1 has not shifted to the same degree as for the dark blue area ( $-1.5 < FC$  ( $\log_2$ )  $< -0.5$  or  $0.5 < FC$  ( $\log_2$ )  $< 1.5$ ). Genes in the dark red area are only imprinted after cellularization (TE1) and show biparental expression for EE ( $-0.5 < FC$  ( $\log_2$ )  $< 0.5$ ). Concurrently, genes in the light red area are MEGs or PEGs for TE1, but parental bias in EE has not shifted to the same degree as for the dark red area ( $-1.5 < FC$  ( $\log_2$ )  $< -0.5$  or  $0.5 < FC$  ( $\log_2$ )  $< 1.5$ ). MEGs = maternally expressed gene; PEGs = paternally expressed genes.

endosperm in previous studies. All markers taken together, a total of 78 imprinted genes (60 MEGs and 18 PEGs; Supplemental Dataset S9) out of 237 total imprinted genes were previously identified (33%), including *FLOWERING WAGENINGEN* (FWA), *SEEDSTICK* (STK), *BANYULS* (BAN), *FIDDLEHEAD* (FDH), *YUCCA10* (YUC10), *HOMEODOMAIN GLABROUS 3* (HDG3) PEG3, and *PEG6* (Pignatta et al., 2014; Del Toro-De León and Köhler, 2019; Hornslien et al., 2019; Picard et al., 2021).

The high degree of uniquely identified imprinted genes when contrasting our study with other reports as well as between these reports are likely explained by methodological difference and experimental set-up. Most studies are conducted using different ecotype accessions, using different time points of seed development and using different tissue and RNA extraction methods (Supplemental Table S4). Furthermore, genes not identified in this study may also have been excluded by filtering steps, such as the absence

of SNPs in accessions or biased accession expression. This is indeed a limitation for all studies restricted to experimental set-ups with few ecotype accessions. Excluding the genes that did not pass filtering steps in our set-up, and therefore could not be analyzed, increased the overlap substantially (Supplemental Figure S6, Supplemental Dataset S10). This suggests that experimental set-up and bioinformatic analysis considerably impact the resulting identified imprinted genes. Importantly, the largest overlap with our data is observed with studies that used similar endosperm tissue extraction methods (Del Toro-De León and Köhler, 2019; Picard et al., 2021).

### Overlap between temporal and spatial endosperm domains

As a next step, we compared the identities of imprinted genes recognized by our different FANS marker set-ups (EE, TE1, ESR) (Figure 4). This identified many genes that were uniquely called as imprinted in only one spatial or temporal domain (Supplemental Dataset S11). On the other hand, only one MEG overlapped among EE, ESR, and TE1 (Supplemental Dataset S11). This gene, *SEEDSTICK* (*STK*), retains its imprinting state throughout seed development and ESR endosperm differentiation, and the relatively low number may suggest that imprinting is highly dynamic throughout seed development. Interestingly, *STK*, previously described as a transcription factor controlling ovule and seed integument identity (Ezquer et al., 2016), has recently also been identified to be imprinted in the endosperm by other FANS studies (Del Toro-De León and Köhler, 2019; Picard et al., 2021) and directly or indirectly regulates *BAN* and *TRANSPARENT TESTA 8* (*TT8*), respectively (Mizzotti et al., 2014), both genes also identified as MEGs in this study.

Temporal differential imprinting dynamics is frequently observed in the comparison of allele-specific profiles between early- and late-stage markers (EE versus TE1 and EE versus ESR), where a low number of MEGs and PEGs overlap between early and late stages (four MEGs and two PEGs in both comparisons) (Figure 4, Supplemental Dataset S11). In contrast, the allele-specific profiles of the corresponding stage markers (TE1 versus ESR) were highly concordant, sharing 26 and three MEGs and PEGs, respectively (Figure 4). This suggests high accuracy of FANS since ESR represents a subdomain of the TE1 marker. Nevertheless, more than three times the number of loci (77 MEGs and 27 PEGs) were imprinted only in the TE1 or in the ESR domains, suggesting spatial-specific imprinting (Supplemental Dataset S11). Although the overall expression levels, or lack of expression is not taken into account, both the temporal and spatial comparison suggest that genomic imprinting is highly dynamic, both during differentiation and development of the endosperm.

### Temporal dynamic regulation of genomic imprinting

In order to investigate if the parental-specific allelic expression is truly regulated, and not just a consequence of stage-

specific repression or activation of genes, we next analyzed temporal regulation of imprinting by directly comparing the allelic expression values of imprinted genes between early and late developmental stages (EE versus TE1). To this end, we required that a gene should be expressed in both FANS panels, but identified as imprinted (MEG or PEG) in one seed developmental stage, whereas identified as a biparentally expressed gene (BEG) in the other seed developmental stage. Genes found to be imprinted at one time point and not expressed at the other time point were therefore not assessed in this analysis (Supplemental Dataset S12).

The maternal:paternal fold change of (early) EE or (late) TE1 significantly imprinted genes (EE, TE1 and both EE and TE1; Supplemental Dataset S12) were represented in a scatter plot (Figure 5). Three output scenarios can be proposed from this analysis; genes that are imprinted at both developmental time points, genes that are imprinted early but are biparentally expressed at the later developmental time point and genes that are BEGs at the early developmental time point but imprinted at the later developmental time point.

In our analysis, almost half of the genes (21/56) show a similar parental expression bias in EE and TE1 (Figure 5, diagonal green area), suggesting that they retain their imprinted state throughout seed development.

We identified eight MEGs and three PEGs, significantly imprinted in EE, that show biallelic expression in TE1 (Figure 5, horizontal dark blue area; Table 1A). Several of these EE specific MEGs have previously been identified as imprinted at a similar developmental stage (Del Toro-De León and Köhler, 2019; Picard et al., 2021), including *MYB DOMAIN PROTEIN 7* (*MYB7*), *BETA GALACTOSIDASE 1* (*BGAL1*), and *TON1 RECRUITING MOTIF 11* (*TRM11*) (Supplemental Dataset S9). Most MEGs in EE show an increased expression in TE1 suggesting a reactivation of the paternal allele, presumably without any decrease in maternal expression (Table 1A). This is the case for all EE MEGs that are biparentally expressed in TE1 except for *TRM11* and *AT3G58950* (Supplemental Dataset S6). For the majority of these MEGs, we thus infer that methylation marks are lost from paternal alleles, potentially gradually until later stages of seed development. Loss of methylation marks can be achieved by actively removing them exemplified by the 5mC DNA glycosylase family (Penterman et al., 2007). From our data, we observe a > 20-fold increased expression of the gene encoding the DNA glycosylase REPRESSOR OF SILENCING 1 (*ROS1*) from early- to late-stage developing endosperm (Supplemental Dataset S3), highlighting this protein as a candidate for relaxing the imprint on paternal alleles towards later stages of seed development. However, the canonical DNA glycosylase DEMETER (*DME*), generally described to act in the central cell (Ibarra et al., 2012; Park et al., 2020), is in our data also found to be increasingly expressed towards later stages of endosperm development (Supplemental Dataset S3).

Another possibility is a discontinued maintenance of DNA methylation which will partly or completely lead to the loss of methyl groups from a given allele depending on their

**Table 1** Temporal specific imprinted genes. Identified genes that show temporal dynamic regulation of genomic imprinting. A, Genes that are imprinted in endosperm at early developmental stage (EE) and that are biparentally expressed in endosperm at late developmental stage (TE1). B, Genes that are imprinted in TE1 and that are biparentally expressed in EE. Parental bias is represented by log<sub>2</sub> Fold Change (FC)-values between maternal and paternal informative reads within the domain.

AGI	Parental bias log <sub>2</sub> (mat/pat)		Symbol	Gene	Differential Expression <sup>c</sup> –log <sub>2</sub> (EE versus TE1)
	EE <sup>a</sup>	TE1 <sup>b</sup>			
<b>A</b>					
AT1G13390	–3.3	–0.2		Translocase subunit seca	–0.1
AT1G23070	–2.6	0.2	LAZ1H2	Lazarus1 Homolog 2	0.5
AT5G41790	–2.3	0.1	CIP1	COP1-interacting protein	0.6 <sup>d</sup>
AT3G58950	2.6	0.2		F-box/RNA-like superfamily protein	–0.2
AT5G20120	3.0	0.3		Testis- and ovary-specific PAZ domain protein	1.3 <sup>d</sup>
AT5G54200	3.8	0.2	WDD1	Transducin/WD40 repeat-like superfamily protein	2.4 <sup>d</sup>
AT2G26300	3.9	0.2	GPA1	G protein alpha subunit 1	2.4 <sup>d</sup>
AT4G18070	3.9	–0.3		Suppressor	2.4 <sup>d</sup>
AT3G13750	4.9	0.4	BGAL1	Beta-galactosidase 1	2.8 <sup>d</sup>
AT4G23020	4.9	–0.1	TRM11	TON1 recruiting motif 11	0.7
AT2G16720	5.5	0.3	MYB7	MYB domain protein 7	1.0 <sup>d</sup>
<b>B</b>					
AT1G11940	0.0	3.3		Core-2/l-branching beta-1,6-N- acetylglucosaminyltransferase family protein	0.0
AT1G16750	–0.3	3.5		Putative GPI-anchored adhesin-like protein	–4.3 <sup>d</sup>
AT1G19210	–0.4	3.5	ERF17	ERF/AP2 transcription factor 17	–0.3

<sup>a</sup>Significant imprinted MEGs and PEGs.

<sup>b</sup>No significant parental bias.

<sup>c</sup>Differential gene expression output between EE and TE1 total expression for selected genes is given by -FC (log<sub>2</sub>) and values represent higher expression in EE (negative FC) or TE1 (positive FC).

<sup>d</sup>Significant differentially expressed genes between EE and TE1.

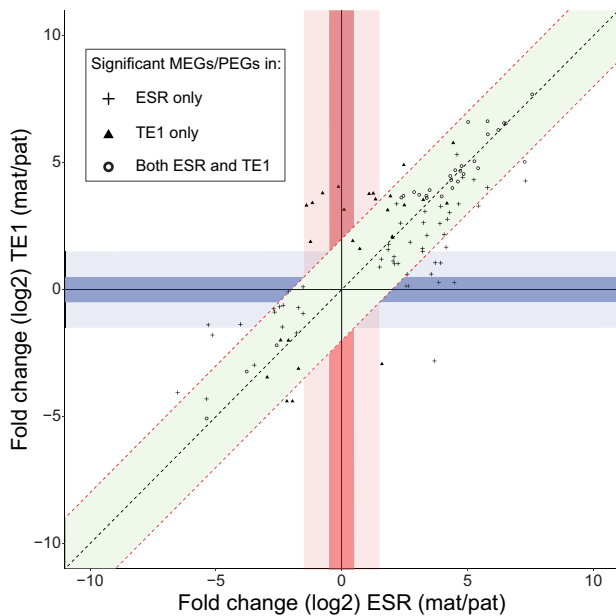
cytosine context in the DNA. Differential expression analysis showed that the genes encoding the two methyltransferases mainly shown to be responsible for maintenance of DNA methylation METHYLTRANSFERASE 1 (MET1) and CHROMOMETHYLASE 3 (CMT3), are expressed at lower levels in TE1 than in EE (Supplemental Dataset S3). This is also in line with previous observations of the activity of these methyltransferases in the endosperm (Jullien et al., 2012).

For EE PEGs, differential gene expression analysis between EE and TE1 (Table 1A, Supplemental Dataset S3) revealed similar expression levels, and show that there is a simultaneous decrease of paternal and an increase of maternal expression (Supplemental Dataset S6). Such a dynamic gene regulation scenario possibly involves several regulation mechanisms acting both through DNA demethylation, as described above, as well as through histone modifications. Histone modifications through the FIS-PRC2-complex are believed to mainly regulate the maternal allele of paternally expressed imprinted genes (Makarevich et al., 2008; Batista and Köhler, 2020). Although only demonstrated for regulation in early seed development it is possible that regulation of maternal alleles of PEGs are regulated by similar, or other, yet unknown mechanisms, early, and also at later seed developmental stages (Hornslien et al., 2019).

Interestingly, three MEGs, significantly imprinted at the late developmental stage (TE1), displayed unbiased parental expression at the early endosperm developmental stage (EE)

(Figure 5, vertical dark red area; Table 1B). Two out of these three MEGs show no significant change in overall expression level comparing EE and TE1 stages (Table 1B, Supplemental Dataset S3). In both cases, the paternal allele is completely silenced at the late time point, whereas the maternal allele is constantly expressed (Supplemental Dataset S6). These results suggest that imprinting marks that affect silencing of parental alleles are not automatically set in the gametes of the gametophyte, but may also be established at later seed developmental stages. Traditionally, the paternal alleles of MEGs are thought to be silenced by DNA methylation, and in this scenario, *de novo* DNA methylation would be required to mediate silencing of the paternal allele in late seed development. The main methyltransferase shown to be involved in a *de novo* methylation is DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2) which indeed is highly up-regulated (>ten-fold) in our TE1 expression profile compared with our EE expression profile (Supplemental Dataset S3), in line with previous endosperm expression studies (Belmonte et al., 2013). Further on, a recent study showed that the MET1 homolog MET3 has an increasing expression level towards later stages of seed development (Tirot et al., 2022). Expression profiling of *met3* mutant seeds compared with wild-type seeds did however only lead to the deregulation of a few genes, none of which overlapped with the candidate genes identified by us. In the same study, MET3 was also shown to be over-expressed in a mutant of





**Figure 6** Identification of spatial-specific imprinted genes. Scatter plot representation of the maternal:paternal fold change of significantly called imprinted genes in total endosperm (TE1), in the embryo surrounding region (ESR) or in both. Genes called as MEGs or PEGs in one domain; ESR (plus) or TE1 (triangle) or both domains ESR and TE1 (circle) were included. Genes in the green area show a similar parental bias in both ESR and TE1 ( $\Delta FC (\log_2) < 2$ ) indicating that their imprinting state is maintained throughout endosperm development reported by ESR and TE1. Genes in the dark blue area are imprinted in ESR but biparentally expressed in the overall total endosperm (TE1;  $-0.5 < FC (\log_2) < 0.5$ ). Similarly, genes in the dark red area are imprinted in the overall total endosperm (TE1) but show biparental expression in the ESR ( $-0.5 < FC (\log_2) < 0.5$ ). Genes in the light blue area are MEGs or PEGs for the ESR, but the parental bias in TE1 has not shifted to the same degree as for the dark blue area ( $-1.5 < FC (\log_2) < -0.5$  or  $0.5 < FC (\log_2) < 1.5$ ). Concurrently, genes in the light red area are MEGs or PEGs for TE1, but parental bias in the ESR has not shifted to the same degree as for the dark red area ( $-1.5 < FC (\log_2) < -0.5$  or  $0.5 < FC (\log_2) < 1.5$ ). MEGs = maternally expressed gene; PEGs = paternally expressed genes.

the FIS-PRC2 protein FERTILIZATION INDEPENDENT ENDOSPERM (FIE), and the authors proposed that this elevated MET3 activity is causing higher DNA methylation levels in the CG-context in the *fie* endosperm, as shown previously (Bouyer et al., 2017). This, taken together with results presented here, supports our observation that canonical epigenetic marks on parentally biased genes in the endosperm are not necessarily established in the gametophytes.

In order to substantiate if the hypothetical role of MET, DRM, and CMT DNA methyltransferases as well as DEMETER-like DNA glycosylases could confer allele-specific regulation of the identified temporal- or endosperm domain-specific imprinting patterns observed, we analyzed community datasets (Hsieh et al., 2011; Stroud et al., 2013; He et al., 2022) to investigate expression patterns of candidate genes in corresponding mutants (and higher order mutant

combinations of these) from various tissues (seedling, leaf, and endosperm). The expression analysis shows that expression of several of the identified candidate genes are indeed highly deregulated in the mutant panels (Supplemental Figure S7, Supplemental Dataset S13). Both MEGs and PEGs are affected by mutation of all three classes of DNA methyltransferases as well as DNA glycosylase. Overall, this analysis is laying the ground for future experiments to investigate potential mechanisms for dynamic regulation of genomic imprinting.

### Spatial dynamic regulation of genomic imprinting

In contrast to the temporal developmental stage comparison (EE versus TE1), the comparison between imprinted loci from the same developmental stage but from partially overlapping endosperm domains (TE1 versus ESR) resulted in 29 imprinted loci (26 MEGs and three PEGs) with the same imprinting status in both domains (Figure 4). Nonetheless, spatial-specific imprinting dominated, and a total of 104 genes (77 MEGs and 27 PEGs) were identified as imprinted in only TE1 or ESR domains (Figure 4, Supplemental Dataset S11). We wanted to identify imprinted loci that displayed an actual bias in their parental contribution to the two domains investigated, and that were not merely caused by the lack of expression in one of the domains. To this end, we performed a direct comparison of the allelic expression in the two domains. Genes that were not significantly expressed in both spatial domains compared were not assessed (Supplemental Dataset S14) and the remaining loci were required to be significantly identified as an MEG or PEG in one or both endosperm domains and at the same time display expression with no parental bias in the other domain.

The maternal:paternal fold changes of significantly imprinted genes (ESR, TE1, and both ESR and TE1; Supplemental Dataset S14) were visualized in a scatter plot (Figure 6). Three types of spatial imprinting dynamics could be hypothesized from this analysis; loci that display the same imprinting status in the ESR and TE1 domains, loci that are MEGs or PEGs in TE1 and BEGs in ESR, and conversely MEGs or PEGs in ESR and BEGs in TE1.

As expected, due to the overlap between the domains, an absolute majority of genes (80/110) that are imprinted in ESR, TE1, or both, and at the same time expressed in both domains show similar parental expression bias in ESR and TE1 (Figure 6, diagonal green area). Twenty-nine loci are significantly identified as imprinted in both the ESR and TE1 domain (Figure 6, circles in diagonal green area). The results also suggest that many of the loci called as significantly imprinted in only one of the two domains (Figure 4, Supplemental Dataset S14) have a similar bias in the other domain, although not significant (Figure 6, triangle or plus in diagonal green area).

Interestingly, we identified five genes significantly imprinted in the ESR (four MEGs and one PEG), that are expressed in a biallelic manner in TE1 (Figure 6, horizontal dark blue area; Table 2A, Supplemental Dataset S14). One

**Table 2** Spatial-specific imprinted genes. Identified genes that show spatial dynamic regulation of genomic imprinting. A, Genes that are imprinted in the embryo surrounding region (ESR) and that are overall biparentally expressed in the total endosperm (TE1). B, Genes that are imprinted in the TE1 but that are biparentally expressed in the ESR. Parental bias is represented by log<sub>2</sub> Fold Change (FC)-values between maternal and paternal informative reads within the domain.

AGI	Parental bias log <sub>2</sub> (mat/pat)		Symbol	Gene	Differential Expression <sup>c</sup> –log <sub>2</sub> (ESR versus TE1)
	ESR <sup>a</sup>	TE1 <sup>b</sup>			
<b>A</b>					
AT5G53140	–2.1	–0.1		Protein phosphatase 2C family protein	0.3
AT4G09970	2.6	0.1		Transmembrane protein	–1.0 <sup>d</sup>
AT4G12080	2.6	0.1	AHL1	AT-hook motif nuclear-localized protein 1	1.0
AT2G23470	3.9	0.3	RUS4	Root UV-B sensitive 4	0.7
AT4G12690	4.5	0.3		DUF868 family protein	0.1
<b>B</b>					
AT2G38480	0.1	3.2	CASPL4B1	CASP-like protein 4B1	–1.2
AT5G50990	–0.1	4.1		Tetratricopeptide repeat (TPR)-like superfamily protein	0.3

<sup>a</sup>Significant imprinted MEGs and PEGs.

<sup>b</sup>No significant parental bias.

<sup>c</sup>Differential gene expression output between ESR and TE1 total expression for selected genes is given by -FC (log<sub>2</sub>) and values represent higher expression in ESR (negative FC) or TE1 (positive FC).

<sup>d</sup>Significant differentially expressed genes between ESR and TE1.

of these loci, *AHL1*, has indeed previously been identified as imprinted in the endosperm (Del Toro-De León and Köhler, 2019; Hornslien et al., 2019) (Supplemental Dataset S9). Albeit, these results may indicate that endosperm domain-specific analysis, such as by the use of FANS, is imperative to successfully identify imprinted genes that otherwise would be masked by overall biparental expression in the endosperm. Differential expression analysis between ESR and TE1 showed that most loci were not significantly changed in overall expression (Table 2A, Supplemental Dataset S3). For all genes identified as MEGs or PEG in ESR and BEGs in TE1, the silenced allele in ESR is higher expressed in TE1 resulting in biparental expression, while the expressed allele of the identified MEGs or PEG is at comparable levels in both TE1 and ESR (Supplemental Dataset S6). Furthermore, we similarly identified two significantly imprinted MEGs in TE1 that are expressed in a biallelic manner in ESR due to higher expression of the paternal allele (Figure 6, horizontal dark red area; Table 2B, Supplemental Dataset S14). To determine if these differences are caused by a reactivation of a silenced allele or the silencing of an expressed allele, it is necessary to know their expression and imprinting level in earlier developmental stages. For the identified genes (Table 2), we can only assess four out of the seven in our EE data set (Supplemental Dataset S6). Although not significant, we see candidates with the same parental bias as well as biparentally expressed genes in EE for these genes suggesting that both reactivation and silencing mechanisms are acting on the different genes in a domain-specific manner in the ESR and the TE1.

Since the ESR is part of the TE1, the imprinting state of TE1 is affected by parental-specific expression in the ESR. For genes that are imprinted in the ESR but identified as biallelic in the TE1 (Table 2A), we presume that the contribution of

the ESR domain to the TE1 profile is minimal and that it will be masked by the TE1 contribution. Therefore, it is not unexpected that genes that are only imprinted in a sub domain of the endosperm are not detected as imprinted when looking at the endosperm as a whole. For genes that are imprinted in the total endosperm, but biallelic in the embryo surrounding region (Table 2B) we infer that the allelic bias must be established in a region of TE1 not including ESR. Indirectly, our results therefore indicate that this class of allelic expression identifies genes that are imprinted in a subdomain of TE1, excluding the ESR. Detection of imprinted genes in subregions of the endosperm as shown here, further indicates that the mechanisms underlying establishment, maintenance or release of imprinting may also act in a highly spatio-specific manner in the endosperm.

## Conclusion

Using FANS, we have successfully captured H2A-GFP domain-specific nuclei from the seed, allowing the generation of pure endosperm expression profiles devoid of embryo and/or seed coat contamination which has been a challenge for studying imprinted genes in the endosperm. In addition, our data identify loci that are spatially differentially imprinted in specific domains of endosperm. This suggests that regulation of genomic imprinting in the endosperm is more dynamic than previously reported and our findings are not readily explained by the canonical mechanisms for genomic imprinting.

In a systematic approach, genes have been assessed for their imprinting state at different seed developmental stages. We demonstrate temporal, developmental stage-specific imprinting for several loci. Genes were found to be imprinted in the syncytial endosperm while being biparentally expressed

at a later seed developmental stage, as previously shown in single-gene studies (Ngo et al., 2012). This observation naturally follows the general assumption on how imprinting is established in gametes (Jullien et al., 2006; Gehring and Satyaki, 2017; Batista and Köhler, 2020), and that allelic bias is observed from early seed developmental stages. However, the systematic analysis provided here demonstrates that while many genes that are imprinted in early stages are completely silenced and switched off at later stages, some genes still retain a dynamic regulation by reactivation of the silenced allele or intricate balancing of the parental expression. Moreover, if expression of a gene remains at a constant level throughout endosperm development while the imprinting state changes, it is implied that both parental alleles are actively silenced and reactivated, suggesting that expression from the maternal and paternal allele is dynamically balanced.

In particular, for the temporal aspect of this analysis we detected imprinted genes that indeed follow a traditional imprinting pattern, by being detected as imprinted genes early in seed development and subsequently being biparentally expressed, or completely silenced, at later stages of seed development. Interestingly, we identified several genes to be biallelically expressed at early stages of seed development that acquire imprinted expression patterns through time. Most known mechanisms associated with genomic imprinting are primarily associated with the establishment of imprinting marks in gametes, which do not explain this type of regulation. In order to accommodate such type of regulation, a secondary mechanism able to distinguish parental alleles must be present to establish imprinting marks at later developmental stages. The RdDM pathway, resulting in small RNA-directed *de novo* DNA methylation could be a potential regulatory mechanism for such a dynamic process (Kirkbride et al., 2019). Here, we show, supported by other studies, that there are candidate methyltransferases and glycosylases as well as RdDM mechanisms to accommodate the emergence of imprinted gene expression in later stages of seed development.

Furthermore, we show spatial dynamic regulation of genomic imprinting by providing direct and indirect evidence that some genes are imprinted only in a certain domain (ESR) of the endosperm or imprinted in a different endosperm domain than the ESR. Similar mechanisms as suggested above could be involved to reactivate the silenced allele in a domain-specific manner. However, further investigation of this type of spatial imprinting is required to determine if the allele-specific regulation is established early and then differentiate in distinct domains, or, if the imprinting is established *de novo* after the specification of endosperm domain identities.

The future application of specific marker lines and FANS, as demonstrated in this report, allow for systematic dissection of endosperm subdomains in both a spatial and temporal manner. This will enhance our knowledge and understanding of discrete functions attributed to

subdomains, and contribute to the future discovery of novel imprinted genes that are not readily identified by investigating the complete endosperm. Ultimately, high-resolution analysis of domain-specific parent-of-origin allelic expression may also contribute to resolve the evolutionary role of imprinting in the endosperm.

## Materials and methods

### Plant material and growth conditions

*Arabidopsis* (*Arabidopsis thaliana*) WT accessions Columbia (Col-0) and Tsushima (Tsu-1) seedlings were grown on MS-2 (Murashige and Skoog medium with 2% w/v sucrose) plates in a 16-h-light/8-h-dark cycle at 22°C for 10 days prior to transferring to soil. Plants were further grown on a 16-h-light/8-h-dark cycle at 18°C. For the FANS experiment, WT accession Tsu-1 and endosperm domain-specific marker lines (Col-0) were grown on a 16-h-light/8-h-dark cycle at 20–22°C.

### Identification and generation of endosperm marker lines

Endosperm domain-specific marker lines were designed utilizing a two-component construct system where mGAL4-VP16 is expressed under control of a selected promoter sequence. The GAL4-VP16 transcription factor then activates expression of HISTONE 2A-GFP through the UAS regulatory element (Olvera-Carrillo et al., 2015). Candidate promoters were selected based on publicly available microarray data (Le et al., 2010) and on gene expression patterns (Winter et al., 2007). The promoters (Supplemental Table S1) were obtained as Gateway cloning compatible amplicons from SAP collection (Benhamed et al., 2008) or amplified from genomic DNA using designated primers (Supplemental Table S5). Promoter fragments were recombined into the pDONRP4P1r vector in a gateway BP reaction (Invitrogen). Subsequently, the promoter entry vectors were assembled together with the pENL1-GAL4-VP16-L2 entry vector into the pB-9FH2A-UAS-7m24GW destination vector in a multisite gateway LR reaction (Invitrogen), generating PROMOTER-OF-INTEREST:GAL4-VP16>>UAS:H2A-GFP constructs. The expression clones were transformed into *Agrobacterium tumefaciens* C58C1 (pMP90) competent cells using electroporation, which was used to transform WT Col-0 plants using the floral dip method (Clough and Bent, 1998). T1 plants were selected on antibiotics and segregation analysis was performed on T2 plants. All further analyses were performed with homozygous single-locus T3 plants of the following marker lines: Early Endosperm (EE; AT5G09370; proEE>>H2A-GFP); Embryo Surrounding Region (ESR; AT3G45130; proESR>>H2A-GFP); DAL (AT4G31060; proDAL>>H2A-GFP); Total Endosperm (TE1; AT4G00220; proTE1>>H2A-GFP).

## Imaging

For fluorescent microscopy, siliques of EE were manually dissected and seeds were imaged in water using a Zeiss Axioplan 2 imaging microscope equipped with a Zeiss Axio cam HDR camera. For confocal imaging of ESR, DAL, and TE1 siliques were manually dissected and developing seeds were fixed in 4% w/v paraformaldehyde (PFA), embedded in 5% w/v agarose and 200  $\mu\text{m}$  sections were obtained using a vibratome. Confocal images were acquired using an Axio Observer coupled to a LSM710 scanner with a Plan-Apochromat 20x/0.8 objective. GFP was excited with the 488 nm laser line of the Argon laser and the emission was detected between 495 and 545 nm. Autofluorescence in the seed was excited with the 633 nm and emission was detected between 638 and 721 nm. For staging of samples, light microscopy of developmental seed stages was performed using an Axioplan2 Imaging microscope equipped with a Zeiss Axio cam HDR camera. Wild-type Col-0 and *Tsu-1* plants were emasculated two days prior to crossing with pollen from the same individual. Siliques were manually dissected at different time points using a stereomicroscope and seeds were mounted on a microscopy slide in a clearing solution of chloral hydrate in 30% v/v glycerol as described previously (Grini et al., 2002).

## Fluorescence-Activated Nuclear Sorting

Closed flower buds were emasculated two days prior to crossing with pollen from the paternal donor. Siliques were dissected at 4 days after pollination (4 DAP—EE) and 7 days after pollination (7 DAP—ESR; DAL; TE1) using a stereomicroscope and seeds were collected and chopped on ice in 50  $\mu\text{l}$  Galbraith buffer (Galbraith et al., 1983) with 0.5% v/v Triton X-100. Three different biological replicates were obtained with 40 siliques (EE; ESR) or 20 siliques (DAL; TE1) each. More Galbraith + Triton X-100 buffer (850  $\mu\text{l}$ ) was added before filtering through a 40  $\mu\text{m}$  Partec filter. Propidium iodide (10  $\mu\text{g}/\mu\text{l}$ ) was added and GFP-positive and -negative nuclei were sorted directly into 500  $\mu\text{l}$  RLT lysis buffer (Qiagen RNeasy Plant Mini Kit) containing 1% v/v 2-mercaptoethanol using a BD FACSAria™ III Cell Sorter (BD Biosciences).

## RNA isolation, library preparation and RNA sequencing of sorted nuclei

RNA was isolated as described in the RNeasy Plant Mini Kit (Qiagen) and amplified with SMART-Seq v4 Ultra Low Input RNA kit (Clontech; version “091817”). Sequencing libraries were prepared with NEBnext Ultra DNA Library Prep Kit (New England Biolabs; version 6.0–2/18) using NEBNext Multiplex Oligos for Illumina-Dual Index Primers Set 1 (#E7600S) and equimolar amounts of libraries from three biological replicates (except for DAL-GFP+; two biological replicates) were pooled and sequenced 150 bp paired end on a NovaSeq6000 S4 flow cell on one lane.

## Reverse transcription quantitative polymerase chain reaction

Nuclei from DAL were sorted for GFP-positive and GFP-negative nuclei. Nuclei from ESR were sorted for GFP-positive nuclei and GFP-negative nuclei, and the latter were re-sorted for 3C and 6C to capture only endosperm nuclei. RNA was extracted (Qiagen Micro RNeasy kit; Qiagen), cDNA was synthesized (iScript cDNA Synthesis Kit; BioRad) and RT-qPCR was performed for GFP with UBIQUITIN10 (UBQ10) as calibrator. The RT-qPCR was done in duplicate on a Lightcycler 480 (Roche) with SYBR green for detection (2.5  $\mu\text{l}$  of mastermix, 0.25  $\mu\text{l}$  of 5  $\mu\text{M}$  of each forward and reverse primer, 1  $\mu\text{l}$  H<sub>2</sub>O and 1  $\mu\text{l}$  of cDNA). Data were analyzed with the  $-\Delta\Delta\text{CT}$  method (Schmittgen and Livak, 2008), statistical analysis was done using the R package “dplyr” (Wickham, 2018) and output was visualized in RStudio using the package “ggplot2” (Wickham, 2016). The following primers were used for the RT-qPCR: GFP forward 5' GACGGC AACTACAAGACCCG 3'; GFP reverse 5' TTCAGCTCGAT GCGTTTAC 3'; UBQ10 forward 5' TTCTGCCATCCT CCAACTGC 3'; UBQ10 reverse 5' CACCCTCCACTTGGT CCTCA 3'.

## RNA isolation and sequencing for the polishing of 4 DAP and 7 DAP target sequences of Col-0 and *Tsu-1*

Closed flower buds were emasculated two days prior to manual self-crossing. For the 4 DAP targets, siliques were dissected and RNA was isolated at 4 DAP for both WT accessions (Hornslien et al., 2019). Since growth conditions varied slightly between plants used for building the WT expression libraries and ecotype accession target sequences, and marker line crosses, seed developmental stage was determined by manually crossing the WT plants. The embryo developmental stage (walking stick) was determined by light microscopy as described above and corresponded to 8/9 DAP (*Tsu-1*) and 10/11 DAP (Col-0). Silique dissection was performed using a stereomicroscope and seeds were harvested in MagNA Lyser Green Beads tubes (Roche) tubes were collected into liquid nitrogen. Three different biological replicates were obtained from four different mother plants and 12 siliques per replicate. RNA was isolated as described in the Spectrum Total Plant RNA Kit (Sigma Aldrich) manual, Protocol A. During step 1, 1 ml Lysis Solution/2-mercaptoethanol solution was added to the tubes with the developing seeds. The tubes were shaken in a MagNA Lyser Instrument (Roche) at 7000 rounds per minute (rpm) for 15 sec, centrifuged at 13,000 rpm in 4°C for 15 sec and then placed at  $-20^{\circ}\text{C}$  for 2 min. This procedure was repeated three times before proceeding with the protocol. On-column DNase digestion was performed as described with On-Column DNaseI Digestion Set (Sigma Aldrich). RNA libraries were prepared using the Strand-Specific TruSeq™ RNA-Seq Library Prep (Illumina). Sequencing, 150 bp paired end, was performed over two lanes using an Illumina HiSeq 4000.

## Bioinformatics

A detailed description of bioinformatic analyses is provided as [supplementary methods](#) and scripts used in this study have been deposited to Github at <https://github.com/PaulGrini/vanEkelenburg>. In short, WT target sequences for Col-0 and Tsu-1 were generated and polished using pilon (Walker et al., 2014). Reads from heterozygous RNA endosperm domain marker lines or from WT Col-0 and Tsu-1 were mapped to the target sequences with bowtie2 (Langmead and Salzberg, 2012). Marker lines were subjected to differential gene expression analysis and principal component analysis using DESeq2 (Love et al., 2014) to assess marker transcriptome sensitivity and replicate similarity, respectively. Differential gene expression analysis between WT Col-0 and Tsu-1 was performed using DESeq2 to identify ecotype-specific expression bias. The Informative Read Pipeline (IRP) (Hornslien et al., 2019) was used to determine and extract informative reads. A read pair is informative, if it maps with indels to one target sequence but without indels to the other target sequence with at most one SNP. A read pair is also informative, if it maps without indels to both target sequences, but mapping to one of the transcriptomes results in fewer SNPs and the larger SNP count is more than twice the smaller count (i.e., 0 versus 1 SNP or 2 versus 5 SNPs). Statistical analysis to identify parent-of-origin-specific differential gene expression bias was performed on informative reads with the R package limma (Ritchie et al., 2015). Bias was measured as fold change (log<sub>2</sub>) and significance was measured as *P*-value < 0.05 after adjustment for multiple tests.

Gene ontology (GO) term enrichment analysis was performed on the identified imprinted genes of the FANS markers using DAVID Bioinformatics Resources version 2021 (Huang et al., 2009a, 2009b) to identify significantly enriched GO terms using the EASE score (*P*-value < 0.05), a modified one-tailed Fisher's exact *t*-test (Hosack et al., 2003).

## Image analysis and figure preparation

Images were processed using Fiji (Schindelin et al., 2012) and assembled in Adobe Illustrator 2021 (Adobe Systems Incorporated, San Jose, USA). Heatmaps were generated using Heatmapper (Babicki et al., 2016).

## Accession numbers

All sequences generated in this study have been deposited in the National Center for Biotechnology Information Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) with project number PRJNA806405.

## Supplemental data

**Supplemental Figure S1.** Fluorescence-activated nuclear sorting (FANS), ploidy analysis, and marker identity verification.

**Supplemental Figure S2.** Spatio-temporal and seed developmental stage differential gene expression.

**Supplemental Figure S3.** Schematic overview of the applied informative read pipeline.

**Supplemental Figure S4.** Ecotype-specific gene expression analysis between Col-0 and Tsu-1 at different seed stages.

**Supplemental Figure S5.** Classification of genes based on fold change (FC) and statistical analysis.

**Supplemental Figure S6.** Overlap of identified imprinted genes with other studies increases after excluding genes that were not analyzed in this study.

**Supplemental Figure S7.** Gene expression of temporal- and spatio-regulated imprinted genes is affected in mutants of DNA methyltransferases and glycosylases.

**Supplemental Table S1.** Expression analysis of endosperm domain-specific promoter marker line candidates.

**Supplemental Table S2.** Whole-seed transcriptomes are enriched in seed coat and endosperm-specific genes compared with FANS transcriptomes.

**Supplemental Table S3.** Gene Ontology (GO) term enrichment analysis.

**Supplemental Table S4.** Summary of experimental set-up of various imprinting analysis studies.

**Supplemental Table S5.** Cloning primers of the endosperm domain marker construct lines.

**Supplemental Table S6.** Library normalization factor for marker line replicates.

Supplemental dataset files ([Supplemental Dataset 1–14](#)) have been deposited to Github at <https://github.com/PaulGrini/vanEkelenburg>.

**Supplemental Dataset S1.** Sequencing information.

**Supplemental Dataset S2.** Enrichment analysis of seed coat and embryo-specific targets.

**Supplemental Dataset S3.** Domains Differential Gene Expression.

**Supplemental Dataset S4.** Homozygous gene filter.

**Supplemental Dataset S5.** Ecotype-specific gene filter.

**Supplemental Dataset S6.** Limma output EE, ESR and TE1.

**Supplemental Dataset S7.** All MEGs and PEGs for EE, ESR and TE1.

**Supplemental Dataset S8.** Identified imprinted genes in other studies.

**Supplemental Dataset S9.** All domains—overlapping MEGs and PEGs with other studies.

**Supplemental Dataset S10.** Overlap in imprinted genes when only looking at genes present in our dataset.

**Supplemental Dataset S11.** Domains specific and overlapping MEGs and PEGs

**Supplemental Dataset S12.** Seed stage-specific imprinting.

**Supplemental Dataset S13.** Differentially imprinted genes in mutant studies.

**Supplemental Dataset S14.** Domain-specific imprinting.

## Funding

This work was supported by the Norwegian Research Council (FRIPRO grant no. 276053 to P.E.G.), the Vlaams Instituut voor Biotechnologie Omics@VIB program cofounded by EU FP7 Marie Skłodowska-Curie Action (MSCA) to M.F. and European Research Council (ERC) StG PROCELLDEATH 639234 and CoG EXECUT.ER 864952 grants to M.K.N.

*Conflict of interest statement.* The authors declare no conflict of interest.

## References

- Adams S, Vinkenoog R, Spielman M, Dickinson HG, Scott RJ (2000) Parent-of-origin effects on seed development in Arabidopsis thaliana require DNA methylation. *Development* **127**: 2493–2502
- Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS (2016) Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* **44**: W147–W153
- Batista RA, Köhler C (2020) Genomic imprinting in plants—revisiting existing models. *Genes Dev* **34**: 24–36
- Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, et al. (2013) Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proc Natl Acad Sci U S A* **110**: E435–E444
- Benhamed M, Martin-Magniette M-L, Taconnat L, Bitton F, Servet C, De Clercq R, De Meyer B, Buyschaert C, Rombauts S, Villarroel R, et al. (2008) Genome-scale Arabidopsis promoter array identifies targets of the histone acetyltransferase GCN5. *Plant J* **56**: 493–504
- Berger F, Grini PE, Schnittger A (2006) Endosperm: an integrator of seed growth and development. *Curr Opin Plant Biol* **9**: 664–670
- Bethke PC, Libourel IGL, Aoyama N, Chung Y-Y, Still DW, Jones RL (2007) The Arabidopsis aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. *Plant Physiol* **143**: 1173–1188
- Birchler JA (1993) Dosage analysis of maize endosperm development. *Annu Rev Genet* **27**: 181–204
- Bjerkan KN, Hornslien KS, Johannessen IM, Krabberød AK, van Ekelerenburg YS, Kalantarian M, Shirzadi R, Comai L, Brysting AK, Bramsiepe J, et al. (2020) Genetic variation and temperature affects hybrid barriers during interspecific hybridization. *Plant J* **101**: 122–140
- Boisnard-Lorig C, Colon-Carmona A, Bauch M, Hodge S, Doerner P, Bancharel E, Dumas C, Haseloff J, Berger F (2001) Dynamic analyses of the expression of the HISTONE::YFP fusion protein in Arabidopsis show that syncytial endosperm is divided in mitotic domains. *The Plant Cell* **13**: 495
- Bouyer D, Kramdi A, Kassam M, Heese M, Schnittger A, Roudier F, Colot V (2017) DNA methylation dynamics during early plant life. *Genome Biol* **18**: 179
- Brown RC, Lemmon BE, Nguyen H, Olsen O-A (1999) Development of endosperm in Arabidopsis thaliana. *Sex Plant Reprod* **12**: 32–42
- Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ (1997) Fertilization-independent seed development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **94**: 4223–4228
- Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**: 735–743
- Deal RB, Henikoff S (2010) The INTACT method for cell type-specific gene expression and chromatin profiling in Arabidopsis thaliana. *Nat Protoc* **6**: 56–68
- Dekkers BJW, Willems L, Bassel GW, van Bolderen-Veldkamp RP (marieke), Ligterink W, Hilhorst HWM, Bentsink L (2011) Identification of reference genes for RT-qPCR expression analysis in Arabidopsis and tomato seeds. *Plant Cell Physiol* **53**: 28–37
- Del Toro-De León G, Köhler C (2019) Endosperm-specific transcriptome analysis by applying the INTACT system. *Plant Reprod* **32**: 55–61
- Denay G, Creff A, Moussu S, Wagnon P, Thévenin J, Gérentes M-F, Chambrier P, Dubreucq B, Ingram G (2014) Endosperm breakdown in Arabidopsis requires heterodimers of the basic helix-loop-helix proteins ZHOUP1 and INDUCER OF CBP EXPRESSION 1. *Development* **141**: 1222–1227
- Doll NM, Royek S, Fujita S, Okuda S, Chamot S, Stintzi A, Widiez T, Hothorn M, Schaller A, Geldner N, et al. (2020) A two-way molecular dialogue between embryo and endosperm is required for seed development. *Science* **367**: 431–435
- Elzquer I, Mizzotti C, Nguema-Ona E, Gotté M, Beauzamy L, Viana VE, Dubrulle N, de Oliveira A C, Caporali E, Koroney A-S, et al. (2016) The developmental regulator SEEDSTICK controls structural and mechanical properties of the Arabidopsis seed coat. *The Plant Cell* **28**: 2478–2492
- Feil R, Berger F (2007) Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet* **23**: 192–199
- Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **220**: 1049–1051
- Gehring M, Satyaki PR (2017) Endosperm and imprinting, inextricably linked. *Plant Physiol* **173**: 143–154
- Grini PE, Jürgens G, Hülskamp M (2002) Embryo and endosperm development is disrupted in the female gametophytic capulet mutants of Arabidopsis. *Genetics* **162**: 1911–1925
- He L, Huang H, Bradai M, Zhao C, You Y, Ma J, Zhao L, Lozano-Durán R, Zhu J-K (2022) DNA methylation-free Arabidopsis reveals crucial roles of DNA methylation in regulating gene expression and development. *Nat Commun* **13**: 1335
- Hornslien KS, Miller JR, Grini PE (2019) Regulation of parent-of-origin allelic expression in the endosperm. *Plant Physiol* **180**: 1498–1519
- Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol* **4**: R70
- Hsieh T-F, Shin J, Uzawa R, Silva P, Cohen S, Bauer MJ, Hashimoto M, Kirkbride RC, Harada JJ, Zilberman D, et al. (2011) Regulation of imprinted gene expression in Arabidopsis endosperm. *Proc Natl Acad Sci U S A* **108**: 1755–1762
- Huang DW, Sherman BT, Lempicki RA (2009a) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57
- Huang DW, Sherman BT, Lempicki RA (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**: 1–13
- Huang S, An YQ, McDowell JM, McKinney EC, Meagher RB (1997) The Arabidopsis ACT11 actin gene is strongly expressed in tissues of the emerging inflorescence, pollen, and developing ovules. *Plant Mol Biol* **33**: 125–139
- Ibarra CA, Feng X, Schoft VK, Hsieh T-F, Uzawa R, Rodrigues JA, Zemach A, Chumak N, Machlicova A, Nishimura T, et al. (2012) Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**: 1360–1364
- Ingouff M, Haseloff J, Berger F (2005) Polycomb group genes control developmental timing of endosperm. *Plant J* **42**: 663–674
- Ingram GC (2010) Family life at close quarters: communication and constraint in angiosperm seed development. *Protoplasma* **247**: 195–214
- Jacob Y, Mongkolsirawatana C, Veley KM, Kim SY, Michaels SD (2007) The nuclear pore protein AtTPR is required for RNA homeostasis, flowering time, and auxin signaling. *Plant Physiol* **144**: 1383–1390

- Jullien PE, Kinoshita T, Ohad N, Berger F** (2006) Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *Plant Cell* **18**: 1360–1372
- Jullien PE, Susaki D, Yelagandula R, Higashiyama T, Berger F** (2012) DNA methylation dynamics during sexual reproduction in Arabidopsis thaliana. *Curr Biol* **22**: 1825–1830
- Kirkbride RC, Lu J, Zhang C, Mosher RA, Baulcombe DC, Chen ZJ** (2019) Maternal small RNAs mediate spatial-temporal regulation of gene expression, imprinting, and seed development in Arabidopsis. *Proc Natl Acad Sci U S A* **116**: 2761–2766
- Kiyosue T, Ohad N, Yadegari R, Hannon M, Dinneny J, Wells D, Katz A, Margossian L, Harada JJ, Goldberg RB, et al.** (1999) Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc Natl Acad Sci U S A* **96**: 4186–4191
- Lafon-Placette C, Köhler C** (2014) Embryo and endosperm, partners in seed development. *Curr Opin Plant Biol* **17**: 64–69
- Langmead B, Salzberg SL** (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359
- Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Horvath S, et al.** (2010) Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proc Natl Acad Sci U S A* **107**: 8063–8070
- Li J, Berger F** (2012) Endosperm: food for humankind and fodder for scientific discoveries. *New Phytol* **195**: 290–305
- Long Y, Liu Z, Jia J, Mo W, Fang L, Lu D, Liu B, Zhang H, Chen W, Zhai J** (2021) FlnRNA-seq: protoplasting-free full-length single-nucleus RNA profiling in plants. *Genome Biol* **22**: 66
- Love MI, Huber W, Anders S** (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550
- Makarevich G, Villar CBR, Erilova A, Köhler C** (2008) Mechanism of PHERES1 imprinting in Arabidopsis. *J Cell Sci* **121**: 906–912
- Mizzotti C, Ezquer I, Paolo D, Rueda-Romero P, Guerra RF, Battaglia R, Rogachev I, Aharoni A, Kater MM, Caporali E, et al.** (2014) SEEDSTICK is a master regulator of development and metabolism in the Arabidopsis seed coat. *PLoS Genet* **10**: e1004856
- Moreno-Romero J, Santos-González J, Hennig L, Köhler C** (2017) Applying the INTACT method to purify endosperm nuclei and to generate parental-specific epigenome profiles. *Nat Protoc* **12**: 238–254
- Ngo QA, Baroux C, Guthörl D, Mozerov P, Collinge MA, Sundaresan V, Grossniklaus U** (2012) The Armadillo repeat gene ZAK IXIK promotes Arabidopsis early embryo and endosperm development through a distinctive gametophytic maternal effect. *Plant Cell* **24**: 4026–4043
- Nowack MK, Ungru A, Bjerkan KN, Grini PE, Schnittger A** (2010) Reproductive cross-talk: seed development in flowering plants. *Biochem Soc Trans* **38**: 604–612
- Olsen O-A** (2004) Nuclear endosperm development in cereals and Arabidopsis thaliana. *Plant Cell* **16**(Suppl): S214–27
- Olvera-Carrillo Y, Van Bel M, Van Hautegeem T, Fendrych M, Huysmans M, Simaskova M, van Durme M, Buscaill P, Rivas S, Coll NS, et al.** (2015) A conserved core of programmed cell death indicator genes discriminates developmentally and environmentally induced programmed cell death in plants. *Plant Physiol* **169**: 2684–2699
- Opsahl-Ferstad HG, Le Deunff E, Dumas C, Rogowsky PM** (1997) ZmEs, a novel endosperm-specific gene expressed in a restricted region around the maize embryo. *Plant J* **12**: 235–246
- Park K, Lee S, Yoo H, Choi Y** (2020) DEMETER-mediated DNA demethylation in gamete companion cells and the endosperm, and its possible role in embryo development in Arabidopsis. *J Plant Biol* **63**: 321–329
- Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL** (2007) DNA demethylation in the Arabidopsis genome. *Proc Natl Acad Sci U S A* **104**: 6752–6757
- Picard CL, Povilus RA, Williams BP, Gehring M** (2021) Transcriptional and imprinting complexity in Arabidopsis seeds at single-nucleus resolution. *Nat Plants* **7**: 730–738
- Pignatta D, Erdmann RM, Scheer E, Picard CL, Bell GW, Gehring M** (2014) Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene imprinting. *Elife* **3**: e03198
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK** (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**: e47
- Satyaki PRV, Gehring M** (2017) DNA methylation and imprinting in plants: machinery and mechanisms. *Crit Rev Biochem Mol Biol* **52**: 163–175
- Satyaki PRV, Gehring M** (2019) Paternally acting canonical RNA-directed DNA methylation pathway genes sensitize Arabidopsis endosperm to paternal genome dosage. *Plant Cell* **31**: 1563–1578
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al.** (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**: 676–682
- Schmittgen TD, Livak KJ** (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**: 1101–1108
- Schon MA, Nodine MD** (2017) Widespread contamination of Arabidopsis embryo and endosperm transcriptome data sets. *Plant Cell* **29**: 608–617
- Scott RJ, Spielman M, Bailey J, Dickinson HG** (1998) Parent-of-origin effects on seed development in Arabidopsis thaliana. *Development* **125**: 3329–3341
- Shirzadi R, Andersen ED, Bjerkan KN, Gloeckle BM, Heese M, Ungru A, Winge P, Koncz C, Aalen RB, Schnittger A, et al.** (2011) Genome-wide transcript profiling of endosperm without paternal contribution identifies parent-of-origin-dependent regulation of AGAMOUS-LIKE36. *PLoS Genet* **7**: e1001303
- Strasburger E** (1900) Einige Bemerkungen zur Frage nach der "doppelten Befruchtung" bei den Angiospermen. *Bot Zeitung (Berlin)* **58**: 293–316
- Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE** (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**: 352–364
- Tanaka H, Onouchi H, Kondo M, Hara-Nishimura I, Nishimura M, Machida C, Machida Y** (2001) A subtilisin-like serine protease is required for epidermal surface formation in Arabidopsis embryos and juvenile plants. *Development* **128**: 4681–4689
- Tirot L, Bonnet DMV, Jullien PE** (2022) DNA methyltransferase 3 (MET3) is regulated by Polycomb group complex during Arabidopsis endosperm development. *Plant Reprod* **35**: 141–151
- Van Hautegeem T, Waters AJ, Goodrich J, Nowack MK** (2015) Only in dying, life: programmed cell death during plant development. *Trends Plant Sci* **20**: 102–113
- Vu TM, Nakamura M, Calarco JP, Susaki D, Lim PQ, Kinoshita T, Higashiyama T, Martienssen RA, Berger F** (2013) RNA-directed DNA methylation regulates parental genomic imprinting at several loci in Arabidopsis. *Development* **140**: 2953–2960
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, et al.** (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**: e112963
- Weijers D, Van Hamburg J-P, Van Rijn E, Hooykaas PJJ, Offringa R** (2003) Diphtheria toxin-mediated cell ablation reveals interregional communication during Arabidopsis seed development. *Plant Physiol* **133**: 1882–1892
- Wickham H** (2018) dplyr: A Grammar of Data Manipulation. R package version 0.7.6. <https://CRAN.R-project.org/package=dplyr>
- Wickham H** (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York

- Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ** (2007) An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One* **2**: e718
- Wolff P, Jiang H, Wang G, Santos-González J, Köhler C** (2015) Paternally expressed imprinted genes establish postzygotic hybridization barriers in *Arabidopsis thaliana*. *Elife* **4**: e10074
- Xiao W, Brown RC, Lemmon BE, Harada JJ, Goldberg RB, Fischer RL** (2006) Regulation of seed size by hypomethylation of maternal and paternal genomes. *Plant Physiol* **142**: 1160–1168
- Xin M, Yang R, Li G, Chen H, Laurie J, Ma C, Wang D, Yao Y, Larkins BA, Sun Q**, et al. (2013) Dynamic expression of imprinted genes associates with maternally controlled nutrient allocation during maize endosperm development. *Plant Cell* **25**: 3212–3227
- Yang S, Johnston N, Talideh E, Mitchell S, Jeffree C, Goodrich J, Ingram G** (2008) The endosperm-specific ZHOUP1 gene of *Arabidopsis thaliana* regulates endosperm breakdown and embryonic epidermal development. *Development* **135**: 3501–3509
- Zhang S, Wang D, Zhang H, Skaggs MI, Lloyd A, Ran D, An L, Schumaker KS, Drews GN, Yadegari R** (2018) FERTILIZATION-INDEPENDENT SEED-polycomb repressive complex 2 plays a dual role in regulating type I MADS-box genes in early endosperm development. *Plant Physiol* **177**: 285–299