Journal of Experimental Botany, Vol. 77, No. 7 pp. 2045–2058, 2026
https://doi.org/10.1093/jxb/eraf294 Advance Access Publication 28 June 2025

EXPERT VIEW

Advances in seed omics

Lucas Auroux[1,2], Lim Chee Liew[1,2], James Whelan[3,4], and Mathew G. Lewsey[1,2,5,6,7,*]

[1] La Trobe Institute for Sustainable Agriculture and Food, AgriBio, La Trobe University, Melbourne, VIC 3083, Australia
[2] La Trobe Institute for Molecular Sciences, La Trobe University, Melbourne, VIC 3083, Australia
[3] State Key Laboratory of Plant Environmental Resilience, College of Life Sciences, Zhejiang University, Hangzhou, 310058, P.R. China
[4] Provincial International Science and Technology Cooperation Base on Engineering Biology, Zhejiang University, Haining, 314400, P.R. China
[5] Australian Research Council Industrial Transformation Research Hub for Medicinal Agriculture, AgriBio, La Trobe University, Melbourne, VIC 3083, Australia
[6] Australian Research Council Centre of Excellence in Plants for Space, AgriBio, La Trobe University, Melbourne, VIC 3083, Australia
[7] Australian Research Council Industrial Transformation Research Hub for Protected Cropping, La Trobe University, Melbourne, VIC 3083, Australia

* Correspondence: M.Lewsey@latrobe.edu.au

Received 12 March 2025; Accepted 23 June 2025

Editor: Christophe Bailly, Sorbonne Université, France

Abstract

Seeds provide 70% of the global food supply, making them crucial for food security. Understanding the molecular mechanisms governing seed development, dormancy, and germination has become increasingly urgent as climate change impacts crop productivity. Over the past 3 years, omics technologies have transformed our understanding of seed biology through comprehensive molecular profiling at unprecedented resolution. This review synthesizes recent advances in seed biology enabled by cutting-edge omics applications in Arabidopsis and crops. We examine how the integration of epigenomics, genomics, transcriptomics, proteomics, and metabolomics analysis has enabled reconstruction of the complex regulatory networks controlling seed development, dormancy, and germination. The recent emergence of single-cell and spatial technologies has been revolutionary, uncovering previously unknown cell types and tissue-specific regulatory mechanisms. Key discoveries include the identification of critical phosphorylation networks, metabolic transitions, and hormone signalling activity in seeds. Advanced genomic approaches have provided insights into crop domestication and trait control, while proteomics and metabolomics have been used to characterize essential regulatory modules controlling dormancy release and germination. These findings provide valuable molecular frameworks for developing climate-resilient crops and enhanced seed vigour through targeted genetic improvements, as well as optimized agricultural practices for ensuring global food security.

Keywords: Development, dormancy, germination, multi-omics integration, single-cell genomics, spatial transcriptomics, systems biology.

Introduction

Seeds are nature's solution for plant reproduction and species dispersal, encapsulating both the genetic information and nutritional reserves necessary for the successful development of offspring. Beyond their evolutionary significance, seeds serve as the world's primary food source, accounting for 70% of human caloric intake, either directly or indirectly through livestock feed. The transition from embryonic to vegetative development (seed to seedling) involves meticulously coordinated molecular and cellular processes across three key stages: seed development (morphogenesis and maturation), dormancy, and germination (Holdsworth *et al.*, 2008; Reed *et al.*, 2022; Verma *et al.*, 2022).

Angiosperm seed development begins with double fertilization in the embryo sac (female gametophyte), where one sperm cell fuses with the haploid egg cell to form the embryo, while the other sperm cell fuses with the central cell (containing two haploid polar nuclei in most flowering plants) to form the endosperm (Williams and Friedman, 2002; Ingram and Gutierrez-Marcos, 2015). This gives rise to three major compartments: the diploid embryo that will develop into the future seedling, the triploid endosperm that serves as a nutrient reserve, and the maternal seed coat that provides protection (Maheshwari, 1950; Baroux and Grossniklaus, 2019). As development advances to seed maturation, the embryo undergoes a series of precise developmental stages, establishing its body pattern and accumulating storage compounds, while the endosperm supports embryo growth (Baroux and Grossniklaus, 2019). Seeds also acquire desiccation tolerance during maturation and enter seed dormancy, a metabolically quiescent state, until germination (Graeber *et al.*, 2012).

Seed dormancy is an adaptive trait that regulates the timing and uniformity of seed germination by preventing seeds from germinating, even under favourable environmental conditions, ensuring seed germination occurs during optimal conditions for successful seedling establishment (Bewley, 1997; Nonogaki, 2014; Nee *et al.*, 2017). Low or deep dormancy, respectively, can lead to pre-harvest sprouting or poor crop emergence, both of which significantly reduce overall crop yield. Dormancy is governed by a complex interplay of plant hormones, with abscisic acid (ABA) maintaining dormancy and gibberellins (GA) promoting germination (Graeber *et al.*, 2012). The depth of dormancy is influenced by both genetic factors and environmental cues experienced by the mother plant during seed development, and storage conditions of the mature seed (Penfield, 2017). This plasticity in dormancy levels enables plants to optimize their reproductive success across diverse environments (Yan and Chen, 2020).

Germination marks the critical transition from a dormant seed to seedling establishment (Rajjou *et al.*, 2012). This process begins with water uptake (imbibition) and encompasses three distinct phases: rapid initial water uptake, a plateau phase of metabolic preparation, and a final phase of embryo expansion leading to radicle emergence (Steinbrecher and Leubner-Metzger, 2017). Successful germination depends on seed vigour, i.e. the set of traits determining the potential for rapid and uniform emergence under diverse environmental conditions (Finch-Savage and Bassel, 2016). Germination involves large-scale dynamic re-programming of gene expression (transcription and translation) and the epigenome, repair of cellular damage accumulated during storage, mobilization of stored reserves, and cell wall modifications to support embryo growth (Nakabayashi *et al.*, 2005; Rajjou *et al.*, 2012; Bassel, 2016; Steinbrecher and Leubner-Metzger, 2017; Narsai *et al.*, 2017a; Narsai *et al.*, 2017b). Seeds sense environmental cues, particularly temperature, moisture, and light (Van Zanten *et al.*, 2011; Narsai *et al.*, 2017a; Narsai *et al.*, 2017b; Yan and Chen, 2020). These factors influence internal molecular networks to regulate the timing of germination and success of seedling establishment.

Despite decades of research using traditional molecular biology approaches, fundamental questions in seed biology remain unanswered: how do individual tissues/cell types communicate and coordinate their activities during development, dormancy release, and germination? What regulatory networks govern the transition between developmental stages? And how is environmental information integrated into these regulatory networks? Understanding these complex molecular mechanisms has become increasingly important for developing climate-resilient crops and ensuring global food security. At the same time, the advent of omics technologies, i.e. comprehensive, system-wide approaches to studying biological molecules, has revolutionized seed biology research and accelerated progress. Omics technologies offer several advantages over traditional approaches, including the ability to capture a holistic view of the molecular landscape, identify novel molecular components and pathways (unbiased by prior knowledge), and integrate multiple biological levels (proteins, transcripts, metabolites, etc.), and to achieve this at tissue or single-cell resolution. They also enable the analysis of dynamic molecular changes over time, generate large datasets for computational modelling, and identify a broader range of genetic and epigenetic targets for crop breeding and engineering (Fig. 1).

This review synthesizes key findings on the application and integration of omics technologies in studies of seeds, at both bulk and the single-cell resolution. We highlight how these advances have helped elucidate the molecular mechanisms underlying seed development, dormancy, and germination, offering novel opportunities for crop improvement in an era of climate uncertainty.

The omics revolution in seed biology

Omics encompasses diverse molecular profiling techniques, including epigenomics (DNA modifications and chromatin organization), genomics (DNA sequence), transcriptomics (RNA abundance and sequence), proteomics (protein abundance and modifications), and metabolomics (metabolite profiles) (Fig. 1). In recent years, omics studies and multi-omics

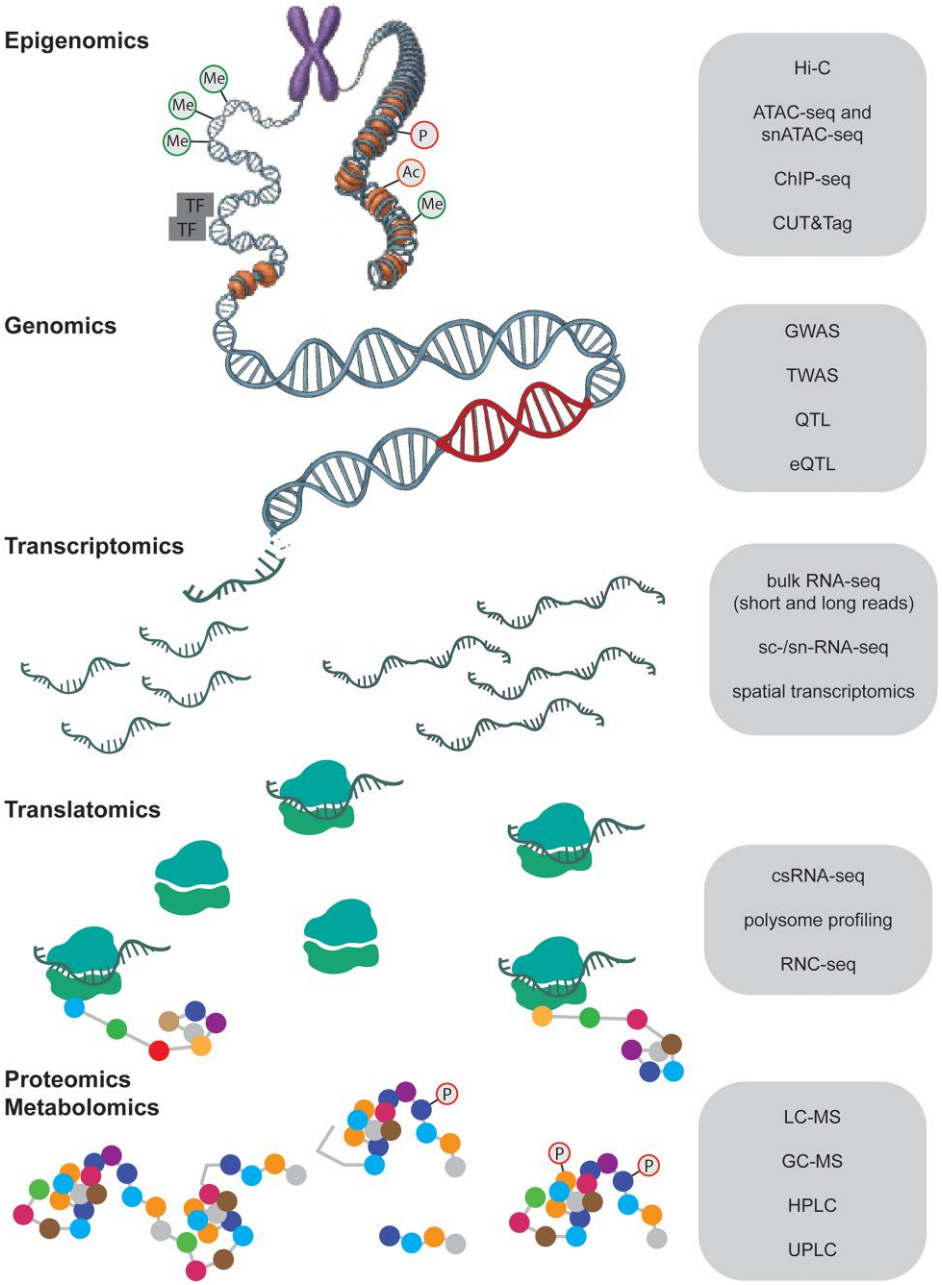

Fig. 1. Hierarchical organization of omics technologies applied to seed biology research (2022–2025). The figure illustrates the multi-level molecular approaches employed in seed biology studies, from genome organization to metabolic output. Epigenomics approaches detect transcription factor (TF) binding and chromatin modifications including methylation (Me), acetylation (Ac), and phosphorylation (P), using High-throughput Chromosome Conformation Capture (Hi-C) to map three-dimensional genome architecture, Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) at the bulk and single-nucleus resolution (single-nucleus ATAC-seq, snATAC-seq) to profile chromatin accessibility, and Chromatin Immunoprecipitation sequencing/Cleavage Under Targets and Tagmentation (ChIP-seq/CUT&Tag) for mapping protein-DNA interactions and histone modifications. Genomics approaches feature DNA sequence analysis (blue helices) and variant region detection (red regions), employing Genome-Wide Association Studies (GWAS), Transcriptome-Wide Association Studies (TWAS), expression quantitative trait loci (QTL), and expression QTL (eQTL) methodologies to associate genetic variation with phenotypic traits and gene expression. Transcriptomics depicts RNA molecules at different levels: bulk RNA-seq (short and long reads), capped-small RNA sequencing (csRNA-seq) for nascent transcription analysis, single-cell/single-nucleus RNA-seq (sc/snRNA-seq) shown as individual RNA molecules, and spatial transcriptomics to analyse RNA distribution and gene regulation across cells and tissues. Translatomics visualizes RNA-ribosome complexes representing nascent translation, analysed through polysome profiling, and ribosome nascent chain complex sequencing (RNC-seq) for capturing translation dynamics. Proteomics and metabolomics gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), and ultra-performance liquid chromatography (UPLC) techniques for comprehensive profiling of proteins including post-translational modifications such as phosphorylated sites (marked as 'P'), and a wide range of metabolites.

data integration, i.e. data integration of multiple molecular types/modalities (multi-modal data integration), have significantly advanced our understanding of seed biology, particularly the intricate regulatory mechanisms governing the onset of plant life.

Dynamic remodelling of chromatin occurs during developmental transitions in seeds

Arabidopsis thaliana (Arabidopsis) seeds undergo nuclear and chromatin re-organization throughout their development and germination (Xiao *et al.*, 2006 ; Van Zanten *et al.*, 2011; Footitt *et al.*, 2015). Early seed development involves modifications of DNA methylation patterns, which affect embryo and endosperm development and ultimately seed size (Xiao *et al.*, 2006). During subsequent maturation and germination, embryonic nuclei go through architectural changes, including chromatin compaction, leading to a significant reduction in nuclear size during maturation (Van Zanten *et al.*, 2011). These structural modifications are further modulated during dormancy through changes to histone modifications (Footitt *et al.*, 2015). Nuclear expansion and chromatin decondensation then occur during germination (Van Zanten *et al.*, 2011). The functional significance of these chromatin dynamics is evidenced by the phenotypes of mutants defective in chromatin remodelling factors failing to repress embryonic traits (Ogas *et al.*, 1999; Tang *et al.*, 2008). Loss-of-function mutations in the CHD3-group chromatin remodelling factor *PICKLE* result in inappropriate accumulation of seed storage protein and mis-expression of embryo-specific genes in seedlings (Ogas *et al.*, 1999). Mutations in the chromatin remodelling ATPase *BRAHMA* leads to the ectopic expression of a subset of embryogenesis-related genes in leaf tissue causing defects in seedling establishment (Tang *et al.*, 2008). Together, these findings highlight the critical role of CHD3- and SWI/SNF-class chromatin re-modellers, such as PICKLE and BRAHMA, in repressing embryonic gene expression during post-germinative development, thereby ensuring proper seedling establishment. Furthermore, the epigenetic regulation of *DELAY OF GERMINATION 1* (*DOG1*), a major regulator of seed dormancy, illustrates the connection between chromatin states and germination potential (Bentsink *et al.*, 2006; Bouyer *et al.*, 2011; Footitt *et al.*, 2015). *DOG1* expression is regulated by a complex interplay of histone modifications, notably the antagonistic marks trimethylation of lysine 4 of histone H3 (H3K4me3) and trimethylation of lysine 27 of histone H3 (H3K27me3), whose deposition at the *DOG1* locus correlates with expression changes during dormancy cycling and early seedling development (Bouyer *et al.*, 2011; Footitt *et al.*, 2015). These findings indicate that extensive re-programming of genome organization and chromatin accessibility is essential for coordinating developmental transitions in seeds.

Epigenomics, the study of DNA or chromatin modifications that do not affect the DNA sequence itself, has provided insights into how plants regulate gene expression during seed development, dormancy, and germination. Epigenomics encompass several high-throughput profiling techniques that analyse DNA modifications and chromatin organization (Box 1). Combined analyses using the Cleavage Under Targets and Tagmentation (CUT&Tag) and Assay for

Box 1. High-throughput epigenomic profiling technologies applied to seed biology studies

(1) **Histone modification mapping:** Chromatin Immunoprecipitation followed by sequencing (ChIP-seq) relies on antibody-based precipitation of cross-linked chromatin complexes (Solomon *et al.*, 1988). Cleavage Under Targets and Tagmentation (CUT&Tag) employs *in situ* antibody-targeted tagmentation, providing an enhanced signal-to-noise ratio and requiring significantly smaller input material (Kaya-Okur *et al.*, 2019).

(2) **Chromatin accessibility assay:** The Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) leverages the preferential integration of Tn5 transposase into accessible DNA regions to map genome-wide open and closed chromatin regions (Buenrostro *et al.*, 2013, 2015). Open chromatin regions are generally associated with enhancer elements or the promoters of actively expressed genes and can be used to identify potential *cis*-regulatory elements (CREs) associated with expressed genes (Auroux *et al.*, 2025). Comparative statistical analysis of ATAC-seq data can identify differentially accessible regions, i.e. loci that show significant differences in chromatin accessibility between experimental conditions, tissue types, or developmental stages. These regions indicate dynamic regulatory activity and frequently coincide with functional elements such as promoters, enhancers, and transcription factor binding sites that influence gene expression.

(3) **Genome-wide chromatin interaction mapping:** High-throughput Chromosome Conformation Capture (Hi-C) is used to investigate chromosome conformation. This method probes three-dimensional genome architecture by quantifying physical interactions between genomic loci through the proximity ligation of cross-linked chromatin fragments (Lieberman-Aiden *et al.*, 2009; van Berkum *et al.*, 2010).

Transposase-Accessible Chromatin using sequencing (ATAC-seq) methods have been used to elucidate a three-phase model of epigenome regulation during wheat embryogenesis (Zhao *et al.*, 2023). The early phase is characterized by extensive histone modification re-programming, followed by a mid-phase dominated by transcription factor-mediated regulation, and culminating in a late phase marked by chromatin condensation (Zhao *et al.*, 2023). The CUT&Tag technology used in this study enabled precise mapping of eight histone modifications/variants and RNA Polymerase II occupancy across eight developmental stages between 0 and 22-days post-anthesis, with minimal nuclei input (~1000 nuclei).

In Arabidopsis, the essential role of the histone variant H3.3 in establishing chromatin accessibility and equipping seeds with the competence for germination and seedling establishment was established using a combination of omics technologies (Zhao *et al.*, 2022). Integration of chromatin immunoprecipitation followed by sequencing (ChIP-seq), ATAC-seq, and strand-specific RNA-seq demonstrated that H3.3 exhibits a seed-specific distribution pattern at 5' gene ends and co-operates with H2A.Z to enhance chromatin accessibility (Zhao *et al.*, 2022). While H3.3 was found to be dispensable for embryogenesis, it was crucial for post-embryonic development, particularly in regulating the expression of hormone-related genes involved in GA synthesis and ABA catabolism (Zhao *et al.*, 2022).

The role(s) of chromosome conformation in regulation of seed oil content (SOC) in *Brassica napus* (oilseed rape) seeds has been studied using High-throughput Chromosome Conformation Capture (Hi-C) (Zhang *et al.*, 2024). Two lines with contrasting SOC levels were compared in a multi-omics analysis that combined Hi-C, ATAC-seq, and RNA-seq. Differentially accessible regions of chromatin and differentially expressed genes were enriched in SOC-related quantitative trait loci (QTLs) and associated genomic regions (Zhang *et al.*, 2024). Fine mapping was applied subsequently to identify the candidate gene *BnaA09g48250D*, where structural variation impacts SOC. Overexpression and knockout of this gene significantly affected SOC in transgenic lines. These findings highlight the role of 3D genome architecture and structural variations in SOC regulation and how chromatin architecture influences agronomically important traits (Zhang *et al.*, 2024).

Epigenomics studies have also focused on the well-known polycomb repressive complex 2 (PRC2) in the context of seed germination. PRC2 is an evolutionary conserved chromatin modifier that mediates the deposition of H3K27me3, a histone mark primarily associated with transcriptional silencing of developmental genes (Turck *et al.*, 2007; Zhang *et al.*, 2007; Mozgová *et al.*, 2017; Miller *et al.*, 2021). While PRC2 was not essential for seed viability, it was crucial for establishing the H3K27me3 repressive mark, with its loss leading to enhanced dormancy, germination defects, and severe developmental defects post-germination (Zhang *et al.*, 2007; Bouyer *et al.*, 2011). More recently it was demonstrated that PRC2 plays a role in a chromatin-level silencing mechanism that regulates Arabidopsis seed germination (Yang *et al.*, 2022). This process is controlled through its synergistic action with the RNA-binding protein RZ-1, which is known to repress the transcription of the *FLOWERING LOCUS C* (Z. Wu *et al.*, 2016b; Yang *et al.*, 2022). Furthermore, cross-linked nuclear immunoprecipitation mass spectrometry revealed the progressive silencing of *9-CIS-EPOXYCAROTENOID DIOXYGENASE 6*, a key player in ABA biosynthesis, through co-ordinated changes in histone modifications including H3K27me3, H3K4me3, and H3 acetylation (Yang *et al.*, 2022). The study also demonstrated that multiple regulatory layers govern this process, with RZ-1 directly interacting with chromatin-modifying factors, such as MSI1 (PRC2-associated protein), HDA19 (histone deacetylase), AL proteins (H3K4me3 readers), NCED6, and ABI3 nascent transcripts (ABA response) (Yang *et al.*, 2022). The dynamics of accumulation of H3K27me3 during the seed-to-seedling transition were studied using ChIP-seq, with RZ-1 found to facilitate establishment of this histone mark (Yang *et al.*, 2022).

The use of genome-wide approaches to identify genetic determinants of seed traits

Genetic mapping has been pivotal in identifying key loci that regulate seed dormancy and germination phenotypes (Koornneef *et al.*, 2002). Conventional quantitative trait loci (QTL) mapping, which relies on linkage analysis in segregating populations, played a foundational role in early efforts to dissect complex traits. In Arabidopsis, QTL mapping using accessions with contrasting dormancy phenotypes—Landsberg *erecta* (low dormancy) and Cape Verde Islands (strong dormancy)—led to the identification of several *DOG* loci, with *DOG1* emerging as a major regulator of seed dormancy (Alonso-Blanco *et al.*, 2003; Bentsink *et al.*, 2006). This discovery marked a major advance in our understanding of dormancy regulation and established a molecular framework for subsequent investigations in diverse species. As genomic technologies have evolved, genome-wide approaches have significantly expanded the scope and resolution of genetic analyses, building upon the foundational insights provided by traditional QTL mapping.

Various high-throughput genomics approaches can now be used to understand genetic variation associated with traits such as germination, seed size or weight, and seed composition. Genome-wide association studies (GWAS) identify statistical associations between genetic variants and phenotypic traits across populations by comparing single nucleotide polymorphisms (SNPs) with observed trait variations, enabling the discovery of genetic loci linked to complex traits (Hirschhorn and Daly, 2005). Transcriptome-wide association studies (TWAS) extend this approach by integrating gene expression data with

genetic variation to identify expression-trait associations, shedding light on the regulatory mechanisms underlying phenotypic variation and improving trait-associated genes identification (Gusev *et al.*, 2016). Expression QTL (eQTL) analysis identifies genetic variants associated with gene expression levels, providing QTL mapping powerful tools for dissecting the genetic architecture of complex traits (Mackay *et al.*, 2009). Together, such genomics approaches enable marker-assisted selection for breeding programmes by pin-pointing genetic markers associated with desirable seed traits, providing a foundation for improving seed performance.

A combined approach using GWAS, TWAS, and eQTL mapping was applied to analyse 421 soybean cultivars and identify loci associated with seed weight and oil content (Yuan *et al.*, 2024). TWAS prioritized 22 candidate genes, including *REGULATOR OF WEIGHT AND OIL OF SEED 1* (*GmRWOS1*), a sodium pump protein verified to significatively increase both seed weight and oil content (Yuan *et al.*, 2024). Similarly, a GWAS study of 164 Japanese rice varieties incorporating genotype-by-environment interactions identified the *GF14 h* gene, a 14-3-3 protein along with its interacting partners bZIP protein, OREB1, and MOTHER OF FT AND TFL 2 (MFT2), as key regulators of seed germination (Yoshida *et al.*, 2022). Notably, loss-of-function alleles of *GF14 h* had been selected in northern regions of East Asia, demonstrating geographical adaptation to the specific climate of that region (Yoshida *et al.*, 2022).

Long-read DNA sequencing technologies, which can sequence large single molecules without fragmentation, have revolutionized genome assembly in crop species (Sedlazeck *et al.*, 2018; Alejo-Jacuinde *et al.*, 2023). These technologies produce reads spanning tens of kilobases and offer superior resolution of repetitive regions and structural variations compared to short-read sequencing technologies. Long-read platforms such as PacBio Sequel II and Oxford Nanopore have been used to produce high-quality genome assemblies for many agriculturally important seed crops including mung bean, oil-Camellia, and chia (*Salvia hispanica*) (Lin *et al.*, 2022; C. Liu *et al.*, 2022a; Alejo-Jacuinde *et al.*, 2023). These high-quality genome assemblies have enabled the identification of candidate genes associated with key traits for crop improvement and breeding, including grain quality, yield, starch and oil content, and disease resistance. High quality genome assemblies, coupled with decreases in sequencing cost and increases in output, have also enabled the construction of pan-genomes, which are a collection of genetic sequences from different populations of a species that capture the near-complete genomic diversity of that species. For example, a maize pan-genome has been assembled from 721 individuals, comprising 507 maize inbred lines, 31 landrace individuals, and 183 teosinte individuals (Gui *et al.*, 2022). From this, 274 649 genome structural variations were identified, many with higher heritability than SNPs or indels, and the specificity or dispensability of genes between ecotypes could be assessed. Notably, genes specific to the ancestral teosinte were enriched in stress response pathways while maize-specific genes were associated with grain nutritional content and flavour pathways (Gui *et al.*, 2022).

Single-cell and spatially resolved transcriptomics uncover cellular heterogeneity in seeds

Transcriptomics technologies, such as RNA-seq and microarrays, have enabled comprehensive analysis of gene expression patterns that regulate seed development, dormancy, and germination (Girke *et al.*, 2000; Bassel *et al.*, 2011; T. Wu *et al.*, 2016a). These approaches provide insights into how genes are expressed, revealing the cellular functions and molecular regulatory networks at different seed developmental stages, as well as gene expression changes in response to environmental cues (light, temperature), hormonal signals (ABA, GA), and various abiotic stresses (Ligterink *et al.*, 2007; Auge *et al.*, 2009; T. Wu *et al.*, 2016a). Most recently, single-cell and spatially resolved transcriptomics studies have provided remarkable insights—previously obscured in bulk tissue analysis—into the cellular heterogeneity of developing and germinating seeds of Arabidopsis, barley (*Hordeum vulgare*), rice, soybean, and *Gossypium bickii* (an Australian wild cotton species), at the resolution of a single cell (Box 2) (Lee *et al.*, 2023, Preprint; Peirats-Llobet *et al.*, 2023; Sun *et al.*, 2023; Liew *et al.*, 2024; Yao *et al.*, 2024; Zhang *et al.*, 2025).

Cottonseed (*Gossypium* spp.) is a valuable source of high-quality protein and oil, but also contains pigment glands that produce gossypol, a substance that is toxic to both humans and animals (de Peyster, 2014). These pigment glands are located in the cotyledons of the embryo (Stanford and Viehoever, 1918). To better understand pigment gland development, cotyledon morphogenesis in seeds of the gland-less and gossypol-free cottonseed relative *Gossypium bickii* was examined using single-cell RNA-seq (Sun *et al.*, 2023). The study identified three previously undefined pigment gland cell subtypes (parenchyma cells, secretory cells, and pre-apoptotic cells), as well as novel regulatory factors with roles in light signalling and gibberellic acid in modulating pigment gland development, which was only possible with the resolution provided by single-cell technology (Sun *et al.*, 2023). Single-cell transcriptomic analysis has also been applied to germinating Arabidopsis seeds, mapping the dynamic progression of cell type-specific gene expression programmes over time (Liew *et al.*, 2024). The study advanced our understanding of germination dynamics by identifying cell types and cell states with distinct transcriptional profiles and revealed a previously unknown common early transcriptional state across embryonic cells before the transition to cell-type specific programmes (Box 2) (Liew *et al.*, 2024).

Spatially resolved transcriptomics (spatial transcriptomics) are technologies that enable transcriptome-wide analysis of tissue samples and preserve information about where genes are

Box 2. Key discoveries in seed transcriptomics reveal new insights in (cell type-specific) transcriptional and regulatory programmes

- Integrated multi-omics atlas reveals the basis of seed development.

  Zhang *et al*. (2024) integrated single-cell RNA-seq and ATAC-seq analysis in soybean, identifying 103 distinct cell types including three previously undefined endosperm subtypes. Their multi-omics analysis approach linked chromatin accessibility dynamics to cell-type specific gene expression, shedding light on seed development.

- A common transcriptional state precedes cell-type differentiation in germinating seeds.

  Liew *et al*. (2024) demonstrated that Arabidopsis embryonic cells transition through a common early transcriptional state before adopting cell-type specific differentiation programmes, providing insight into how cellular identity is established during germination thanks to predicted transcriptional regulators.

- Spatial transcriptomics uncovers nutrient transport dynamics during germination.

  Peirats-Llobet *et al*. (2023) and Yao *et al*. (2024) mapped tissue-specific transcriptional programmes in barley and rice. Their work revealed how spatially distinct cell populations coordinate nutrient mobilization and hormone signalling during germination. Yao *et al*. (2024) integrated spatial and single-cell transcriptomics in rice embryos, identifying previously unknown scutellum cell subtypes and their specific roles in hormone signalling.

- RNA processing machinery controls seed dormancy through the DOG1 pathway.

  Li *et al*. (2023) demonstrated that the pre-mRNA 3' end processing factor FIP1 promotes seed dormancy by regulating DOG1 and ABA-responsive genes, highlighting post-transcriptional regulation as a key mechanism in the dormancy-to-germination transition.

- Alternative splicing regulates seed dormancy.

  L. Liu *et al*. (2022b) identified 342 differentially spliced genes associated with secondary dormancy in oilseed rape cultivars with high secondary dormancy potential. Their analysis showed an enrichment of spliceosome components among these genes, suggesting post-transcriptional regulation as a novel mechanism in seed dormancy regulation.

expressed within a tissue (Auroux *et al*., 2025). There are two main approaches: sequencing-based methods using molecular barcodes to encode a transcript's location, and imaging-based methods that quantify transcript abundance through fluorescent detection (Y. Wang *et al*., 2023b; Auroux *et al*., 2025). Several studies have captured dynamic spatiotemporal transcriptome profiles at both cellular resolution and within the context of tissue architecture by applying a combination of single-cell and spatial methods, offering a deeper understanding of the complex biological processes involved in seed development and germination (Box 2) (Lee *et al*., 2023, Preprint; Peirats-Llobet *et al*., 2023; Yao *et al*., 2024; Zhang *et al*., 2025). These studies successfully identified previously uncharacterized rare cell types, characterized novel cell type-specific markers and transcription factors, predicted spatiotemporal gene regulatory networks, and proposed the roles of different cell types during development and germination across different plant species (Box 2) (Lee *et al*., 2023, Preprint; Peirats-Llobet *et al*., 2023; Yao *et al*., 2024; Zhang *et al*., 2025). A notable finding from these studies was the identification of sub-domain specific gene expression patterns [localized gene expression within a distinct spatial region (sub-domains) of a tissue or organ], including aquaporins, auxin transporters, and lipid transfer proteins—with the latter serving as markers to define novel cell types or sub-types within the scutellum (Peirats-Llobet *et al*., 2023; Yao *et al*., 2024).

Non-coding and post-transcriptional regulatory mechanisms influence seed biology

Capped-small RNA-seq (csRNA-seq) is a high-throughput technology that uses total RNA as starting material to detect transcription start sites (TSSs) of both stable and unstable RNAs at single-nucleotide resolution (Duttke *et al*., 2019).

By specifically targeting 5'-capped RNA species, this technique enables precise genome-wide mapping of transcription initiation events (Duttke *et al.*, 2019). The method's capacity to capture nascent transcription and unstable non-coding RNAs provides insight into real-time transcriptional dynamics and regulatory element activity across the genome (Duttke *et al.*, 2019; Tremblay *et al.*, 2024). More than 30 000 TSSs and accessible chromatin regions were identified using a combination of csRNA-seq and ATAC-seq to study transcription initiation during Arabidopsis seed germination (Tremblay *et al.*, 2024). Non-coding transcription was found to be far more common than previously realized, and many novel transcriptional regulatory elements such as bidirectional promoters and enhancers were mapped (Tremblay *et al.*, 2024).

Polysome profiling and ribosome nascent chain complex sequencing (RNC-seq) enable the identification of mRNA actively engaged in translation by capturing those bound to a single ribosome (monosome) or multiple ribosomes (polysome) using RNA fractionation via sucrose gradients or ultracentrifugation (Mustroph *et al.*, 2009). When applied to seeds, these techniques demonstrated that primed seeds contain significantly more mRNA-ribosome complexes than dry seeds, leading to a faster reactivation of translation during germination (Gran *et al.*, 2025). This pre-association facilitates rapid reactivation of translation during germination, suggesting that translational control begins even before visible metabolic re-awakening (Gran *et al.*, 2025). Furthermore, these techniques have identified specific RNA-binding proteins, such as ARGININE-GLYCINE-GLYCINE RNA BINDING PROTEIN B (RGGB), that appear to coordinate translational re-programming during seed priming and early germination (Bleckmann *et al.*, 2023; Gran *et al.*, 2025). Together, these findings complement single-cell transcriptional approaches by elucidating how post-transcriptional regulation, particularly of stored mRNAs, fine-tunes gene expression during the critical transition from quiescence to metabolic activation.

Post-transcriptional regulation influences RNA stability, splicing, editing, and translation, thereby modulating gene expression and cellular responses beyond direct transcriptional control. Alternative splicing has emerged as an additional regulatory layer for secondary dormancy in oilseed rape, where 342 differentially spliced genes were found to be associated with secondary dormancy by comparing splicing patterns between cultivars with high and low secondary dormancy potential. This suggests that modulating alternate splicing could serve as a strategy to reduce secondary dormancy and improve seed quality (Box 2) (L. Liu *et al.*, 2022b). Post-transcriptional regulation is also important during Arabidopsis seed dormancy (Box 2) (Li *et al.*, 2023). FACTOR INTERACTING WITH POLY(A) POLYMERASE1 (FIP1), a component of the pre-mRNA 3' end processing machinery, positively regulates seed dormancy through both the DOG1 and ABA pathways, revealing complex interactions between transcriptional and post-transcriptional control mechanisms (Box 2) (Li *et al.*, 2023).

Protein phosphorylation and metabolite networks drive seed state transitions

The advancement of transcriptomics has enabled high-throughput profiling of gene expression in seed biology. However, transcript abundances alone do not provide a complete picture of cellular processes. Translation is essential for successful germination—if translation is inhibited, seeds will not germinate—and differential protein phosphorylation plays a central role in regulating the activity of numerous proteins and signalling pathways (Xiaojian *et al.*, 2018). Furthermore, metabolite abundances change rapidly as seeds become active (L. Wang *et al.*, 2023a). Mapping global protein and metabolite abundance is consequently needed to complement transcriptomics data and offer a comprehensive view of cellular function. Proteomic and phosphoproteomic approaches have enabled comprehensive characterization of protein abundance dynamics and regulatory phosphorylation networks with depth and precision during seed development, dormancy, and germination (Yan *et al.*, 2022). Proteomics primarily employs mass spectrometry (MS) or liquid chromatography-MS (LC-MS) for protein identification and quantification, whilst metabolomics encompasses a broader array of analytical platforms such as gas chromatography-MS (GC-MS). These metabolomic platforms profile hundreds of primary and secondary metabolites, both in a targeted and an un-targeted manner (Yan *et al.*, 2022). Phosphoproteomics uses specialized MS platforms and enrichment strategies to enhance phosphopeptide detection sensitivity and coverage (Baudouin *et al.*, 2022; Yamashita and Umezawa, 2022).

Specific regulatory modules controlling dormancy release and germination processes have been identified by integration of proteomic approaches (Baudouin *et al.*, 2022; Pan *et al.*, 2023). In *Amomum tsaoko* (Chinese black cardamom), modifications in MAPK signalling cascades and hormone signal transduction pathways during dormancy release by warm stratification were identified using coupled proteome and transcriptome analysis (Pan *et al.*, 2023). Protein phosphorylation plays a critical role in both ABA and GA signalling pathways. For example, phosphorylation of ABA INSENSITIVE 5 (ABI5) in Arabidopsis prevents its degradation, promoting an ABA response, whereas phosphorylation of the rice DELLA protein OsSLR1 negatively regulates GA signalling (Dai and Xue, 2010; Dai *et al.*, 2013). Phosphorylation networks exhibit remarkable dynamism during Arabidopsis and rice germination (Baudouin *et al.*, 2022; Hu *et al.*, 2024). A comprehensive set of 10 244 phosphopeptides from 2546 phosphoproteins, including 110 protein kinases and key dormancy regulators like DOG1, were identified in Arabidopsis (Baudouin *et al.*, 2022). Extensive early protein dephosphorylation events followed by condition-specific phosphorylation patterns were also observed (Baudouin *et al.*, 2022).

Sucrose non-fermenting-1-related protein kinase 1 (SnRK1) is involved in metabolic re-programming during

energy limitation and in germination regulation but has distinct roles in monocotyledonous seeds compared with dicotyledonous seeds. In dicotyledons like Arabidopsis, SnRK1 promotes seed dormancy through the ABA signalling pathway and the master regulator of seed maturity FUSCA3, which enhances ABA synthesis while suppressing GA biosynthesis (Nambara *et al.*, 2000; Gazzarrini *et al.*, 2004; Tsai and Gazzarrini, 2012). Conversely, reduced SnRK1 abundance in the monocots rice and wheat causes delayed germination due to its role in α-amylase activation, which is essential for starch mobilization during early seedling establishment (Laurie *et al.*, 2003; Lu *et al.*, 2007; Hu *et al.*, 2022). The function of SnRK1 during rice germination has been further elucidated using phosphoproteomics analysis (Hu *et al.*, 2024). SnRK1 activity was modulated using a chemical activator, allowing 8387 unique specific locations where a phosphate group can be attached or removed (phosphosites) to be identified across the rice proteome (Hu *et al.*, 2024). SnRK1 was shown to co-ordinate water transport, Na^+ flux, and ABA signalling during rice germination through the phosphorylation of specific phosphosites (Hu *et al.*, 2024). Additionally, 20 SnRK1-dependent phosphosites in 16 proteins were validated and shown to modulate starch degradation and transport protein function during grain filling (Hu *et al.*, 2022).

Metabolomics have been used to investigate seed biology in several species, including grain development of qingke [Tibetan hulless barley (*Hordeum vulgare* L. var. *nudum* Hook. f.)], as well as germination of quinoa (*Chenopodium quinoa*), pigmented rice, and barley (Hao *et al.*, 2022; Yao *et al.*, 2022; Tiozon *et al.*, 2023; Sybilska *et al.*, 2024; Wang *et al.*, 2024). Three key anthocyanin compounds and their regulatory pathways were identified during qingke grain development, revealing the regulatory network governing anthocyanin synthesis (Wang *et al.*, 2024). Stage-specific metabolic changes were observed across five developmental stages of quinoa germination, ranging from dry seed to cotyledon expansion (Hao *et al.*, 2022). These included dynamic alterations in storage reserve mobilization, cell wall re-modelling, and amino acid metabolism (Hao *et al.*, 2022). Similarly, over 600 metabolites were identified in germinated pigmented rice, with significant increases in free phenolics, dipeptides, and micronutrients, compared to non-germinated samples (Tiozon *et al.*, 2023).

Regulatory complexity in seed biology can be decoded by integrating multi-omics data

Integrative analysis of multiple seed omics data modalities is a new and promising approach to understanding seed biology at a systems scale. For instance, Hi-C epigenomic technology has been coupled with genomic analysis such as GWAS and long-read sequencing to construct high-quality chromosome-scale reference genomes for oil-Camellia and mung bean (Lin *et al.*, 2022; C. Liu *et al.*, 2022a). A combination of GWAS with metabolomics and mineral profiling (by inductively coupled plasma optical emission spectroscopy) has been used to identify genetic markers associated with nutritional enhancement during pigmented rice germination (Tiozon *et al.*, 2023). The resulting machine learning model could predict multi-nutritional properties (Tiozon *et al.*, 2023). Distinct temporal phases associated with metabolic transitions within the first 4 hours of rice imbibition were identified using label-free shotgun proteomics coupled with metabolomics and microarray-based transcriptomics (Sano *et al.*, 2022). Striking tissue-specific protein accumulation patterns were observed throughout rice germination: 417 proteins were differentially accumulated in the embryo compared to 121 in the endosperm, with 109 of the endosperm proteins being degraded during germination, highlighting the importance of the degradation of proteins stored in the endosperm during germination to support embryo growth and seedling establishment (Sano *et al.*, 2022). Most recently, integrative analyses of single-cell and/or spatial transcriptomics data have emerged as an approach to interrogate cell-specific activity. For example, soybean seed and silique development has been analysed by a combination of single-nucleus RNA-seq, single-cell ATAC-seq, and spatial transcriptomics (Zhang *et al.*, 2025). This integrative multi-omics approach identified the three regions of the soybean seed (embryo, endosperm, and seed coat), and the three endosperm cell subtypes (micropylar, peripheral, and chalazal), as well as specific cell types such as the seed coat endothelium and the seed coat inner integument (Box 2) (Zhang *et al.*, 2025). Similarly, single-nucleus RNA-seq combined with spatial transcriptomics in Arabidopsis seeds and siliques revealed nine groups of embryonic cells, each with a common transcriptome, including one potential embryo cell subtype or embryo developmental cell stage (Lee *et al.*, 2023, Preprint).

Emerging frontiers: single-cell multi-omics and systems approaches in seed biology

The application and integration of advanced omics technologies have drastically enhanced our understanding of seed biology in multiple species, revealing intricate molecular dynamics associated with seed development, dormancy, and germination (Table 1). The convergence of high-throughput sequencing technologies, advanced mass spectrometry platforms, and sophisticated computational integrative approaches has enabled insights across multiple scales, from individual cells to whole tissues (Fig. 1). Single-cell and spatial technologies have emerged as particularly transformative, offering cell type-specific resolution of molecular processes previously obscured in bulk tissue analyses. The application of single-cell RNA-seq and single-cell ATAC-seq has revealed previously uncharacterized cell populations and regulatory networks in multiple crop species (Lee *et al.*, 2023, Preprint; Sun *et al.*, 2023; Liew *et al.*, 2024; Zhang *et al.*, 2025). These approaches, often combined with spatial transcriptomics, have illuminated the

Table 1. Systematic categorization of omics technologies in recent seed biology research (2022–2025)

Omics	Epigenomics					Genomics			Transcriptomics					Translat-omics		Proteo-mics		Metabolomics
Technique	Hi-C	ATAC-seq	snATAC-seq	ChIP-seq	CUT&Tag	GWAS	TWAS	eQTL	RNA-seq	csRNA-seq	long-read sequencing	sc/snRNA-seq	Spatially resolved	Polysome profiling	RNC-seq	MS	Phosphoproteomics	
Zhang et al., 2024	✓	✓							✓									
Zhao *et al.*, 2023				✓	✓													
Yang *et al.*, 2022				✓					✓									
Zhang et al., 2025		✓	✓									✓	✓					
Lee et al., 2023, Preprint												✓	✓					
Peirats-Llobet *et al.*, 2023													✓					
Sun *et al.*, 2023									✓			✓						
Liew *et al.*, 2024									✓			✓						
Yao et al., 2024												✓	✓					
Tremblay et al., 2024		✓								✓								
Yao et al., 2022																✓		✓
Gran *et al.*, 2025														✓	✓			
Wang et al., 2024																✓		✓
Tiozon et al., 2023						✓												✓
Lin *et al.*, 2022	✓					✓		✓	✓									
C. Liu *et al.*, 2022a	✓					✓			✓		✓							
Alejo-Jacuinde et al., 2023	✓										✓					✓		✓
Gui et al., 2022						✓			✓									
L. Liu *et al.*, 2022b									✓									
Hao *et al.*, 2022																		
Sybilska *et al.*, 2024																		✓
Yuan *et al.*, 2024						✓	✓	✓	✓									
Yoshida *et al.*, 2022						✓												
Sano et al., 2022									✓							✓		✓
Pan et al., 2023									✓							✓		
Baudouin *et al.*, 2022																✓	✓	

(*continued*)

Table 1. Continued.

Omics	Epigenomics					Genomics			Transcriptomics					Translat-omics		Proteo-mics		Metabolomics
Technique	Hi-C	ATAC-seq	snATAC-seq	ChIP-seq	CUT&Tag	GWAS	TWAS	eQTL	RNA-seq	csRNA-seq	long-read sequencing	sc/snRNA-seq	Spatially resolved	Polysome profiling	RNC-seq	MS	Phosphoproteomics	
Hu et al., 2024																	✓	
Yamashita and Umezawa, 2022																	✓	

This table categorizes omics technologies used in seed biology research from 2022 to 2025. Technologies are grouped into major omics categories (columns) and specific techniques (sub-columns). Each row corresponds to an individual study, with checkmarks indicating the use of particular techniques. Omics categories and techniques: Epigenomics, High-throughput Chromosome Conformation Capture (Hi-C), Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) and single-nucleus ATAC-seq (snATAC-seq), Chromatin Immunoprecipitation followed by sequencing (ChIP-seq), Cleavage Under Targets and Tagmentation (CUT&Tag); Genomics, Genome-Wide Association Studies (GWAS), Transcriptome-Wide Association Studies (TWAS), expression quantitative trait loci (eQTL); Transcriptomics, capped-small RNA sequencing (csRNA-seq), bulk RNA sequencing (RNA-seq), single-cell/-nucleus RNA-seq (sc/snRNA-seq), spatially resolved transcriptomics; Translatomics: polysome profiling and ribosome nascent chain complex sequencing (RNC-seq); Proteomics: mass spectrometry (MS) and phosphoproteomics; Metabolomics, gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), ultra-performance liquid chromatography (UPLC). Italicized references indicate studies employing multi-omics integration analysis approaches.

diversity of cell-type specific gene expression programmes during seed development and germination (Yao *et al.*, 2024; Zhang *et al.*, 2025).

Nonetheless, key challenges remain. We have learned the sequence of molecular events once germination has been initiated by using current omics techniques. Physiological and genetic studies have taught us the triggers for germination. However, the molecular cascade connecting stimulus to cell-type communication and co-ordination of dormancy regulation, as well as the molecular cascade initiating germination remain unknown.

Emerging technologies, such as single-cell multi-omics, Hi-C sequencing, spatial metabolomics, and subcellular protein mapping promise to address these gaps. Applications of single-cell multi-omics experimental approaches, i.e. simultaneously profiling transcriptomes, epigenomes, and proteins within individual cells, could reveal coordinated regulatory mechanisms governing cell-type specific responses (Auroux *et al.*, 2025). These insights are increasingly urgent in the context of climate change, where understanding seed responses to environmental cues is crucial for developing climate-resilient crops. Future research focusing on diverse germplasm collections could identify key determinants of seed longevity and stress resilience, crucial for seed banks and agriculture.

While significant progress has been made, seed biology research still holds great potential for further discovery and advancement. The field is poised for transformative discoveries through the convergence of single-cell technologies and systems biology approaches. The rapid advances in seed omics are transforming our understanding of seed biology, offering new opportunities for improving agricultural productivity and sustainability. By integrating multi-omics data, researchers are unlocking the genetic and biochemical secrets of seeds, paving the way for the development of crops that can meet the challenges of a changing climate and a growing global population. As omics technologies continue to evolve, the future of seed science promises even greater discoveries that will illuminate the birth of plant life and shape the future of agriculture and food security.

Conflict of interest

The authors declare no competing interests.

Funding

This work was supported by grants from the Australian Research Council to MGL (CE230100015, DP220102840, IH180100006, IH240100024). JW is supported by a Yangtze River Scholar award from the Ministry of Education The People Republic of China. LA is supported by a La Trobe University Graduate Research Scholarship and a La Trobe University Full Fee Research Scholarship.

References

Alejo-Jacuinde G, Najera-Gonzalez HR, Chavez Montes RA, Gutierrez Reyes CD, Barragan-Rosillo AC, Perez Sanchez B, Mechref Y, Lopez-Arredondo D, Yong-Villalobos L, Herrera-Estrella L. 2023. Multi-omic analyses reveal the unique properties of chia (*Salvia hispanica*) seed metabolism. Communications Biology **6**, 820.

Alonso-Blanco C, Bentsink L, Hanhart CJ, Blankestijn-de Vries H, Koornneef M. 2003. Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis thaliana*. Genetics **164**, 711–729.

Auge GA, Perelman S, Crocco CD, Sanchez RA, Botto JF. 2009. Gene expression analysis of light-modulated germination in tomato seeds. The New Phytologist **183**, 301–314.

Auroux L, Liew LC, Lewsey MG. 2025. An overview of single-cell high-throughput technology in plants. In: Chen J-T, ed. Guide to plant single-cell technology. Amsterdam: Academic Press, 1–34.

Baroux C, Grossniklaus U. 2019. Chapter twenty-one–seeds–an evolutionary innovation underlying reproductive success in flowering plants. In: Grossniklaus U, ed. Current topics in developmental biology, Vol. **131**. Amsterdam: Academic Press, 605–642.

Bassel GW. 2016. To grow or not to grow? Trends in Plant Science **21**, 498–505.

Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ. 2011. Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. Proceedings of the National Academy of Sciences **108**, 9709–9714.

Baudouin E, Puyaubert J, Meimoun P, Blein-Nicolas M, Davanture M, Zivy M, Bailly C. 2022. Dynamics of protein phosphorylation during Arabidopsis seed germination. International Journal of Molecular Sciences **23**, 7059.

Bentsink L, Jowett J, Hanhart CJ, Koornneef M. 2006. Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. Proceedings of the National Academy of Sciences **103**, 17042–17047.

Bewley JD. 1997. Seed germination and dormancy. The Plant Cell **9**, 1055–1066.

Bleckmann A, Spitzlberger N, Denninger P, ***et al.*** 2023. Cytosolic RGG RNA-binding proteins are temperature sensitive flowering time regulators in Arabidopsis. Biological Chemistry **404**, 1069–1084.

Bouyer D, Roudier F, Heese M, ***et al.*** 2011. Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. PLoS Genetics **7**, e1002014.

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods **10**, 1213–1218.

Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: a method for assaying chromatin accessibility genome-wide. Current Protocols in Molecular Biology **109**, 21.29.1-21.29.9.

Dai C, Xue HW. 2010. Rice early flowering1, a CKI, phosphorylates DELLA protein SLR1 to negatively regulate gibberellin signalling. The EMBO Journal **29**, 1916–1927.

Dai M, Xue Q, McCray T, ***et al.*** 2013. The PP6 phosphatase regulates ABI5 phosphorylation and abscisic acid signaling in Arabidopsis. The Plant Cell **25**, 517–534.

de Peyster A. 2014. Gossypol. In: Wexler P, ed. Encyclopedia of toxicology (third edition). Oxford: Academic Press, 782–785.

Duttke SH, Chang MW, Heinz S, Benner C. 2019. Identification and dynamic quantification of regulatory elements using total RNA. Genome Research **29**, 1836–1846.

Finch-Savage WE, Bassel GW. 2016. Seed vigour and crop establishment: extending performance beyond adaptation. Journal of Experimental Botany **67**, 567–591.

Footitt S, Muller K, Kermode AR, Finch-Savage WE. 2015. Seed dormancy cycling in Arabidopsis: chromatin remodelling and regulation of DOG1 in response to seasonal environmental signals. The Plant Journal **81**, 413–425.

Gazzarrini S, Tsuchiya Y, Lumba S, Okamoto M, McCourt P. 2004. The transcription factor FUSCA3 controls developmental timing in Arabidopsis

through the hormones gibberellin and abscisic acid. Developmental Cell **7**, 373–385.

Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J. 2000. Microarray analysis of developing Arabidopsis seeds. Plant Physiology **124**, 1570–1581.

Graeber K, Nakabayashi K, Miatton E, Leubner-Metzger G, Soppe WJ. 2012. Molecular mechanisms of seed dormancy. Plant, Cell & Environment **35**, 1769–1786.

Gran P, Visscher TW, Bai B, Nijveen H, Mahboubi A, Bakermans LL, Willems LAJ, Bentsink L. 2025. Unravelling the dynamics of seed-stored mRNAs during seed priming. The New Phytologist 247, 2196–2209.

Gui S, Wei W, Jiang C, ***et al.*** 2022. A pan-*Zea* genome map for enhancing maize improvement. Genome Biology **23**, 178.

Gusev A, Ko A, Shi H, ***et al.*** 2016. Integrative approaches for large-scale transcriptome-wide association studies. Nature Genetics **48**, 245–252.

Hao Y, Hong Y, Guo H, Qin P, Huang A, Yang X, Ren G. 2022. Transcriptomic and metabolomic landscape of quinoa during seed germination. BMC Plant Biology **22**, 237.

Hirschhorn JN, Daly MJ. 2005. Genome-wide association studies for common diseases and complex traits. Nature Reviews Genetics **6**, 95–108.

Holdsworth MJ, Bentsink L, Soppe WJJ. 2008. Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. The New Phytologist **179**, 33–54.

Hu Y, Lin Y, Bai J, Xu X, Wang Z, Ding C, Ding Y, Chen L. 2024. AMPK activator 991 specifically activates SnRK1 and thereby affects seed germination in rice. Journal of Experimental Botany **75**, 2917–2932.

Hu Y, Liu J, Lin Y, ***et al.*** 2022. Sucrose nonfermenting-1-related protein kinase 1 regulates sheath-to-panicle transport of nonstructural carbohydrates during rice grain filling. Plant Physiology **189**, 1694–1714.

Ingram G, Gutierrez-Marcos J. 2015. Peptide signalling during angiosperm seed development. Journal of Experimental Botany **66**, 5151–5159.

Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. 2019. CUT&tag for efficient epigenomic profiling of small samples and single cells. Nature Communications **10**, 1930.

Koornneef M, Bentsink L, Hilhorst H. 2002. Seed dormancy and germination. Current Opinion in Plant Biology **5**, 33–36.

Laurie S, McKibbin RS, Halford NG. 2003. Antisense SNF1-related (SnRK1) protein kinase gene represses transient activity of an alpha-amylase (alpha-amy2) gene promoter in cultured wheat embryos. Journal of Experimental Botany **54**, 739–747.

Lee TA, Nobori T, Illouz-Eliaz N, Xu J, Jow B, Nery JR, Ecker JR. 2023. A single-nucleus atlas of seed-to-seed development in Arabidopsis. BioRxiv doi: 10.1101/2023.03.23.533992. [Preprint].

Li Y, Chen F, Yang Y, Han Y, Ren Z, Li X, Soppe WJJ, Cao H, Liu Y. 2023. The Arabidopsis pre-mRNA 3' end processing related protein FIP1 promotes seed dormancy via the DOG1 and ABA pathways. The Plant Journal **115**, 494–509.

Lieberman-Aiden E, van Berkum NL, Williams L, ***et al.*** 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science **326**, 289–293.

Liew LC, You Y, Auroux L, ***et al.*** 2024. Establishment of single-cell transcriptional states during seed germination. Nature Plants **10**, 1418–1434.

Ligterink W, Kodde J, Lammers M, Dassen H, van der Geest A, de Maagd R, Hilhorst H. 2007. Stress-inducible gene expression and its impact on seed and plant performance: a microarray approach. In: Adkins SW, Ashmore S, Navie SC, eds. Seeds: biology, development and ecology. Proceedings of the eighth international workshop on seeds, Brisbane, Australia, May 2005. Wallingford: CABI, 139–148.

Lin P, Wang K, Wang Y, ***et al.*** 2022. The genome of oil-Camellia and population genomics analysis provide insights into seed oil domestication. Genome Biology **23**, 14.

Liu C, Wang Y, Peng J, ***et al.*** 2022a. High-quality genome assembly and pan-genome studies facilitate genetic discovery in mung bean and its improvement. Plant Communications **3**, 100352.

Liu L, Wu D, Gu Y, Liu F, Liu B, Mao F, Yi X, Tang T, Zhao X. 2022b. Comprehensive profiling of alternative splicing landscape during secondary dormancy in oilseed rape (*Brassica napus* L.). Molecular Breeding: New Strategies in Plant Improvement **42**, 44.

Lu CA, Lin CC, Lee KW, Chen JL, Huang LF, Ho SL, Liu HJ, Hsing YI, Yu SM. 2007. The SnRK1A protein kinase plays a key role in sugar signaling during germination and seedling growth of rice. The Plant Cell **19**, 2484–2499.

Mackay TF, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. Nature Reviews Genetics **10**, 565–577.

Maheshwari P. 1950. An introduction to the embryology of angiosperms. New York, NY: McGraw-Hill.

Miller SA, Damle M, Kim J, Kingston RE. 2021. Full methylation of H3K27 by PRC2 is dispensable for initial embryoid body formation but required to maintain differentiated cell identity. Development **148**, dev196329.

Mozgová I, Muñoz-Viana R, Hennig L. 2017. PRC2 represses hormone-induced somatic embryogenesis in vegetative tissue of *Arabidopsis thaliana*. PLoS Genetics **13**, e1006562.

Mustroph A, Zanetti ME, Jang CJH, Holtan HE, Repetti PP, Galbraith DW, Girke T, Bailey-Serres J. 2009. Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. Proceedings of the National Academy of Sciences **106**, 18843–18848.

Nakabayashi K, Okamoto M, Koshiba T, Kamiya Y, Nambara E. 2005. Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. The Plant Journal **41**, 697–709.

Nambara E, Hayama R, Tsuchiya Y, Nishimura M, Kawaide H, Kamiya Y, Naito S. 2000. The role of ABI3 and FUS3 loci in *Arabidopsis thaliana* on phase transition from late embryo development to germination. Developmental Biology **220**, 412–423.

Narsai R, Gouil Q, Secco D, Srivastava A, Karpievitch YV, Liew LC, Lister R, Lewsey MG, Whelan J. 2017a. Extensive transcriptomic and epigenomic remodelling occurs during *Arabidopsis thaliana* germination. Genome Biology **18**, 172.

Narsai R, Secco D, Schultz MD, Ecker JR, Lister R, Whelan J. 2017b. Dynamic and rapid changes in the transcriptome and epigenome during germination and in developing rice (*Oryza sativa*) coleoptiles under anoxia and re-oxygenation. The Plant Journal **89**, 805–824.

Nee G, Xiang Y, Soppe WJ. 2017. The release of dormancy, a wake-up call for seeds to germinate. Current Opinion in Plant Biology **35**, 8–14.

Nonogaki H. 2014. Seed dormancy and germination-emerging mechanisms and new hypotheses. Frontiers in Plant Science **5**, 233.

Ogas J, Kaufmann S, Henderson J, Somerville C. 1999. PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in Arabidopsis. Proceedings of the National Academy of Sciences **96**, 13839–13844.

Pan C, Yao L, Yu L, Qiao Z, Tang M, Wei F, Huang X, Zhou Y. 2023. Transcriptome and proteome analyses reveal the potential mechanism of seed dormancy release in *Amomum tsaoko* during warm stratification. BMC Genomics **24**, 99.

Peirats-Llobet M, Yi C, Liew LC, Berkowitz O, Narsai R, Lewsey MG, Whelan J. 2023. Spatially resolved transcriptomic analysis of the germinating barley grain. Nucleic Acids Research **51**, 7798–7819.

Penfield S. 2017. Seed dormancy and germination. Current Biology **27**, R874–R878.

Rajjou L, Duval M, Gallardo K, Catusse J, Bally J, Job C, Job D. 2012. Seed germination and vigor. Annual Review of Plant Biology **63**, 507–533.

Reed RC, Bradford KJ, Khanday I. 2022. Seed germination and vigor: ensuring crop sustainability in a changing climate. Heredity **128**, 450–459.

Sano N, Lounifi I, Cueff G, Collet B, Clement G, Balzergue S, Huguet S, Valot B, Galland M, Rajjou L. 2022. Multi-omics approaches unravel specific features of embryo and endosperm in rice seed germination. Frontiers in Plant Science **13**, 867263.

Sedlazeck FJ, Lee H, Darby CA, Schatz MC. 2018. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. Nature Reviews Genetics **19**, 329–346.

Solomon MJ, Larsen PL, Varshavsky A. 1988. Mapping protein-DNA interactions *in vivo* with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell **53**, 937–947.

Stanford EE, Viehoever A. 1918. Chemistry and histology of the glands of the cotton plant, with notes on the occurrence of similar glands in related plants. Journal of Agricultural Research **8**, 688–689.

Steinbrecher T, Leubner-Metzger G. 2017. The biomechanics of seed germination. Journal of Experimental Botany **68**, 765–783.

Sun Y, Han Y, Sheng K, Yang P, Cao Y, Li H, Zhu QH, Chen J, Zhu S, Zhao T. 2023. Single-cell transcriptomic analysis reveals the developmental trajectory and transcriptional regulatory networks of pigment glands in *Gossypium bickii*. Molecular Plant **16**, 694–708.

Sybilska E, Collin A, Haddadi BS, Mur LAJ, Beckmann M, Guo W, Simpson CG, Daszkowska-Golec A. 2024. The cap-binding complex modulates ABA-responsive transcript splicing during germination in barley (*Hordeum vulgare*). Scientific Reports **14**, 18278.

Tang X, Hou A, Babu M, ***et al.*** 2008. The Arabidopsis BRAHMA chromatin-remodeling ATPase is involved in repression of seed maturation genes in leaves. Plant Physiology **147**, 1143–1157.

Tiozon RJN, Sreenivasulu N, Alseekh S, Sartagoda KJD, Usadel B, Fernie AR. 2023. Metabolomics and machine learning technique revealed that germination enhances the multi-nutritional properties of pigmented rice. Communications Biology **6**, 1000.

Tremblay BJM, Santini CP, Cheng Y, Zhang X, Rosa S, Questa JI. 2024. Interplay between coding and non-coding regulation drives the Arabidopsis seed-to-seedling transition. Nature Communications **15**, 1724.

Tsai AY, Gazzarrini S. 2012. AKIN10 and FUSCA3 interact to control lateral organ development and phase transitions in Arabidopsis. The Plant Journal **69**, 809–821.

Turck F, Roudier F, Farrona S, Martin-Magniette ML, Guillaume E, Buisine N, Gagnot S, Martienssen RA, Coupland G, Colot V. 2007. Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genetics **3**, e86.

van Berkum NL, Lieberman-Aiden E, Williams L, Imakaev M, Gnirke A, Mirny LA, Dekker J, Lander ES. 2010. Hi-C: a method to study the three-dimensional architecture of genomes. Journal of Visualized Experiments **39**, e1869.

Van Zanten M, Koini MA, Geyer R, Liu Y, Brambilla V, Bartels D, Koornneef M, Fransz P, Soppe WJJ. 2011. Seed maturation in *Arabidopsis thaliana* is characterized by nuclear size reduction and increased chromatin condensation. Proceedings of the National Academy of Sciences **108**, 20219–20224.

Verma S, Pardha V, Attuluri S, Robert HS. 2022. Transcriptional control of Arabidopsis seed development. Planta **255**, 90–90.

Wang L, Zhu Y, Jiang J, Tan G, Ma Q, Zhang H. 2023a. Dynamic changes in the levels of metabolites and endogenous hormones during the germination of *Zanthoxylum nitidum* (Roxb.) DC. Seeds. Plant Signaling & Behavior **18**, 2251750.

Wang Y, Liu B, Zhao G, Lee Y, Buzdin A, Mu X, Zhao J, Chen H, Li X. 2023b. Spatial transcriptomics: technologies, applications and experimental considerations. Genomics **115**, 110671.

Wang Y, Yao Y, Cui Y, An L, Li X, Bai Y, Ding B, Yao X, Wu K. 2024. Unveiling the mysteries of HvANS: a study on anthocyanin biosynthesis in qingke (*Hordeum vulgare* L. var. Nudum hook. f.) seeds. BMC Plant Biology **24**, 637.

Williams JH, Friedman WE. 2002. Identification of diploid endosperm in an early angiosperm lineage. Nature **415**, 522–526.

Wu T, Yang C, Ding B, Feng Z, Wang Q, He J, Tong J, Xiao L, Jiang L, Wan J. 2016a. Microarray-based gene expression analysis of strong seed dormancy in rice cv. N22 and less dormant mutant derivatives. Plant Physiology and Biochemistry **99**, 27–38.

Wu Z, Zhu D, Lin X, ***et al.*** 2016b. RNA binding proteins RZ-1B and RZ-1C play critical roles in regulating Pre-mRNA splicing and gene expression during development in Arabidopsis. The Plant Cell **28**, 55–73.

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 Physiology **142**, 1160–1168.

Xiaojian Y, Xin W, Setsuko K. 2018. Phosphoproteomics: protein phosphorylation in regulation of seed germination and plant growth. Current Protein & Peptide Science **19**, 401–412.

Yamashita K, Umezawa T. 2022. Phosphoproteomic approaches to evaluate ABA signaling. Methods in Molecular Biology **2462**, 163–179.

Yan A, Chen Z. 2020. The control of seed dormancy and germination by temperature, light and nitrate. The Botanical Review **86**, 39–75.

Yan S, Bhawal R, Yin Z, Thannhauser TW, Zhang S. 2022. Recent advances in proteomics and metabolomics in plants. Molecular Horticulture **2**, 17.

Yang D, Zhao F, Zhu D, Chen X, Kong X, Wu Y, Chen M, Du J, Qu LJ, Wu Z. 2022. Progressive chromatin silencing of ABA biosynthesis genes permits seed germination in Arabidopsis. The Plant Cell **34**, 2871–2891.

Yao J, Chu Q, Guo X, ***et al.*** 2024. Spatiotemporal transcriptomic landscape of rice embryonic cells during seed germination. Developmental Cell **59**, 2320–2332.e5.

Yao X, Yao Y, An L, Li X, Bai Y, Cui Y, Wu K. 2022. Accumulation and regulation of anthocyanins in white and purple Tibetan hulless barley (*Hordeum vulgare* L. var. nudum Hook. f.) revealed by combined *de novo* transcriptomics and metabolomics. BMC Plant Biology **22**, 391.

Yoshida H, Hirano K, Yano K, ***et al.*** 2022. Genome-wide association study identifies a gene responsible for temperature-dependent rice germination. Nature Communications **13**, 5665.

Yuan X, Jiang X, Zhang M, Wang L, Jiao W, Chen H, Mao J, Ye W, Song Q. 2024. Integrative omics analysis elucidates the genetic basis underlying seed weight and oil content in soybean. The Plant Cell **36**, 2160–2175.

Zhang L, Liu L, Li H, He J, Chao H, Yan S, Yin Y, Zhao W, Li M. 2024. 3D genome structural variations play important roles in regulating seed oil content of *Brassica napus*. Plant Communications **5**, 100666.

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. 2007. Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. PLoS Biology **5**, e129–e129.

Zhang X, Luo Z, Marand AP, Yan H, Jang H, Bang S, Mendieta JP, Minow MAA, Schmitz RJ. 2025. A spatially resolved multi-omic single-cell atlas of soybean development. Cell **188**, 550–567.e19.

Zhao L, Yang Y, Chen J, ***et al.*** 2023. Dynamic chromatin regulatory programs during embryogenesis of hexaploid wheat. Genome Biology **24**, 7.

Zhao T, Lu J, Zhang H, Xue M, Pan J, Ma L, Berger F, Jiang D. 2022. Histone H3.3 deposition in seed is essential for the post-embryonic developmental competence in Arabidopsis. Nature Communications **13**, 7728.