Seed Science Research

cambridge.org/ssr

Deciphering seed development and germination in the single-cell era

Paul Schouveiler[1,2,3], Lim Chee Liew[2,3], Massimiliano Corso[1], Loïc Rajjou[1], Julia Zinsmeister[1] and Mathew G. Lewsey[2,3,4]

[1]Université Paris-Saclay, INRAE, AgroParisTech, Institute Jean-Pierre Bourgin for Plant Sciences (IJPB), France; [2]La Trobe Institute for Sustainable Agriculture and Food, AgriBio Building, La Trobe University, Australia; [3]Australian Research Council Centre of Excellence in Plants for Space, AgriBio Building, La Trobe University, Australia and [4]Australian Research Council Industrial Transformation Research Hub for Medicinal Agriculture, AgriBio Building, La Trobe University, Australia

Methods Paper

Cite this article: Schouveiler P, Liew LC, Corso M, Rajjou L, Zinsmeister J, Lewsey MG (2025). Deciphering seed development and germination in the single-cell era. *Seed Science Research* 1–10. https://doi.org/10.1017/S0960258524000199

Received: 2 April 2024
Revised: 4 July 2024
Accepted: 10 July 2024

Keywords:
Arabidopsis thaliana; multi-omics integration; seed development; seed germination; single-cell technologies; transcriptomics

Corresponding authors:
Mathew G. Lewsey;
Email: m.lewsey@latrobe.edu.au;
Julia Zinsmeister;
Email: julia.zinsmeister@agroparistech.fr

Abstract

Seeds are complex structures that serve as dispersal units in angiosperms. Seeds consist of three specialized tissues with distinct roles and molecular compositions. Hence, the characterization of the genetic regulators that act within individual seed tissues, and how their activity changes during seed development and germination, has been a primary focus of seed research. However, our knowledge of the spatiotemporal modulation of genetic regulators within seeds, across different seed cell types, has been limited by the resolution of available techniques. In the last few years, the development and application of single-cell technologies in plants have enabled the elucidation of gene networks involved in various developmental processes at the cellular level. Some studies have applied these technologies to seeds, enabling further characterization of seed development and germination at the cellular level. Here, we review the current status of the application of single-cell technologies to seeds and present a workflow for conducting single-cell transcriptomics. Additionally, we discuss the integration of single-cell multi-omics, aiming to demonstrate the potential of single-cell technologies in enhancing our comprehension of the spatiotemporal regulations governing seed development and germination.

Introduction

Seed development and germination are critical steps in the plant life cycle, supporting seed filling and seedling establishment in fluctuating environmental conditions. In angiosperms, seed structure is established after fertilization to produce a resilient organism with three major tissues: the zygotic embryo, the endosperm and the seed coat (Bewley and Black, 1994). Angiosperm seed development starts with the double fertilization of ovule cells by sperm cells. This event also delimits the different seed tissues: the embryo arises from the fertilization of the egg cell by one of the sperm nuclei and the endosperm comes from the fusion of the other sperm nucleus with the nuclei of the two maternal embryo sac central cells. In addition, the maternal ovule integuments become the different layers of the seed coat. These three seed tissues then undergo controlled cell divisions, elongations and differentiations during successive developmental stages. These processes are responsible for the establishment of the multiple cell types of the seed that are crucial for dispersal and germination. Single-cell transcriptomics (single-cell RNA sequencing [scRNA-seq] and single-nucleus RNA-seq [snRNA-seq]) enables profiling of the transcriptome of individual cells isolated from a tissue of interest (reviewed in Shaw et al. (2021)). Recent publications have begun to show the beneficial application of single-cell methods to elucidate *Arabidopsis thaliana* seed complexity during development and germination (Kao et al., 2021; Picard et al., 2021; Lee et al., 2023; Liew et al., 2024). In this review, we discuss the potential of single-cell transcriptomics, in combination with other single-cell 'omics, to deepen our understanding of seed development and germination regulation at the cellular level.

In *A. thaliana,* seed development and germination are characterized by several major steps essential for the establishment and organization of the different seed tissues (Fig. 1A; reviewed in Verma et al. (2022)). Seed development starts with embryo morphogenesis during which the embryo develops quickly through cell divisions and elongations. It creates the plant body plan as embryonic cells specialize into different cell types. The embryo develops two cotyledons, a hypocotyl and a radicle which can respectively be considered as seed leaves, stems and roots. In parallel, a syncytial stage, characterized by nuclear divisions with no cytokinesis, occurs in the endosperm. This is then rapidly followed by endosperm cellularization and differentiation. In addition, the five distinct seed coat cell layers originating from maternal ovule integuments go through rapid growth with cell divisions and elongations. Embryo

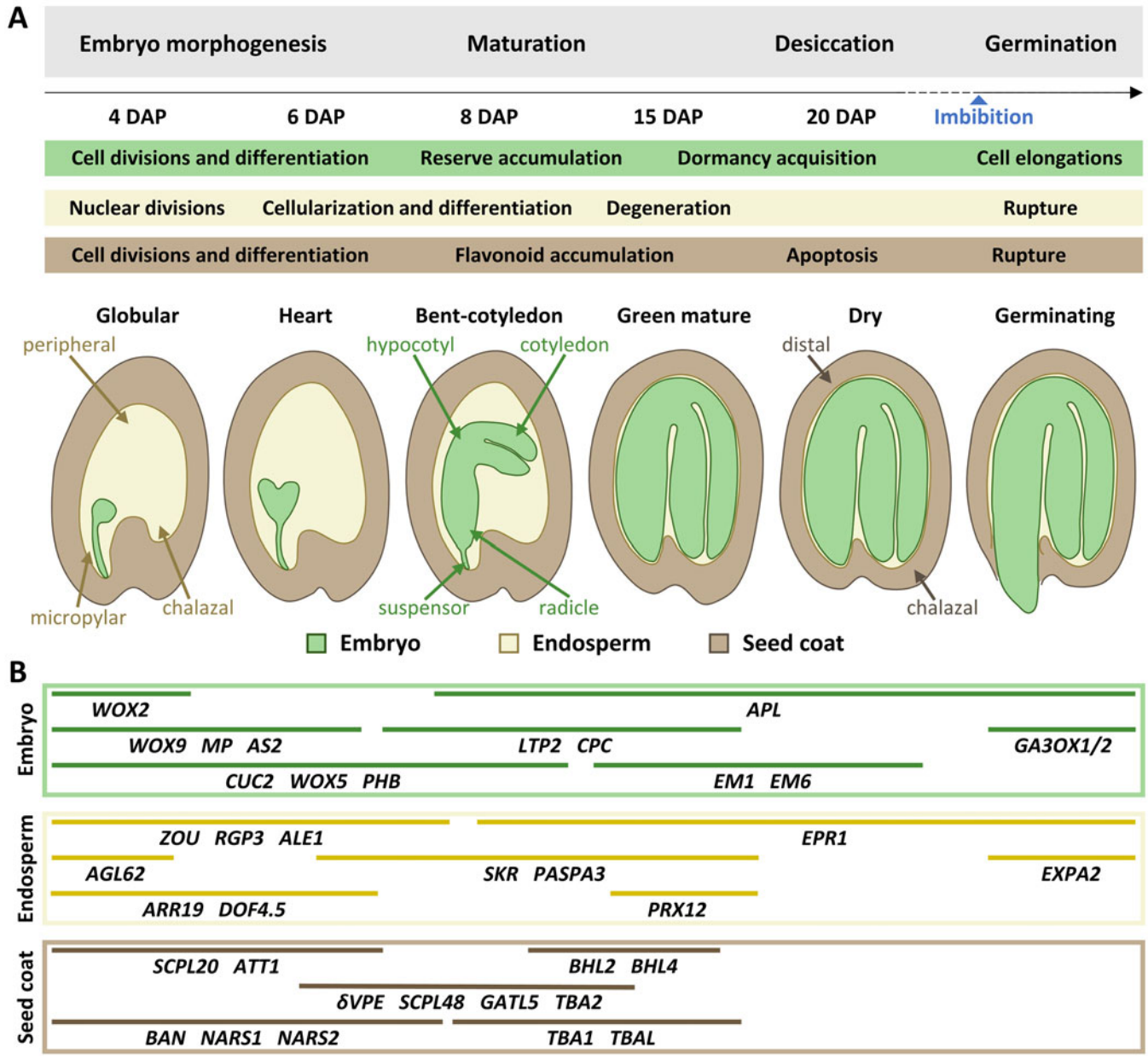

Figure 1. Major stages of seed development and germination, and genetic markers associated with these stages. (A) Seed development and germination stages, and biological processes occurring in the embryo (in green), the endosperm (yellow) and the seed coat (brown). Chronology of biological processes was retrieved from previous publications (Debeaujon et al., 2007; Verma et al., 2022). Represented timeline is indicative and depends on growth conditions (DAP: days after pollination). Subregions of the endosperm, the embryo and the seed coat are, respectively, indicated at globular, bent-cotyledon and dry stages. (B) Examples of genetic markers specifically expressed during some stages of seed development and germination in the embryo (in green), the endosperm (yellow) and the seed coat (brown). A precise description of their spatiotemporal distribution and relevant references can be found for each marker in Supplementary Table S1.

morphogenesis is followed by seed maturation that leads to a dry seed filled with reserve compounds, mainly lipids and proteins, and specialized metabolites (reviewed in Holdsworth et al. (2008)). Maturation is characterized by embryo growth arrest and reserve accumulation in embryonic cotyledons. The endosperm degenerates into a single-cell layer (also referred to as the aleurone layer) during maturation. In the seed coat, the different cell layers differentiate into distinct cell types, which are mostly associated with the accumulation of specialized metabolites, such as tannins and other compounds, in specific layers (reviewed in Debeaujon et al. (2007)). Later, during seed development, seed coat cells die and the different layers crush together. During maturation, seeds become tolerant to subsequent desiccation that corresponds to a drastic water loss in the different seed tissues.

A. thaliana seeds also acquire dormancy, which is a physiological inhibition of germination despite favourable environmental conditions (reviewed in Finch-Savage and Leubner-Metzger (2006)). The endosperm responds to a range of signals such as nitric oxide, gibberellic acid (GA) and abscisic acid (ABA), serving as a critical component for the establishment and maintenance of seed dormancy. The seed developmental steps define seed composition and prepare seeds for dispersal and ultimately germination. Seeds become competent to germinate when

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press

dormancy is released; usually after a few weeks (or months) of dry storage, depending upon the accession of *A. thaliana* and the environmental factors. Under favourable conditions, the germination process begins with water uptake by quiescent and non-dormant dry seeds. Germination ends with seed coat and endosperm rupture, followed by radicle protrusion in the micropylar region (also known as germination *sensu stricto*; Fig. 1A; Bewley and Black, 1994; reviewed in Weitbrecht et al. (2011)). In *A. thaliana*, germination *sensu stricto* completes through elongation in the adjacent transition zone and lower hypocotyl, driven by cell expansion and an increase in nuclear DNA content, not by cell division (Sliwinska et al., 2009; Bassel et al., 2014). Reserves accumulated in cotyledons are used for the resumption of cell division resulting in seedling establishment.

Overall, the study and characterization of seed developmental stages highlight the complexity of *A. thaliana* seed tissues. These tissues contain multiple cell types that are associated with specific composition and, hence, various biological functions. Additionally, seed cell types rapidly go through different cell states during seed development and germination. A cell type can be defined as a group of cells sharing molecular properties and originating from the same developmental programme, while cell states represent variations in molecular phenotypes within a cell type (reviewed in Domcke and Shendure (2023)). In seeds, the different cell types show very distinct and dynamic transcript, protein and metabolite profiles (Belmonte et al., 2013; Barreda et al., 2024). As an example, the seed coat is divided into five different cell layers with different specialized roles that are associated with distinct metabolite accumulations (reviewed in Haughn and Chaudhury (2005)). Seed coat can be further subdivided into two regions: the chalazal region, which connects maternal vascular tissues to the seed, and the distal region, both having very distinct transcriptional profiles (Fig. 1A; Khan et al., 2014). Thus, appropriate technologies should be used to capture molecular activity and changes in individual cells of seeds during seed development and germination.

In the following section, we will show the potential of single-cell transcriptomics to decipher the cell types and states of seed development and germination, and therefore, to address crucial questions in seed biology. We first discuss the advantages of scRNA-seq compared with traditional genetic methods by presenting the few single-cell studies that have been carried out so far in *A. thaliana* seed. Then, we describe a standard workflow for such a single-cell transcriptomic study, and we highlight the added value of integrating this technology with other single-cell and spatial 'omics approaches. This review focuses on *A. thaliana* seed development and germination, but studies carried out in seeds of other plant species are also mentioned.

Methods to elucidate the genetic actors of seed development and germination

In *A. thaliana*, the distinct stages and tissues of seed development and germination are sequentially regulated by complex genetic networks (reviewed in Carrera-Castaño et al. (2020), Baud et al. (2022), and Verma et al. (2022)). In particular, *WUSCHEL-RELATED HOMEOBOX* (*WOX*) transcription factors are key regulators of early embryo morphogenesis controlling body axis establishment (Haecker et al., 2004). Later, in seed development, *LAFL* (*LEAFY COTYLEDON1*, *ABSCISIC ACID INSENSITIVE3*, *FUSCA3*, and *LEAFY COTYLEDON2*) transcription factors trigger seed maturation, while the MBW (MYB, bHLH and WDR) protein complex is crucial for the biosynthesis and accumulation of flavonoid compounds in the seed coat (reviewed in Lepiniec et al. (2018), Baud et al. (2022), and Zumajo-Cardona et al. (2023)). ABA-related genes are known to be central regulators of desiccation tolerance and primary dormancy (reviewed in Nambara and Marion-Poll (2003) and Chauffour et al. (2019)). Additionally, *DELAY OF GERMINATION 1* (*DOG1*) has been identified as a key regulator of seed dormancy acting through both ABA-dependent and ABA-independent pathways (Bentsink et al., 2006; Née et al., 2017). Lastly, germination is controlled by the GA pathway, and in particular, DELLA proteins that act as negative regulators of seed germination (Tyler et al., 2004). These genetic regulators can either be expressed in a specific or in several seed cell types (Belmonte et al., 2013; Peirats-Llobet et al., 2023). Hence, it is crucial to study the spatial and temporal expression dynamics of these actors to better understand the regulation of the processes of the successive stages of seed development and germination. This is a challenging undertaking as seeds are small and progressively accumulate protective layers, such as flavonoids in the seed coat. Methods, such as RNA *in situ* hybridization (ISH) or promoter–reporter fusion assays, have been used to study the spatiotemporal activity of seed gene networks (Haecker et al., 2004; Francoz et al., 2016; Palovaara et al., 2017). Both these methods allow localization of the expression of a gene of interest; by using a labelled complementary RNA probe for ISH or by cloning the promoter of this gene in front of a reporter gene, usually β-glucuronidase (GUS) or green fluorescent protein (GFP) gene, for promoter–reporter fusion assays. However, these methods only offer low throughput and consequently only permit the study of a limited number of genes. In addition, these methods require a prior knowledge of studied genes and do not allow novel gene discovery.

More recently, comprehensive techniques, such as bulk RNA-seq, have also been performed on seeds (Le et al., 2010; Belmonte et al., 2013; Dekkers et al., 2013; Khan et al., 2014; Narsai et al., 2017; Hofmann et al., 2019). Bulk RNA-seq allows the study of global gene expression in the whole seed or in a seed tissue at different stages. Related to this, individual seed tissues can be isolated either manually or by performing laser-capture microdissection which uses laser cutting to extract areas of interest from a biological sample (Dekkers et al., 2013; Betts et al., 2017; Liew et al., 2020). However, bulk RNA-seq does not allow the capture of cell type heterogeneity and transcript signatures within the studied sample. Hence, this technique provides average transcriptome information across a number of cells, and its spatial resolution is limited by the precision of dissection. Additionally, the isolation of seed tissues by laser-capture microdissection is a challenging and time-consuming process that does not allow high throughput.

In the last few years, single-cell transcriptomics methods (scRNA-seq and snRNA-seq) have been developed and applied to several plant organs, including *A. thaliana* roots (Denyer et al., 2019; Shulse et al., 2019) and leaves (Kim et al., 2021; Lopez-Anido et al., 2021), as well as maize and sorghum roots (Guillotin et al., 2023). Single-cell technologies allow us to study the transcriptomic profile of isolated cells and its dynamics during plant development or in response to environment changes.

Several recent studies have demonstrated how scRNA-seq and snRNA-seq can be applied to *A. thaliana* seeds (Kao et al., 2021; Picard et al., 2021; Lee et al., 2023; Liew et al., 2024). Single-cell transcriptomics is a very promising approach for analysis of the many cell types and cell states in *A. thaliana* seeds, with the potential to be applied at different stages of seed development

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press

and germination. For example, snRNA-seq has successfully been performed on *A. thaliana* embryos at the globular stage (Kao et al., 2021). This was done by isolating embryo-specific nuclei using fluorescent activated cell sorting (FACS) in a transgenic line expressing nuclear-localized GFP under the control of the embryo-specific *WOX2* promoter. This enabled analysis of spatial genetic regulation during early *A. thaliana* embryogenesis. At this stage, the different embryonic cell types specialize and establish a basic body plan. Single-nucleus transcriptomics allowed gene expression programmes and transcription factor signatures to be associated with these differentiating cell types. Interestingly, differences in the expression of epigenetic regulators were observed across embryonic cell types. For instance, genes involved in 24-nucleotide small interfering RNAs (siRNAs) biogenesis were preferentially expressed in the suspensor which connects maternal tissues with the embryo, while transcripts involved in siRNA binding and siRNA-dependent gene silencing were enriched in the embryo, and more specifically in shoot meristem initials (Gutzat et al., 2020). This suggests a signalling coordination between seed cell types, with a different site for siRNA production (the suspensor) and response (shoot meristem initials). It has been previously proposed that this mechanism might be involved in the repression of transposon mobility, thus limiting their mutagenic potential in the shoot meristem (reviewed in Mosher and Melnyk (2010)). Hence, this study illustrates how single-cell technology can be used to investigate transcriptional programmes as well as cell communications, and, in particular, signalling pathways involving several cell types, in seeds.

Similarly, a single-nucleus study was carried out on seeds 4 d after pollination and further selected endosperm-specific nuclei based on maternal read percentage (Picard et al., 2021). Two different *A. thaliana* accessions (Col-0 and Cvi) were crossed to study imprinted genes in the endosperm cell subtypes. This revealed the expression patterns of imprinted genes across endosperm cell types, including, notably, that paternally imprinted genes (PEGs) were predominantly expressed in the chalazal endosperm, while maternally imprinted genes (MEGs) were expressed across all endosperm cell types. This study also better characterized transcriptomic profiles in specific endosperm cell types. In particular, genes involved in DNA replication were preferentially expressed in the chalazal endosperm, while no mitotic markers were found, suggesting that this region undergoes endoreduplication. This is consistent with previous observations of large and polyploid nuclei in this compartment (Baroux et al., 2004). Additionally, different transcriptomic profiles were observed within the chalazal endosperm nuclei. The authors further demonstrated that these variations are associated with a subdifferentiation within the chalazal endosperm, which consists of two regions: the chalazal nodules, which delimits the chalazal endosperm, and the chalazal cyst, which interacts with the seed coat. By applying single-cell technology, this study provided the first transcriptomic profiles of specifically of these specific regions (Nguyen et al., 2000).

scRNA-seq was also applied to germinating *A. thaliana* embryos, enabling the discovery that most embryonic cells go through a common transcriptional state at the beginning of germination (Liew et al., 2024). This study enabled the observation that suites of cell-type-specific transcription factors were involved in subsequent transcriptional remodelling, contributing to the establishment of cell-type-specific patterns of gene expression. A subset of these transcription factors was also validated as an influencing germination rate. In addition, the study supported an earlier investigation, suggesting that the provasculature might play a key role in determining the break of seed dormancy (Topham et al., 2017).

Recently, a large-scale snRNA-seq study was carried out on *A. thaliana* organs and tissues at different development stages, including siliques that contain developing seeds and two time points during germination (Lee et al., 2023). Seed cell types, and in particular embryonic ones, were identified in snRNA-seq carried out on siliques. This latter study illustrates the opportunity to study seed development by performing a single-cell analysis either on whole siliques, and to further select seed-specific cells, or on isolated developing seeds. Interestingly, siliques were found to be the organ with the greatest complexity among all *A. thaliana* organs and tissues analysed in this study, indicating the existence of multiple seed and silique cell types and cell states. Furthermore, snRNA-seq was conducted to characterize cell profiles of dissected maize endosperm, showing that this technique can also be transferred to seeds of other plant species including crops (Yuan et al., 2024).

The establishment of single-cell transcriptomics methods for plant tissues has been accompanied by the development of spatial transcriptomics methods, which allow *in situ* transcriptome profiling of sectioned samples (Xia et al., 2022). This technology is short-read sequencing based and offers comprehensive transcriptome analysis. The resolution of spatial transcriptomics depends on the technology used (e.g. 10 µm for Slide-seqV2; 55 µm for 10× Genomics Visium), which is sufficient to profile transcripts in individual cells or group of cells (Ståhl et al., 2016; Stickels et al., 2021). Resolution is constantly improving, so in future, spatial transcriptomics might be used to profile the transcriptome at subcellular levels (Oliveira et al., 2024). The advantage of spatial transcriptomics is that the technologies offer direct spatial information as there is no tissue disruption such as for single-cell studies.

Spatial transcriptomics has recently been performed on germinating barley seeds at six time points, elucidating the spatiotemporal regulation of over 14,000 genes expressed in the first 24 h after imbibition (Peirats-Llobet et al., 2023). This strategy has also been applied to developing wheat grains, as well as to nectarine and peach fruit including their seed (Zhao et al., 2023; Li et al., 2024). However, the resolution of spatial transcriptomics is still low compared with single-cell techniques, which is a major limitation when studying small tissues such as *A. thaliana* seeds (about 500 µm). Alternatively, imaging-based spatial transcriptomics methods have also been adapted for use on plants, such as Multiplexed Error-Robust Fluorescence *in situ* Hybridization (MERFISH; Lee et al., 2023; Nobori et al., 2023). These techniques allow an *in situ* profiling of hundreds of selected target genes at a subcellular resolution using fluorescent probes but require prior knowledge of these target genes. Overall, single-cell and spatial transcriptomics are two complementary approaches to perform transcriptomic profiling of cell types and states in a tissue of interest. These technologies can be integrated to provide a comprehensive spatial description with a precise resolution (Lee et al., 2023).

The above-mentioned studies on seed development and germination show the potential of single-cell transcriptomics to enhance our understanding of seed biology. Single-cell technology enables the establishment of gene expression atlas of seed tissues and stages that can be further utilized to investigate transcriptional programmes. These first results also illustrate how scRNA-seq can both precisely define very specific cell types

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press

(such as the chalazal endosperm; Picard et al., 2021) and unveil common transcriptional states (during germination; Liew et al., 2024). However, at this time, no single-cell transcriptomic study has been carried out to comprehensively cover the whole seed during the different stages of its development. Further single-cell studies will provide more insights into the spatiotemporal regulations of gene expression in *A. thaliana* seeds, as well as in other plant species, and the diversity and interaction among seed cell types. Additionally, complementing single cell with spatial transcriptomics will refine the spatial dimension in these analyses, and increase our comprehension of seed structure. These will certainly be facilitated by the development and the decreasing costs of single cell and spatial technologies in the coming years, enabling an increase in the number of studies in seed research. In the next section, we propose a workflow for scRNA-seq to facilitate future single-cell analyses of seeds.

Workflow for a single-cell transcriptomics study

The workflow of single-cell studies starts with sample preparation during which cells (scRNA-seq), or nuclei (snRNA-seq) are extracted from the tissue of interest (Fig. 2; reviewed in Shaw et al. (2021)). In *A. thaliana*, protoplast extraction has commonly been used to isolate cells by removing cell walls, but nucleus extraction is increasing in popularity (Yoo et al., 2007; Wang et al., 2023). Extraction of protoplasts or nuclei is followed by their separation using microfluidic technologies on droplet-based platforms such as 10x Genomics or DropSeq (Macosko et al., 2015; Zheng et al., 2017). Individual protoplasts or nuclei are encapsulated within droplets, and the transcripts within these droplets are then utilized for library preparation. This process involves reverse transcription, labelling with unique molecular identifiers and cell-specific barcodes, and subsequent amplification. Barcoding is a crucial step that enables the association of reads with their corresponding transcripts and cells of origin in subsequent analyses. The library is then sequenced, and raw reads generated for each cell are processed in a bioinformatic pipeline.

The bioinformatic pipeline starts with read mapping to the relevant species' genome and data preprocessing. Cell quality control is performed during preprocessing. Low-quality cells are removed, which include damaged cells as well as droplets containing multiple cells or no cells (empty droplets). Further processing steps are then usually performed with the R toolkit Seurat (Hao et al., 2024). Seurat has been specifically developed for single-cell transcriptomics processing and analyses. This toolkit takes as input a raw count matrix in which the numbers of transcripts detected in each cell are stored. Several processing steps, such as data normalization and scaling, are usually performed on the raw counts. Then, Seurat can be used to cluster cells using their transcriptomic signatures, based on dimensionality reduction using methods such as principal component analysis (PCA) or uniform manifold approximation and projection (UMAP; Jolliffe, 2011; McInnes et al., 2020). Obtained clusters are next annotated with the help of known marker genes expressed in specific cell types or states. These annotated clusters can be used for subsequent analyses. It is for instance possible with Seurat to extract information on specific genes or cell types. Other tools can also be used to infer cell-type-specific gene co-expressions and regulatory networks from single-cell transcriptomics (Bravo González-Blas et al., 2023; Su et al., 2023). Alternatively, a pseudotime trajectory analysis can be carried out on single-cell data with tools such as Monocle 3.0 (Cao et al., 2019). Pseudotime studies arrange cells based on their transcriptional states and consequently infer a developmental progression. Additional studies can also be performed on single-cell data such as gene ontology enrichment analysis or transcription factor predictions (Lin et al., 2020; Liew et al., 2024).

Challenges for single-cell transcriptomics in seeds

One of the crucial steps of single-cell transcriptomics is sample preparation, which requires careful method design. In particular, a choice between protoplast (scRNA-seq) and nucleus (snRNA-seq) isolation has to be made beforehand. scRNA-seq and snRNA-seq produce comparable results, with each method having its own strengths and drawbacks (Guillotin et al., 2023; Wang et al., 2023). The primary drawback with scRNA-seq is that it presents a bias in capturing some cell types, particularly because of the diverse sizes of plant cells and the variable efficiency of getting protoplasts from different cell types. Additionally, protoplast extraction normally takes a longer time than nuclei extraction and utilizes enzymatic digestion which can trigger changes in gene expression due to stress responses (Birnbaum et al., 2003). On the other hand, snRNA-seq only captures nuclear mRNAs, leading to lower transcript detection due to the inability to capture cytoplasmic mRNAs. These drawbacks should be taken into consideration when designing a single-cell workflow.

So far, in seeds, snRNA-seq has proven its efficiency to profile all cell types of seed development and germination in *A. thaliana* (Kao et al., 2021; Picard et al., 2021; Lee et al., 2023). Additionally, snRNA-seq seems more appropriate for early seed development to capture isolated nuclei from the syncytial endosperm (Picard

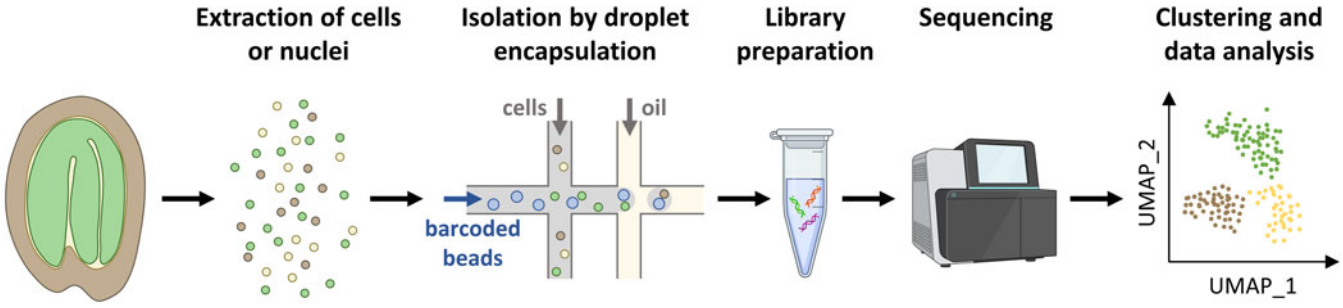

Figure 2. Workflow for single-cell transcriptomics. Cells or nuclei are first extracted from a tissue of interest and isolated by droplet encapsulation. This is coupled with cell- or nucleus-specific tagging of the transcripts with the use of barcoded beads. From these tagged transcripts, a library is prepared and sequenced. In the last step, cells are clustered based on their transcriptomic profile and this clustering serves as a basis for subsequent data analyses.

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press

et al., 2021). However, scRNA-seq could be more relevant for late seed development and germination, as transcripts are progressively stored in cell cytoplasm until germination (Nakabayashi et al., 2005). It has furthermore been demonstrated that transcription is not essential for radicle elongation during germination, emphasizing the significant role of stored mRNAs (Rajjou et al., 2004). The selective translation of a subset of stored mRNAs is a critical regulatory mechanism that allows seeds to rapidly germinate, independent of *de novo* transcription (Galland et al., 2014; Bai et al., 2020; reviewed in Sano et al. (2020)). The specific targeting of certain mRNAs for translation appears to involve both sequence and structural features. However, there is currently a gap in research regarding the spatial regulation of this translational selectivity at the cellular level within seeds. Protoplast isolation from seed tissues can nonetheless be challenging because of seed coat protective layers hindering a homogenous digestion. So far, scRNA-seq has only been used to profile cells of the germinating embryo, and the ability to extract protoplasts from the seed coat or endosperm is still to be demonstrated. Therefore, if the protoplast method is used, a protocol must be optimized that efficiently captures all the different seed cell types. By comparison, spatial transcriptomics has been efficiently utilized to capture all seed development and germination cell types of various plant species (Lee et al., 2023; Peirats-Llobet et al., 2023; Li et al., 2024). However, it is worth noting that this technique is still very costly and is usually applied to a limited number of sections, meaning that the data represent only a portion of a seed.

Another crucial challenge of single-cell transcriptomics is the cell clustering step and correct annotation of obtained clusters, which is essential to perform downstream analyses. The annotation relies on previously selected marker genes, which usually come from existing publications where specific cells or tissues have been analysed, including prior single-cell studies. Liew et al. (2024) curated such a list of embryonic marker genes for seed germination. However, as no single-cell study has been previously carried out across the entirety of seed development, there is no list of marker genes demonstrated to be spatially associated with the different cell types and stages of *A. thaliana* seed development. Thus, we have compiled a list of seed development marker genes to facilitate cluster annotation of future single-cell studies (Fig. 1B; Supplementary Table S1). To do so, genes demonstrated to be specifically associated with seed cell types and stages were selected. Data arise from RNA ISH or promoter–-reporter fusion assays, as well as previous single-cell studies of seed tissues and germination (Francoz et al., 2016; Kao et al., 2021; Picard et al., 2021; Liew et al., 2024). We obtained a list of 369 genes associated with one or a limited number of seed cell types and with various developmental stages (Supplementary Table S1). Seed markers obtained from bulk RNA-seq were not included as they are likely less precise, but these could potentially be used to cross-validate cluster identity (e.g. datasets from Le et al. (2010) and Belmonte et al. (2013)). The list of seed markers might also be complemented by the addition of marker transcripts obtained in other single-cell studies that precisely annotate specific cell types. For instance, to better characterize developing vascular tissues in seed embryos, vasculature markers curated by Kim et al. (2021) could be implemented.

Integration of single-cell multi-omics techniques

In addition to single-cell transcriptomics, several other modalities of single-cell technology have been recently developed in *A. thaliana* and other plant species. For instance, these include single-cell epigenomics, metabolomics and proteomics (Farmer et al., 2021; Marand et al., 2021; Li et al., 2023; Montes et al., 2024). Additionally, spatial metabolomics has also been applied to plants, enabling precise profiling of feature distribution in tissues (Yamamoto et al., 2016). Spatial 'omics now displays cellular resolution, thus providing complementary information to single-cell technologies. However, single-cell and spatial methods have so far been mostly restricted to plant vegetative tissues.

Epigenomic analysis has been carried out at single-cell resolution in plant species. In particular, chromatin accessibility has been mapped in isolated nuclei from *A. thaliana* roots using a single-nucleus Assay for Transposase-Accessible Chromatin sequencing (snATAC-seq; Farmer et al., 2021). To perform snATAC-seq, nuclei are first extracted from a tissue of interest and open chromatin is cleaved and tagged by a transposase. Nuclei are next isolated and transposed DNA fragments are barcoded, with similar methods as for single-cell transcriptomics. Sequencing and analysis of these DNA fragments enables nucleus-specific profiling of the open chromatin in the studied tissue. Hence, snATAC-seq can be used to understand the regulatory landscape of individual cells as well as to identify cell types based on their chromatin accessibility profiles. Single-nucleus epigenomics has not been performed yet in seeds but such analyses would be likely to provide significant new biological insights because of the uniquely dynamic nature of chromatin conformation in seeds across development and germination. For example, it is known that there is allele-specific silencing involved in endosperm parental imprinting (Picard et al., 2021). Nuclear condensation also occurs during seed maturation and is followed by a global epigenomic remodelling when germination takes place (van Zanten et al., 2011; Narsai et al., 2017). These different processes involve chromatin structural changes, and single-cell epigenomics would provide further insights into the relationships between chromatin configuration and reactivation of the genome at the cellular level.

Recently, single-cell metabolomics has also been developed in plants, allowing the quantification of metabolites at single-cell resolution with high throughput (Li et al., 2023). This method relies on protoplast isolation and separation by droplet microfluidics, followed by mass spectrometry (MS). Isolated protoplasts are clustered by PCA, and cluster identity can be determined based on specific metabolomic signatures. Single-cell metabolomics could, therefore, also be helpful to profile metabolite accumulations in different seed cell types. Indeed, it is known that seed tissues are associated with specific metabolite profiles, and single-cell technologies would allow a more precise characterization of their distribution (Zhang et al., 2021). Similarly, single-cell proteomics has also been established in *A. thaliana* roots based on comparable technologies, enabling the identification of 3,763 proteins in 756 root cells (Montes et al., 2024).

Another technology to profile metabolite distribution in plant tissues is spatial metabolomics. Spatial metabolomics relies on MS imaging, which employs either a laser or an electrospray to ionize molecules at specific spots of sectioned tissue. Several technologies are used to perform ionization, the most popular one being matrix-assisted laser desorption/absorption (MALDI). It is then possible to obtain a spatial mapping of the primary or specialized metabolites, or plant hormones, of interest with a resolution up to a few micrometres which is close to single-cell resolution (Kompauer et al., 2017). MALDI MS was, for example, used to quantify ABA in immature common bean seeds, revealing that

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press

it mainly localizes in the embryo (Enomoto et al., 2018). However, MS imaging provides low throughput since it allows the selection of a limited number of ions and is then giving information on a reduced set of metabolites. In addition, the small size of *A. thaliana* seeds represents a limitation to conduct MS imaging with current technologies. Overall, there is a growing interest of single-cell and spatial 'omics in plant research, but only a few studies have applied them so far on seeds. One major challenge of seed research would be to adapt these methods to developing and germinating seeds of various species, including *A. thaliana*.

In addition to performing single-cell and spatial analyses individually, these technologies can also be combined in integrated multi-omic studies (reviewed in Vandereyken et al. (2023)). This approach provides a broad overview of cellular processes in a biological object. Multi-omics integration can either refer to the parallel use of several technologies on a unique sample or to the combined analysis of datasets obtained from multiple 'omics on different samples from the same tissue. For instance, it is possible to perform simultaneous snRNA-seq and snATAC-seq on the same isolated nuclei (Farmer et al., 2021). Alternatively, single-cell and spatial 'omics can also be carried out separately on isolated cells and a section of the tissue of interest, with the subsequent integration of resulting datasets (Lee et al., 2023).

While the integration of single-cell multi-omics techniques is very promising to study biological processes and characterize metabolic pathways, very few publications in plant sciences have presented such single-cell multi-omics due to its challenging nature. Combined snRNA-seq and snATAC-seq have been carried out in *A. thaliana* root and leaf as well as on rice and sorghum shoot (Farmer et al., 2021; Nobori et al., 2023; Swift et al., 2023; Liu et al., 2024). The integrated scRNA-seq and single-cell metabolomics were performed for the first time in the medicinal plant *Catharanthus roseus* and allowed a better characterization of the monoterpene indole alkaloid pathway, including the identification and validation of new genes involved in the biosynthesis of these compounds (Li et al., 2023). Lastly, the integration of single-cell with spatial 'omics has so far been mostly limited to transcriptomics (Lee et al., 2023). Such single-cell multi-omic analyses would be very helpful to better understand *A. thaliana* seed development and germination and to characterize non-genetic regulators. In particular, the phytohormones ABA and GA are, respectively, known to regulate seed development and to trigger germination (reviewed in Shu et al. (2016)). An integrated multi-omic study performed using bulk sample analyses allowed the successive roles of ABA to be deciphered during *A. thaliana* seed development and germination (Chauffour et al., 2019). Thus, transitioning from bulk analyses to single-cell and spatial technologies would offer enhanced resolution and help to fine-tune the spatiotemporal roles of ABA and GA during *A. thaliana* seed development and germination. More generally, the integrated multi-omic studies of the seed at single-cell resolution would allow deeper comprehension of the different roles of seed tissues and their subdomains during seed development and germination. This would enable precise profiling of gene expression, chromatin accessibility and metabolite accumulation in the various cell types and states of seeds.

Several computational tools have been developed to perform single-cell and spatial multi-omic data integration (reviewed in Vandereyken et al. (2023)). To give some examples, the integration of snRNA-seq and snATAC-seq can be carried out with Seurat (Hao et al., 2024). The R package spacexr can be used to integrate single-cell and spatial transcriptomic data (Cable et al., 2022; Li et al., 2022). So far, there is no computational tool specifically designed to integrate single-cell transcriptomics with spatial or single-cell metabolomics. However, it is still possible to extract information from transcriptomic and metabolic data annotated as similar seed cell types, and subsequently perform integration on averaged profiles by using tools such as the mixOmics R package (Welham et al., 2023). Overall, the integrated single multi-omics is a new field in plant science, for which new tools are being developed and are needed.

Perspectives

Single-cell studies are an emerging trend in plant research, providing comprehensive insights into plant cell types and states. In contrast, intensive research in model animal species over the last 15 years has led to the establishment of protocols, analysis pipelines and well-defined cell-type marker genes for single-cell analyses (reviewed in Jovic et al. (2022)). In plants, the extent of application of single-cell studies has varied depending on the growth stage or organ. *A. thaliana* roots and leaves have been quite widely studied through single-cell 'omics, but by comparison *A. thaliana* seed has been analysed much less extensively. This is likely because of the specific composition of seeds, which makes extracting protoplasts or nuclei from them difficult. In this review, we have summarized what has been achieved so far and proposed tools to study seeds at the single-cell level. Further developing single-cell and spatial technologies for seeds could help to answer key questions about seed development and germination: what are the different seed cell types and states, and to what extent do they communicate between each other? What is the role of cell specialization in producing self-sufficient and resilient seeds that will give rise to the next generation? How the various seed cell types individually contribute to germination? And what are the cellular mechanisms underlying genome reorganization during germination?

However, the multiplication of the number of single-cell studies poses challenges. First, there is a need to standardize single-cell techniques to be able to compare the different studies performed on the same or on different *A. thaliana* tissues. This idea was included in the recently proposed Plant Cell Atlas, which aims to comprehensively identify and map all plant cell types (Plant Cell Atlas Consortium et al., 2021). Standardization is, however, complex as each plant tissue presents its own specificities and challenges when single-cell experiments are conducted, due to the distinct biochemistry and composition of different cell types. Recently, Grones et al. (2024) proposed guidelines to set up experimental and computational workflows for single-cell transcriptomics. These recommendations could serve as a basis in the single-cell standardization process. One key, but difficult, objective in standardization is to obtain similar cell clusters from separate studies of the same tissue so that independent studies may be compared. We believe that using established lists of marker genes specific to *A. thaliana* cell types will contribute towards single-cell standardization. Such lists will likely be first curated manually, as done previously for different plant tissues and as we propose here for seed development (Supplementary Table S1). Those lists should also be evolvable by additions of new marker genes provided by future studies or machine learning (Yan et al., 2022). Finally, single-cell studies provide huge amounts of new data on *A. thaliana* cell types, for which new methods of computational analysis, documenting and storage

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press

are needed. Thus, another great challenge of single-cell 'omics is to extract relevant information from the datasets that are generated. Such an enterprise will be facilitated by the increasing number of studies that are being performed and the development of new tools in plants.

Overall, the application of single-cell 'omics to seeds will allow the community to gain a deeper understanding of seed development and germination at the cellular level. Single-cell studies on seeds have so far been mainly limited to *A. thaliana* and have focused on transcriptomics. Therefore, technical development is required for other plant species and other modalities, which might present new challenges. Application of single-cell 'omics to crop species could be particularly useful to enrich our knowledge on crucial agronomical traits, such as grain yield, content and vigour as well as germination efficiency.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0960258524000199.

Funding statement. This work was supported by grants from the Australian Research Council to M.G.L. (CE230100015, DP220102840 and IH180100006). The IJPB benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007). Part of the work presented in this review is supported by a grant from the French national funding agency 'Agence Nationale de la Recherche' to L.R. and M.C. (ANR; 'SUCSEED' project, 20-PCPA-0009 and 'RNASEED' project, 21-CE20-0034) and by a grant from the Labex Saclay Plant Sciences – SPS to M.C. ('Seed MetSpe' project, ANR-10-LABX-0040-SPS).

Competing interests. The authors declare that they have no competing interests.

References

Bai B, van der Horst S, Cordewener JHG, America TAHP, Hanson J and Bentsink L (2020) Seed-stored mRNAs that are specifically associated to monosomes are translationally regulated during germination. *Plant Physiology* **182**, 378–392.

Baroux C, Fransz P and Grossniklaus U (2004) Nuclear fusions contribute to polyploidization of the gigantic nuclei in the chalazal endosperm of *Arabidopsis*. *Planta* **220**, 38–46.

Barreda L, Brosse C, Boutet S, Perreau F, Rajjou L, Lepiniec L and Corso M (2024) Specialized metabolite modifications in Brassicaceae seeds and plants: diversity, functions and related enzymes. *Natural Product Reports* **41**, 834–859.

Bassel GW, Stamm P, Mosca G, Barbier de Reuille P, Gibbs DJ, Winter R, Janka A, Holdsworth MJ and Smith RS (2014) Mechanical constraints imposed by 3D cellular geometry and arrangement modulate growth patterns in the *Arabidopsis* embryo. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8685–8690.

Baud S, Corso M, Debeaujon I, Dubreucq B, Job D, Marion-Poll A, Miquel M, North H, Rajjou L and Lepiniec L (2022) Recent progress in molecular genetics and omics-driven research in seed biology. *Comptes Rendus Biologies* **345**, 61–110.

Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, Le BH, Drews GN, Brady SM, Goldberg RB and Harada JJ (2013) Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E435–E444.

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

Betts NS, Berkowitz O, Liu R, Collins HM, Skadhauge B, Dockter C, Burton RA, Whelan J and Fincher GB (2017) Isolation of tissues and preservation of RNA from intact, germinated barley grain. *The Plant Journal* **91**, 754–765.

Bewley JD and Black M (1994) *Seeds: Physiology of Development and Germination*, 2nd Edn. New York, NY: Springer.

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW and Benfey PN (2003) A gene expression map of the *Arabidopsis* root. *Science* **302**, 1956–1960.

Bravo González-Blas C, De Winter S, Hulselmans G, Hecker N, Matetovici I, Christiaens V, Poovathingal S, Wouters J, Aibar S and Aerts S (2023) SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nature Methods* **20**, 1355–1367.

Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F and Irizarry RA (2022) Robust decomposition of cell type mixtures in spatial transcriptomics. *Nature Biotechnology* **40**, 517–526.

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C and Shendure J (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502.

Carrera-Castaño G, Calleja-Cabrera J, Pernas M, Gómez L and Oñate-Sánchez L (2020) An updated overview on the regulation of seed germination. *Plants* **9**, 703.

Chauffour F, Bailly M, Perreau F, Cueff G, Suzuki H, Collet B, Frey A, Clément G, Soubigou-Taconnat L, Balliau T, Krieger-Liszkay A, Rajjou L and Marion-Poll A (2019) Multi-omics analysis reveals sequential roles for ABA during seed maturation. *Plant Physiology* **180**, 1198–1218.

Debeaujon I, Lepiniec L, Pourcel L and Routaboul J-M (2007) Seed coat development and dormancy. In Bradford K and Nonogaki H (eds), *Annual Plant Reviews Volume 27: Seed Development, Dormancy and Germination*. Oxford, UK: Wiley-Blackwell, pp. 25–49.

Dekkers BJW, Pearce S, van Bolderen-Veldkamp RP, Marshall A, Widera P, Gilbert J, Drost H-G, Bassel GW, Müller K, King JR, Wood ATA, Grosse I, Quint M, Krasnogor N, Leubner-Metzger G, Holdsworth MJ and Bentsink L (2013) Transcriptional dynamics of two seed compartments with opposing roles in *Arabidopsis* seed germination. *Plant Physiology* **163**, 205–215.

Denyer T, Ma X, Klesen S, Scacchi E, Nieselt K and Timmermans MCP (2019) Spatiotemporal developmental trajectories in the *Arabidopsis* root revealed using high-throughput single-cell RNA sequencing. *Developmental Cell* **48**, 840–852.e5.

Domcke S and Shendure J (2023) A reference cell tree will serve science better than a reference cell atlas. *Cell* **186**, 1103–1114.

Enomoto H, Sensu T, Yumoto E, Yokota T and Yamane H (2018) Derivatization for detection of abscisic acid and 12-oxo-phytodienoic acid using matrix-assisted laser desorption/ionization imaging mass spectrometry. *Rapid Communications in Mass Spectrometry* **32**, 1565–1572.

Farmer A, Thibivilliers S, Ryu KH, Schiefelbein J and Libault M (2021) Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in Arabidopsis roots at the single-cell level. *Molecular Plant* **14**, 372–383.

Finch-Savage WE and Leubner-Metzger G (2006) Seed dormancy and the control of germination. *New Phytologist* **171**, 501–523.

Francoz E, Ranocha P, Pernot C, Le Ru A, Pacquit V, Dunand C and Burlat V (2016) Complementarity of medium-throughput in situ RNA hybridization and tissue-specific transcriptomics: case study of *Arabidopsis* seed development kinetics. *Scientific Reports* **6**, 24644.

Galland M, Huguet R, Arc E, Cueff G, Job D and Rajjou L (2014) Dynamic proteomics emphasizes the importance of selective mRNA translation and protein turnover during *Arabidopsis* seed germination. *Molecular & Cellular Proteomics* **13**, 252–268.

Grones C, Eekhout T, Shi D, Neumann M, Berg LS, Ke Y, Shahan R, Cox KL, Gomez-Cano F, Nelissen H, Lohmann JU, Giacomello S, Martin OC, Cole B, Wang J-W, Kaufmann K, Raissig MT, Palfalvi G, Greb T, Libault M and De Rybel B (2024) Best practices for the execution, analysis, and data storage of plant single-cell/nucleus transcriptomics. *The Plant Cell* **36**, 812–828.

Guillotin B, Rahni R, Passalacqua M, Mohammed MA, Xu X, Raju SK, Ramírez CO, Jackson D, Groen SC, Gillis J and Birnbaum KD (2023) A pan-grass transcriptome reveals patterns of cellular divergence in crops. *Nature* **617**, 785–791.

Gutzat R, Rembart K, Nussbaumer T, Hofmann F, Pisupati R, Bradamante G, Daubel N, Gaidora A, Lettner N, Donà M, Nordborg M, Nodine M and Mittelsten Scheid O (2020) *Arabidopsis* shoot stem cells display dynamic transcription and DNA methylation patterns. *The EMBO Journal* **39**, e103667.

Haecker A, Groß-Hardt R, Geiges B, Sarkar A, Breuninger H, Herrmann M and Laux T (2004) Expression dynamics of *WOX* genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. *Development* **131**, 657–668.

Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A, Molla G, Madad S, Fernandez-Granda C and Satija R (2024) Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology* **42**, 293–304.

Haughn G and Chaudhury A (2005) Genetic analysis of seed coat development in *Arabidopsis*. *Trends in Plant Science* **10**, 472–477.

Hofmann F, Schon MA and Nodine MD (2019) The embryonic transcriptome of *Arabidopsis thaliana*. *Plant Reproduction* **32**, 77–91.

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

Jolliffe I (2011) Principal component analysis. In Lovric M (ed.), *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer, pp. 1094–1096.

Jovic D, Liang X, Zeng H, Lin L, Xu F and Luo Y (2022) Single-cell RNA sequencing technologies and applications: a brief overview. *Clinical and Translational Medicine* **12**, e694.

Kao P, Schon MA, Mosiolek M, Enugutti B and Nodine MD (2021) Gene expression variation in *Arabidopsis* embryos at single-nucleus resolution. *Development* **148**, dev199589.

Khan D, Millar JL, Girard IJ and Belmonte MF (2014) Transcriptional circuitry underlying seed coat development in *Arabidopsis*. *Plant Science* **219–220**, 51–60.

Kim J-Y, Symeonidi E, Pang TY, Denyer T, Weidauer D, Bezrutczyk M, Miras M, Zöllner N, Hartwig T, Wudick MM, Lercher M, Chen L-Q, Timmermans MCP and Frommer WB (2021) Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *The Plant Cell* **33**, 511–530.

Kompauer M, Heiles S and Spengler B (2017) Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4-μm lateral resolution. *Nature Methods* **14**, 90–96.

Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Horvath S, Drews GN, Fischer RL, Okamuro JK, Harada JJ and Goldberg RB (2010) Global analysis of gene activity during *Arabidopsis* seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 8063–8070.

Lee TA, Nobori T, Illouz-Eliaz N, Xu J, Jow B, Nery JR and Ecker JR (2023) A single-nucleus atlas of seed-to-seed development in Arabidopsis. *bioRxiv*. doi:10.1101/2023.03.23.533992.

Lepiniec L, Devic M, Roscoe TJ, Bouyer D, Zhou D-X, Boulard C, Baud S and Dubreucq B (2018) Molecular and epigenetic regulations and functions of the LAFL transcriptional regulators that control seed development. *Plant Reproduction* **31**, 291–307.

Li B, Zhang W, Guo C, Xu H, Li L, Fang M, Hu Y, Zhang X, Yao X, Tang M, Liu K, Zhao X, Lin J, Cheng L, Chen F, Xue T and Qu K (2022) Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. *Nature Methods* **19**, 662–670.

Li C, Wood JC, Vu AH, Hamilton JP, Rodriguez Lopez CE, Payne RME, Serna Guerrero DA, Gase K, Yamamoto K, Vaillancourt B, Caputi L, O'Connor SE and Buell CR (2023) Single-cell multi-omics in the medicinal plant *Catharanthus roseus*. *Nature Chemical Biology* **19**, 1031–1041.

Li X, Wan Y, Wang D, Li X, Wu J, Chen K, Han X and Chen Y (2024) Temporal-spatial transcriptomics reveals key gene regulation for grain yield and quality in wheat. *bioRxiv*. doi:10.1101/2024.06.02.596756.

Liew LC, Wang Y, Peirats-Llobet M, Berkowitz O, Whelan J and Lewsey MG (2020) Laser-capture microdissection RNA-sequencing for spatial and temporal tissue-specific gene expression analysis in plants. *Journal of Visualized Experiments* **162**, e61517.

Liew LC, You Y, Auroux L, Oliva M, Peirats-Llobet M, Ng S, Tamiru-Oli M, Berkowitz O, Hong UVT, Haslem A, Stuart T, Ritchie ME, Bassel GW, Lister R, Whelan J, Gouil Q and Lewsey MG (2024) Establishment of single-cell transcriptional states during seed germination. *Nature Plants* **10**, 1418–1434.

Lin C, Ding J and Bar-Joseph Z (2020) Inferring TF activation order in time series scRNA-seq studies. *Public Library of Science Computational Biology* **16**, e1007644.

Liu J, Majeed A, Kumar N, Pajerowska-Mukhtar KM and Mukhtar MS (2024) Single nuclei multiomics reveals the drought-driven gene regulatory atlas in *Arabidopsis*. *bioRxiv*. doi:10.1101/2024.01.11.575118.

Lopez-Anido CB, Vatén A, Smoot NK, Sharma N, Guo V, Gong Y, Anleu Gil MX, Weimer AK and Bergmann DC (2021) Single-cell resolution of lineage trajectories in the *Arabidopsis* stomatal lineage and developing leaf. *Developmental Cell* **56**, 1043–1055.e4.

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A and McCarroll SA (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214.

Marand AP, Chen Z, Gallavotti A and Schmitz RJ (2021) A *cis*-regulatory atlas in maize at single-cell resolution. *Cell* **184**, 3041–3055.e21.

McInnes L, Healy J and Melville J (2020) UMAP: uniform manifold approximation and projection for dimension reduction. *arXiv*. doi:10.48550/arXiv.1802.03426.

Montes C, Zhang J, Nolan TM and Walley JW (2024) Single-cell proteomics differentiates *Arabidopsis* root cell types. *bioRxiv*, doi:10.1101/2024.04.09.588771.

Mosher RA and Melnyk CW (2010) siRNAs and DNA methylation: seedy epigenetics. *Trends in Plant Science* **15**, 204–210.

Nakabayashi K, Okamoto M, Koshiba T, Kamiya Y and 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 and Marion-Poll A (2003) ABA action and interactions in seeds. *Trends in Plant Science* **8**, 213–217.

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

Née G, Kramer K, Nakabayashi K, Yuan B, Xiang Y, Miatton E, Finkemeier I and Soppe WJJ (2017) DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling pathway to control seed dormancy. *Nature Communications* **8**, 72.

Nguyen H, Brown RC and Lemmon B (2000) The specialized chalazal endosperm in *Arabidopsis thaliana* and *Lepidium virginicum* (Brassicaceae). *Protoplasma* **212**, 99–110.

Nobori T, Monell A, Lee TA, Zhou J, Nery J and Ecker JR (2023) Time-resolved single-cell and spatial gene regulatory atlas of plants under pathogen attack. *bioRxiv*. doi:10.1101/2023.04.10.536170.

Oliveira MF, Romero JP, Chung M, Williams S, Gottscho AD, Gupta A, Pilipauskas SE, Mohabbat S, Raman N, Sukovich D, Patterson D, Visium HD Development Team and Taylor SEB (2024) Characterization of immune cell populations in the tumor microenvironment of colorectal cancer using high definition spatial profiling. *bioRxiv*. doi: 10.1101/2024.06.04.597233

Palovaara J, Saiga S, Wendrich JR, van't Wout Hofland N, van Schayck JP, Hater F, Mutte S, Sjollema J, Boekschoten M, Hooiveld GJ and Weijers D (2017) Transcriptome dynamics revealed by a gene expression atlas of the early *Arabidopsis* embryo. *Nature Plants* **3**, 894–904.

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

Picard CL, Povilus RA, Williams BP and Gehring M (2021) Transcriptional and imprinting complexity in *Arabidopsis* seeds at single-nucleus resolution. *Nature Plants* **7**, 730–738.

Plant Cell Atlas Consortium, Jha SG, Borowsky AT, Cole BJ, Fahlgren N, Farmer A, Huang SC, Karia P, Libault M, Provart NJ, Rice SL,

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press

Saura-Sanchez M, Agarwal P, Ahkami AH, Anderton CR, Briggs SP, Brophy JAN, Denolf P, Di Costanzo LF, Exposito-Alonso M, Giacomello S, Gomez-Cano F, Kaufmann K, Ko DK, Kumar S, Malkovskiy AV, Nakayama N, Obata T, Otegui MS, Palfalvi G, Quezada-Rodríguez EH, Singh R, Uhrig RG, Waese J, Van Wijk K, Wright RC, Ehrhardt DW, Birnbaum KD and Rhee SY (2021) Vision, challenges and opportunities for a plant cell atlas. eLife 10, e66877.

Rajjou L, Gallardo K, Debeaujon I, Vandekerckhove J, Job C and Job D (2004) The effect of α-amanitin on the *Arabidopsis* seed proteome highlights the distinct roles of stored and neosynthesized mRNAs during germination. *Plant Physiology* **134**, 1598–1613.

Sano N, Rajjou L and North HM (2020) Lost in translation: physiological roles of stored mRNAs in seed germination. *Plants* **9**, 347.

Shaw R, Tian X and Xu J (2021) Single-cell transcriptome analysis in plants: advances and challenges. *Molecular Plant* **14**, 115–126.

Shu K, Liu X-D, Xie Q and He Z-H (2016) Two faces of one seed: hormonal regulation of dormancy and germination. *Molecular Plant* **9**, 34–45.

Shulse CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, Gouran M, Turco GM, Zhu Y, O'Malley RC, Brady SM and Dickel DE (2019) High-throughput single-cell transcriptome profiling of plant cell types. *Cell Reports* **27**, 2241–2247.e4.

Sliwinska E, Bassel GW and Bewley JD (2009) Germination of *Arabidopsis thaliana* seeds is not completed as a result of elongation of the radicle but of the adjacent transition zone and lower hypocotyl. *Journal of Experimental Botany* **60**, 3587–3594.

Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, Giacomello S, Asp M, Westholm JO, Huss M, Mollbrink A, Linnarsson S, Codeluppi S, Borg Å, Pontén F, Costea PI, Sahlén P, Mulder J, Bergmann O, Lundeberg J and Frisén J (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82.

Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, Arlotta P, Macosko EZ and Chen F (2021) Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nature Biotechnology* **39**, 313–319.

Su C, Xu Z, Shan X, Cai B, Zhao H and Zhang J (2023) Cell-type-specific co-expression inference from single cell RNA-sequencing data. *Nature Communications* **14**, 4846.

Swift J, Luginbuehl LH, Schreier TB, Donald RM, Lee TA, Nery JR, Ecker JR and Hibberd JM (2023) Single nuclei sequencing reveals C4 photosynthesis is based on rewiring of ancestral cell identity networks. *bioRxiv*. doi:10.1101/2023.10.26.562893.

Topham AT, Taylor RE, Yan D, Nambara E, Johnston IG and Bassel GW (2017) Temperature variability is integrated by a spatially embedded decision-making center to break dormancy in *Arabidopsis* seeds. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 6629–6634.

Tyler L, Thomas SG, Hu J, Dill A, Alonso JM, Ecker JR and Sun T (2004) DELLA proteins and gibberellin-regulated seed germination and floral development in *Arabidopsis*. *Plant Physiology* **135**, 1008–1019.

Vandereyken K, Sifrim A, Thienpont B and Voet T (2023) Methods and applications for single-cell and spatial multi-omics. *Nature Reviews Genetics* **24**, 494–515.

van Zanten M, Koini MA, Geyer R, Liu Y, Brambilla V, Bartels D, Koornneef M, Fransz P and 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 of the United States of America* **108**, 20219–20224.

Verma S, Attuluri VPS and Robert HS (2022) Transcriptional control of *Arabidopsis* seed development. *Planta* **255**, 90.

Wang K, Zhao C, Xiang S, Duan K, Chen X, Guo X and Sahu SK (2023) An optimized FACS-free single-nucleus RNA sequencing (snRNA-seq) method for plant science research. *Plant Science* **326**, 111535.

Weitbrecht K, Müller K and Leubner-Metzger G (2011) First off the mark: early seed germination. *Journal of Experimental Botany* **62**, 3289–3309.

Welham Z, Déjean S and Lê Cao K-A (2023) Multivariate analysis with the R package mixOmics. In Burger T (ed.), *Statistical Analysis of Proteomic Data. Methods in Molecular Biology*, vol. **2426**. New York, NY: Humana, pp. 333–359.

Xia K, Sun H-X, Li J, Li J, Zhao Y, Chen L, Qin C, Chen R, Chen Z, Liu G, Yin R, Mu B, Wang X, Xu M, Li X, Yuan P, Qiao Y, Hao S, Wang J, Xie Q, Xu J, Liu S, Li Y, Chen A, Liu L, Yin Y, Yang H, Wang J, Gu Y and Xu X (2022) The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in *Arabidopsis* leaves. *Developmental Cell* **57**, 1299–1310.e4.

Yamamoto K, Takahashi K, Mizuno H, Anegawa A, Ishizaki K, Fukaki H, Ohnishi M, Yamazaki M, Masujima T and Mimura T (2016) Cell-specific localization of alkaloids in *Catharanthus roseus* stem tissue measured with imaging MS and single-cell MS. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 3891–3896.

Yan H, Lee J, Song Q, Li Q, Schiefelbein J, Zhao B and Li S (2022) Identification of new marker genes from plant single-cell RNA-seq data using interpretable machine learning methods. *New Phytologist* **234**, 1507–1520.

Yoo S-D, Cho Y-H and Sheen J (2007) *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols* **2**, 1565–1572.

Yuan Y, Huo Q, Zhang Z, Wang Q, Wang J, Chang S, Cai P, Song KM, Galbraith DW, Zhang W, Huang L, Song R and Ma Z (2024) Decoding the gene regulatory network of endosperm differentiation in maize. *Nature Communications* **15**, 34.

Zhang S, Ghatak A, Bazargani MM, Bajaj P, Varshney RK, Chaturvedi P, Jiang D and Weckwerth W (2021) Spatial distribution of proteins and metabolites in developing wheat grain and their differential regulatory response during the grain filling process. *The Plant Journal* **107**, 669–687.

Zhao Z, Cao K, Qin A, Liu Z, Guan L, Sun S, Liu H, Zhou Y, Yang J, Liu Y, Hu M, Ninkuu V, Sun X and Wang L (2023) Comparative spatial transcriptomics of peach and nectarine fruits elucidates the mechanism underlying fruit trichome development. *bioRxiv*. doi:10.1101/2023.10.03.560746

Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT, Shuga J, Montesclaros L, Underwood JG, Masquelier DA, Nishimura SY, Schnall-Levin M, Wyatt PW, Hindson CM, Bharadwaj R, Wong A, Ness KD, Beppu LW, Deeg HJ, McFarland C, Loeb KR, Valente WJ, Ericson NG, Stevens EA, Radich JP, Mikkelsen TS, Hindson BJ and Bielas JH (2017) Massively parallel digital transcriptional profiling of single cells. *Nature Communications* **8**, 14049.

Zumajo-Cardona C, Gabrieli F, Anire J, Albertini E, Ezquer I and Colombo L (2023) Evolutionary studies of the bHLH transcription factors belonging to MBW complex: their role in seed development. *Annals of Botany* **132**, 383–400.

https://doi.org/10.1017/S0960258524000199 Published online by Cambridge University Press