Title: A Single-Nucleus Atlas of Seed-to-Seed Development in Arabidopsis

Authors: Travis A. Lee[1,2,3]†, Tatsuya Nobori[1,2]†, Natanella Illouz-Eliaz[1,2]†, Jiaying Xu[1,2], Bruce Jow[1,2], Joseph R. Nery[1,2], and Joseph R. Ecker[1,2,3]*

Affiliations:

[1]Plant Biology Laboratory, The Salk Institute for Biological Studies; La Jolla, CA 92037.

[2]Genomic Analysis Laboratory, The Salk Institute for Biological Studies; La Jolla, CA 92037.

[3]Howard Hughes Medical Institute, The Salk Institute for Biological Studies; La Jolla, CA 92037.

* Corresponding author. Email: ecker@salk.edu

Abstract: Extensive studies of the reference plant Arabidopsis have enabled deep understandings of tissues throughout development, yet a census of cell types and states throughout development are lacking. Here, we present a single-nucleus transcriptome atlas of seed-to-seed development employing over 800,000 nuclei, encompassing a diverse set of tissues across ten developmental stages, with spatial transcriptomic validation of the dynamic seed and silique. Cross-organ analyses revealed transcriptional conservation of cell types throughout development but also heterogeneity within individual cell types influenced by organ-of-origin and developmental timing, including groups of transcription factors, suggesting gatekeeping by transcription factor activation. This atlas provides a resource for the study of cell type specification throughout the continuum of development, and a reference for stimulus-response and genetic perturbations at the single-cell resolution.

One-Sentence Summary: A single nucleus atlas of seed-to-seed development in Arabidopsis charts a course through the lifecycle of an organism.

Introduction

Multicellular organisms have evolved various organs to perform specific functions required for the organism to survive and flourish. Plants, unlike most animals, undergo dynamic post-embryonic organogenesis to form new organs over the course of life. Some plant organs are highly specialized and developmental stage-specific, and many organs consist of various cell types with distinct sub-functions. However, some cell types share functionality across diverse organs and developmental stages based on anatomical and physiological features. For instance, epidermal cell types, generally defined as part of the outermost layer of cells of an organ, protect the organ from and interact with external environmental cues, whereas internally localized vascular cells are required for the transport of water and nutrients (*1*, *2*). Yet, we still have a limited understanding of many cell types at the molecular level in the context of organ specificity. High throughput single-cell RNA-sequencing (RNA-seq) has been demonstrated to provide detailed maps of cell types in plants (*3*). Still, its application is currently limited to selected organs, tissues, and cell types (*4*–*8*), with a predominant focus on the Arabidopsis root tip (*9*–*11*), posing a bottleneck toward a comprehensive understanding of cell types and states in this model organism. This motivated us to generate an extensive single-nucleus transcriptome atlas of Arabidopsis development that encompasses all major organs present over the plant's entire life cycle. Global characterization of plant cell types across organs at the molecular level will be paramount for understanding organ development and function. All data and annotations can be accessed through a web portal at http://arabidopsisdevatlas.salk.edu/.

Results

A comprehensive single-nucleus atlas of the Arabidopsis lifecycle

To generate a comprehensive atlas of Arabidopsis development, we collected six distinct organs that encompass a diverse range of tissues present at several developmental stages and transitions throughout the entire life cycle corresponding to developmental roadmaps (*12*–*14*), including imbibed and germinating seeds, three timepoints of seedling development, developing and fully emerged rosettes, the stem (apical, branched area and stem base), a range of flower tissue (unopened flower buds to fully mature flowers), and siliques (immature to fully elongated green siliques) (Fig. 1A). To circumvent the large range of plant cell sizes and reduce heterogeneous sampling effects between tissues, a universal nuclei extraction protocol, amenable for droplet-based single-nuclei sequencing was developed, and further optimized for each tissue (see Supplementary methods). A total of 801,276 nuclei from the ten samples passed accepted droplet-based single nuclei filtering metrics (*15*) and were independently clustered and merged into a global dataset (Fig. 1B). A parallel analysis using higher stringency filtering cutoffs resulting in 432,919 nuclei revealed a similar global structure (Fig. 1C), demonstrating the overall high quality of the larger dataset, but was associated with overall increased cluster number and resolution. While it is possible that the loss of clusters and cell types may occur with the higher stringency dataset, we proceeded with the higher stringency

dataset for downstream analyses but also maintained the larger dataset.

The global clustering of all ten datasets revealed several distinctly separated clusters, with the presence of a large central cluster (Fig. 1C). To determine

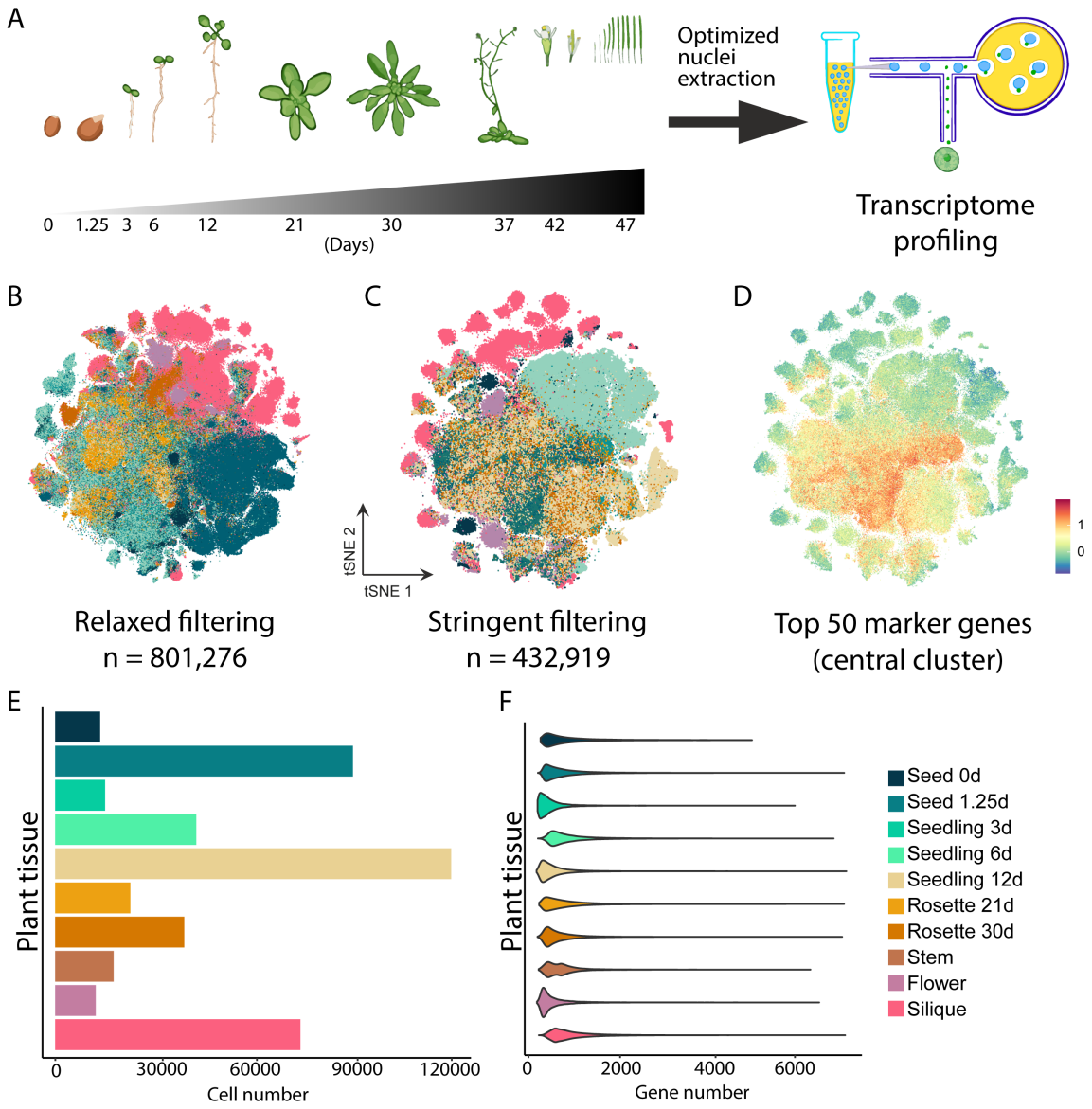

Figure 1. A seed-to-seed atlas of Arabidopsis thaliana.
(**A**) Illustration of the collected tissues over the 10 developmental time points spanning the plant life cycle including: imbibed and germinating seeds (0d and 1.25d), three stages of seedling development (3d, 6d, and 12d), developing and fully expanded rosettes (21d and 30d), the stem (40d) including basal, apical, and branched regions, flower tissue [stages 6-15 (*14*)], and siliques [stages 2-10 (*13*)]. (**B**) relaxed and (**C**) stringent tSNE embeddings of the fully integrated dataset. Nuclei are colored by the dataset of origin. (**D**) Average expression of the top 50 marker genes across the large central cluster of the global tSNE. (**E**) Number of nuclei per dataset. (**F**) Violin plot of genes detected per nucleus in each dataset.

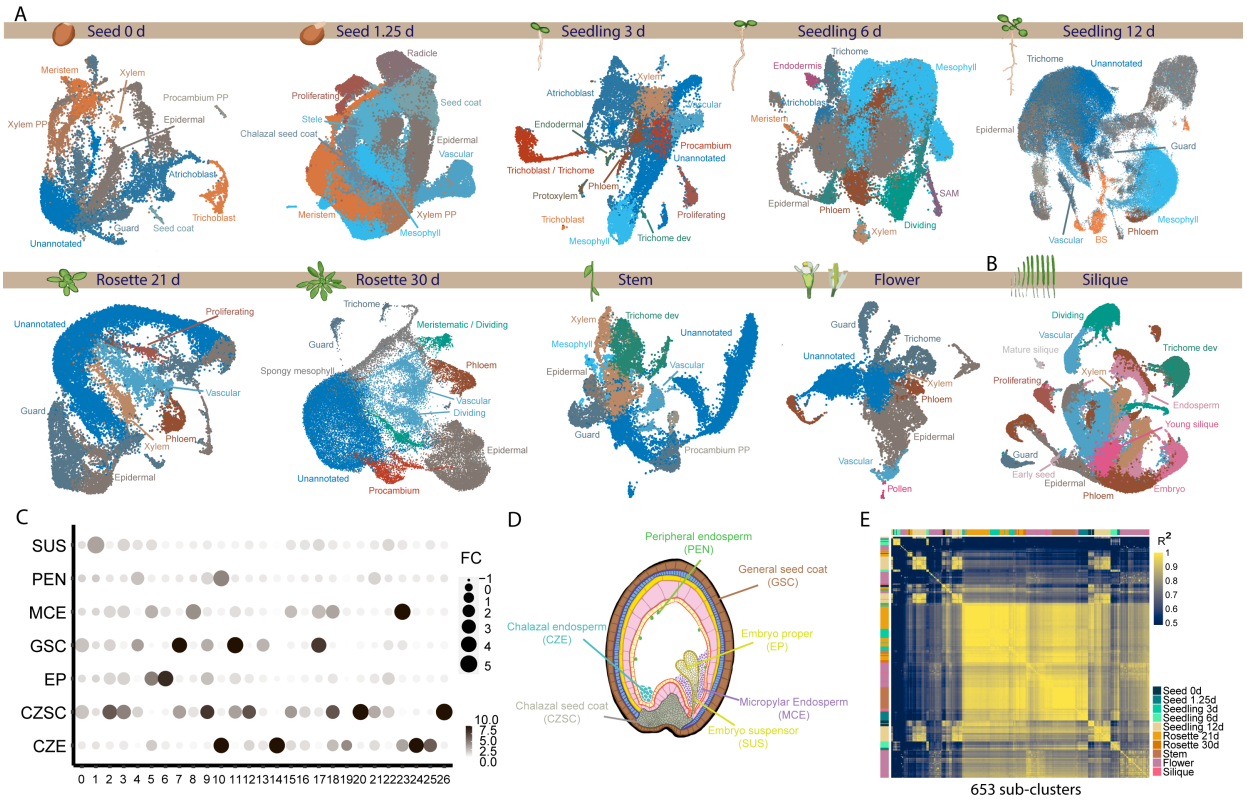

Figure 2. Independent clustering of ten datasets
(**A**) UMAP of each dataset with select high-confidence annotations for each dataset. (**B**) UMAP of the silique dataset. (**C**) Dot plot showing literature marker overlap with major clusters identified in the silique dataset. FC, fold change. (**D**) Illustration of early seed development with cell type annotations. (**E**) Heatmap of sub-cluster markers correlation across the entire dataset. Sub-clusters that contain more than 17,000 genes with zero count were removed from the analysis. The top color bars indicate the organ of origin.

the defining characteristics of the large cluster, we examined the aggregated and individual expression of the top 50 marker genes of these cells (Fig. 1D and figs. S1A-H). Many top markers are involved in photosynthesis, suggesting that this biological process has a dominating influence on the transcriptome of cells with high levels of photosynthetic activity (fig. S1I). To assess the heterogeneity of this central cluster, we removed the abundantly expressed photosynthesis-related genes and re-clustered the resulting data (fig. S2), which slightly improved cluster resolution. However, a large central cluster remained, suggesting a conserved functionality across several populations of photosynthetically active cells.

Overall, we chose to more deeply sample seedling tissue due to the potentially higher quantities of discrete cell types, contributing to the presence of both above- and below-ground tissues (Fig. 1E). Across all tissues and time points, we observed comparable transcriptional complexity between samples and tissues (Fig. 1F). Independent clustering of each dataset revealed a total of 183 clusters, many of which correspond to individual cell types (Figs. 2A-B, fig. S3, and Table S1). As high-resolution cell type maps are not currently available for many of these tissues and developmental time

points, annotation was guided by curated marker genes identified in previous studies (see Methods and Table S2). The organ with the greatest cluster complexity was the silique, where we identified 26 major clusters (Fig. 2B and Table S1), which may arise from the diversity of tissues within the fruit organ (fruit flesh and developing embryonic tissue), and from real-time developmental gradients present both across individual siliques and within each silique (*16*). Further, within individual siliques, diverse stages of embryonic development from fertilization to zygotic to embryonic development can be present along the longitudinal axis of the fruit based on the timing of fertilization of individual embryo sacs. Using previously validated cell type-specific marker genes of the silique (*17*), we were able to annotate many of the clusters to individual cell types of this organ (Fig. 2C and Table S1) along with several cell types that comprise the developing embryo and seeds (Figs. 2B to D).

To explore the transcriptional heterogeneity within and across individual cell types, we systematically performed subclustering of the 183 major clusters identified from all organs, resulting in a total of 653 subclusters representing both unique cell identities or states across plant tissues and development (fig. S4A). Comparing the aggregated transcriptomes (pseudo-bulk) of individual subclusters revealed groups of subclusters derived from individual organs that share conserved transcriptional identities, as well as subcluster groups associated with unique gene expression patterns (Fig. 2E), which was further supported by analyzing overlapping sets of *de novo,* identified cluster marker genes across all 653 subclusters, where we observed particularly high cluster diversity among subclusters of seeds and siliques (fig. S4B). Taken together, our seed-to-seed transcriptome atlas captured known cell types and a diverse set of previously uncharacterized cell populations at the major and subcluster levels.

Reconstructing development along real time and pseudotime

To explore real-time developmental dynamics of individual cell types, we utilized the seedling datasets that comprised identical tissue and organ compositions at three developmental time points (3-, 6-, and 12-day-old-seedlings). Integration of these datasets revealed that all clusters were represented by each timepoint, demonstrating a conservation of cell types across these three stages of seedling development (Fig. 3A and fig. S5). However, despite successful integration, the populations of several clusters were skewed by developmental age, including the annotated root hair clusters. It is well characterized that the timing of root hair (I) specification, (II) emergence, (III) elongation, and (IV) maturation is developmentally gated (*18*), with root hairs first observable around two days of growth, suggesting that the skewed developmental representation reflects varying degrees of root hair differentiation and development; thus our time-resolved seedling atlas provided an opportunity to reconstruct this developmental process by trajectory inference and pseudotime analysis supported by real-time ground truth. Re-clustering of the three annotated root hair clusters revealed heterogeneity within this population of cells (Fig. 3B). Fitting a trajectory and pseudotime to these cells revealed a continuous gradient of root hair maturation, capturing the expression of known root hair developmental stage-

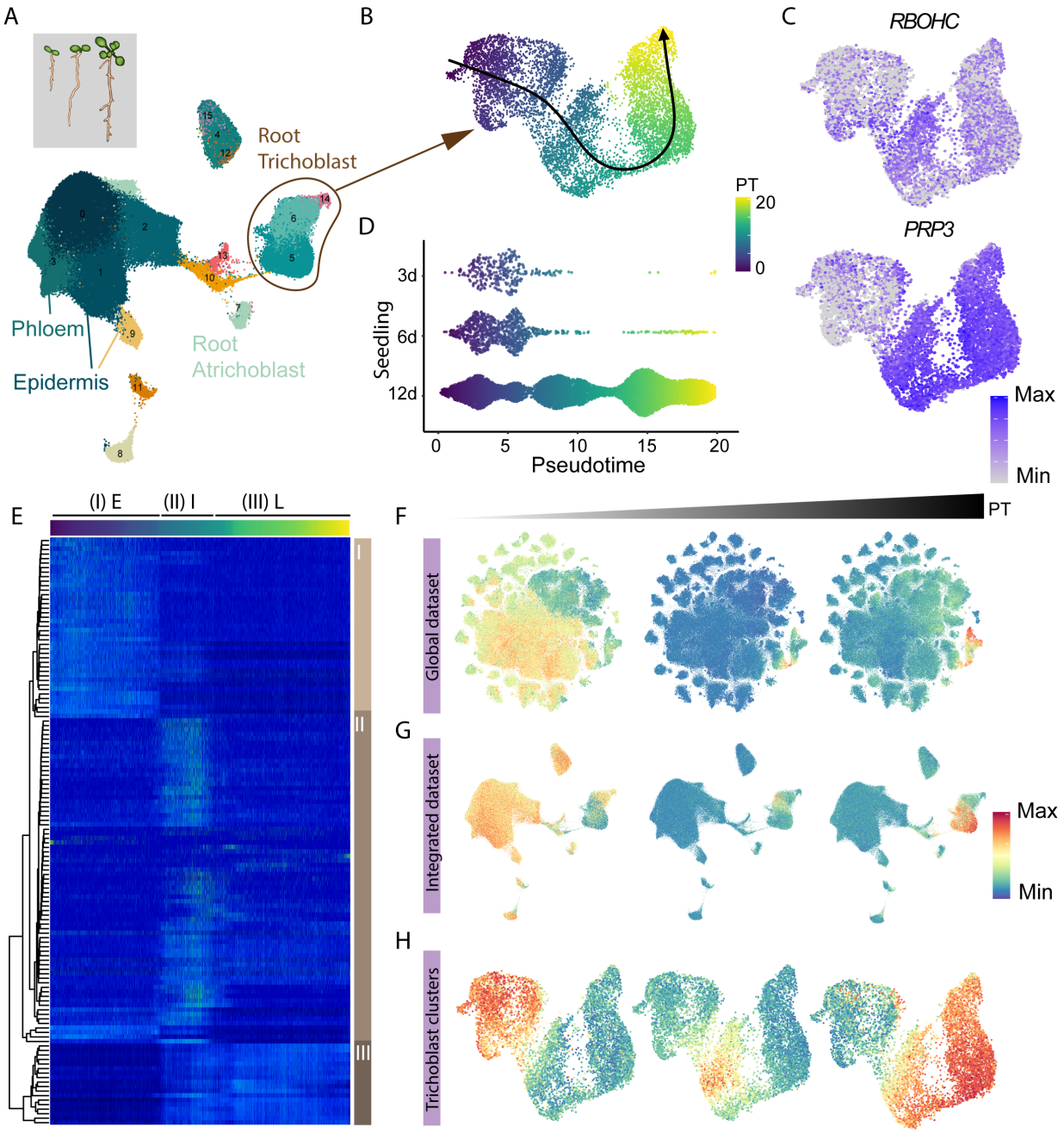

Figure 3. Transcriptional programs of root hair development across the entirety of Arabidopsis development

(**A**) Integrated clustering of the 3d, 6d, and 12d seedling datasets. Root hair (trichoblast) clusters are circled. (**B**) Re-clustering and pseudotime trajectory of annotated root hair clusters. (**C**) Expression of known genes expressed in root hair development (*ROBHC* [AT5G51060] and *PRP3* [AT3G62680]) (**D**) Distribution of root hair cells from 3d, 6d, and 12d seedlings across pseudotime. (**E**) Heatmap of 170 genes differentially expressed along pseudotime of root hair cells. Three major classes of genes corresponding to (I) E=early, (II) I=intermediate, and (III) L=late pseudotime, are depicted. (**F** to **H**) Expression of the three gene modules identified from pseudotime in the (F) global, (G) integrated seedling, and (H) root hair re-clustering.

specific genes (Fig. 3C). Affirming our prior findings, we observed that nuclei of the youngest seedling dataset (3d) were enriched early in the trajectory. In contrast, nuclei of the older seedlings progressively populated intervals of late pseudotime (6d and 12d) and at greater proportions (Fig. 3D). Identification and clustering of genes differentially expressed over pseudotime revealed three expression patterns, with gene modules that largely correlated to these three stages of root hair development (Fig. 3E). Of note, these *de novo* identified gene groups are not enriched for processes related to root hair development but rather broader GO terms such as anatomical structure process and glycerolipid metabolic process, implicating novel roles of those genes in root hair development (fig. S6).

To utilize the breadth of our datasets, we sought to determine if these spatiotemporally expressed gene modules have a conserved role in other tissues and/or cell types. Mapping the expression of each gene module to the integrated seedling and global datasets revealed examples of both highly specific and broad roles of these genes across development (Figs. 3F-H). For example, gene modules corresponding to root hair differentiation in pseudotime (intermediate; II) were exclusively expressed in this population of cells within the context of all tissues and developmental time points assayed (Figs. 3F-H), suggesting that these genes function solely within this trajectory of root hair development. Conversely, genes corresponding to both early and late pseudotime of root hair development were more broadly expressed in other tissues and cell types (Figs. 3F-H), suggesting that the transcriptional programs associated with epidermal cell fate and root hair maturation are shared among other biological programs throughout development. Overall, our time-resolved atlas, including broad tissues and developmental stages, revealed examples of transcriptional programs that are shared across cell types of various tissues and programs uniquely utilized by a single cell type at discrete stages of development.

Cross-tissue analyses of conserved cell types

To determine if transcriptional programs between like-cell types are shared between identical cell types originating from various organs, we evaluated discrete populations of cells identified at the global clustering level (Figs. 1C and 4A). Expression of known marker genes was restricted to individual clusters as evidenced by cluster-specific expression of the stomatal lineage and phloem marker genes, *FAMA* [*AT3G24140*, (*19*)] and *SUC2* [*AT1G22710*, (*20*, *21*)], respectively. The composition of these clusters was well represented by cells from all tissue types assayed (Figs. 4B-D), demonstrating that the transcriptional identity of these cell types is strongly conserved throughout development. Quantification of other distinct clusters revealed that many clusters were well represented by various tissues-of-origin (figs. S7A and B).

Subclustering of the phloem and guard cell lineage populations revealed additional heterogeneity not observed at the global clustering level, whereby cells originating from similar tissues of origin (three seedling datasets, two rosette datasets) co-segregated (Figs. 4B to C). Conversely, guard cell lineage and phloem cells from the terminally differentiated reproductive and germline

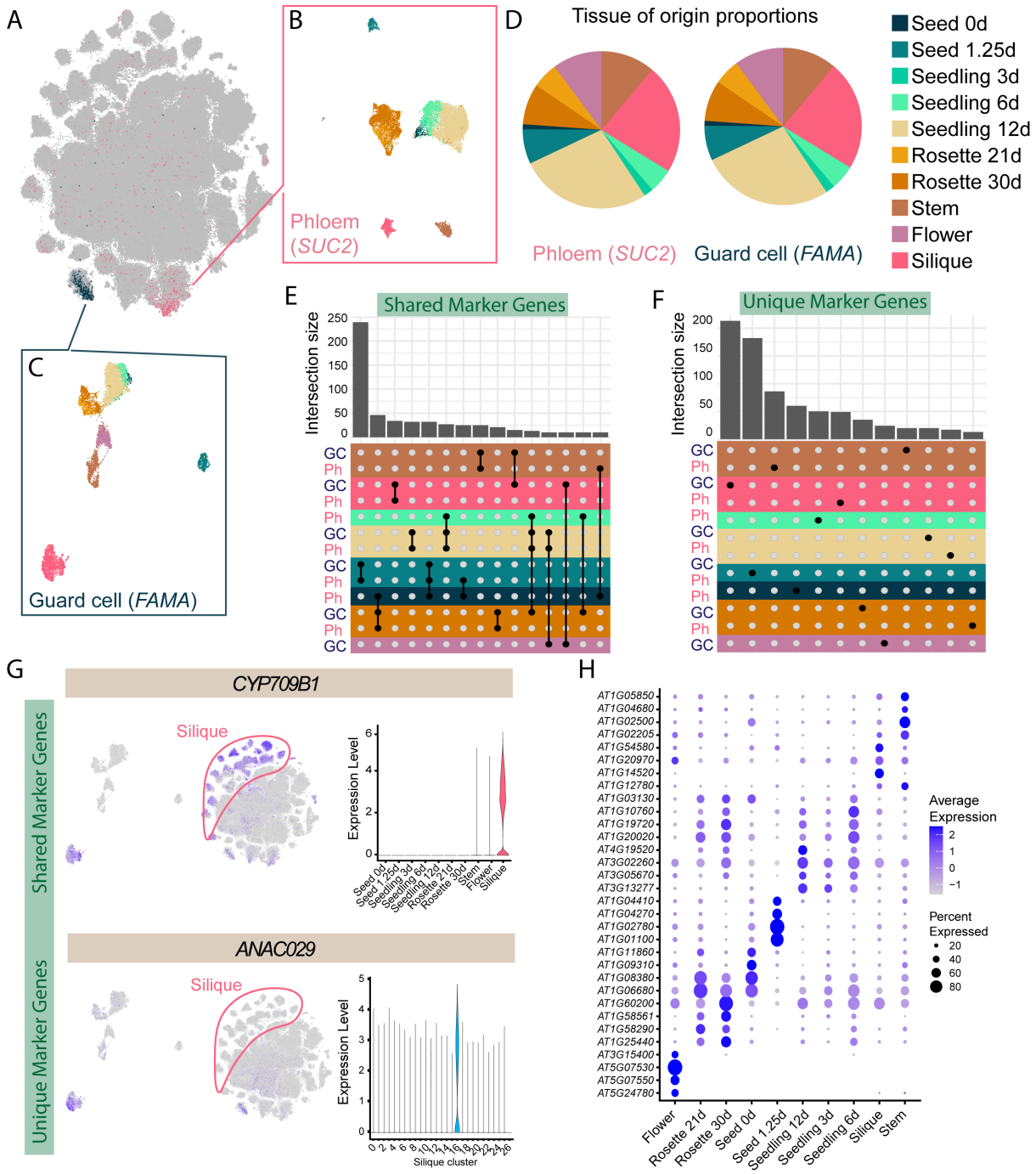

Figure 4. Identification of tissue- and cell type-specific transcriptional signature
(**A**) tSNE of the global dataset with expression levels of the stomatal lineage marker *FAMA* (blue, [*AT3G24140*]) and phloem marker *SUC2* (pink, [*AT1G22710*]). (**B** and **C**) Re-clustering of the (B) phloem and (C) guard cell lineage clusters. Cells are colored by tissue of origin. (**D**) Distribution of cells per dataset in the phloem and guard cell lineage clusters. (**E**) Quantification of subcluster markers shared between phloem and guard cell lineage subclusters. Groups of subclusters with greater than 10 overlapping subcluster markers are depicted. Black dots indicate the overlapping group(s) of subclusters. Color indicates the dataset of origin with the highest proportion of cells for each subcluster.

(**F**) Quantification of markers uniquely identified as a subcluster marker in only one subcluster of phloem and guard lineage cells. Subclusters with at least 10 uniquely identified genes are depicted (**G**) Expression of a representative shared tissue (CYP709B1 [AT2G46960]) and uniquely identified marker (*ANAC029* [*AT1G69490*]) in the stomatal lineage subclustering, global integrated dataset, and expression within each tissue dataset. (H) Expression of four representative tissue markers identified for each tissue.

tissues clustered distinctly from the vegetative tissue, suggesting organ-specific influences on the transcriptional identities of these cell types. Exploring the expression patterns of *de novo* identified subcluster markers of guard cell lineage and phloem cells at the global and cell type levels revealed two classes of genes; (I) organ/tissue-specific - expressed throughout various cell types of an individual tissue (identified as subcluster markers in at least half of the subclusters of a sampled organ or tissue) (Fig. 4E and Table S3), and (II) unique - under the combined regulation of cell type-specific and developmental specific programs (Fig. 4F and Table S3). Extending this analysis to include ten distinct clusters (figs. S7 and S8) further revealed the presence of genes both with tissue level expression patterns and under the combined regulation of cell type and developmental timing (2,941 genes) (Figs. 4G-H and figs. S7C and D).

Of note, we observed the largest quantity of overlapping subcluster markers across cell types of the germinating seed (1.25d) (Table S3). A deeper investigation of conserved markers of germinating seeds revealed that many of these genes with shared expression across cell types encode a diverse range of ribosomal protein subunits (figs. S9 and S10). The unified expression of ribosomal proteins across cell types in the germinating seed dataset coincides with global increases in ribosome biogenesis and translation during seed germination (*22*). While the expression of these ribosomal protein genes is present across several cell types of germinating seeds, this was not observed for all of the other tissues assayed, where subsets of ribosomal proteins were expressed only in cells of specific tissues or in cell type(s) specific expression patterns, supporting the hypothesis that the heterogeneous composition of ribosomes are developmentally regulated (*23*), and may explain the higher levels of homogeneity within this tissue.

Together, these findings suggest the presence of cell populations whereby transcriptional identity is influenced by organ- or tissue-specific transcriptional networks but also unique cell populations with context-specific transcriptional regulation. This may be correlated with cell functionality across tissues, as some cell types and states are conserved regardless of tissue of origin, while others may encounter different environmental stimuli, such as abaxial versus adaxial leaf epidermal guard cells.

Transcription factor specificity across development

We systematically analyzed TF expression across the identified cell populations to understand cell type and organ-specific gene regulatory mechanisms. Many TFs showed organ-specific expression (Figs. 5A and B), but we also observed heterogeneity within organs (figs. S11A-C). For instance, AT3G15510 (*ANAC056*) was expressed specifically and globally in the silique tissue (Fig. 5C); AT3G62340 (*WRKY68*) was also specifically expressed in the silique, within a subset of cells, but also

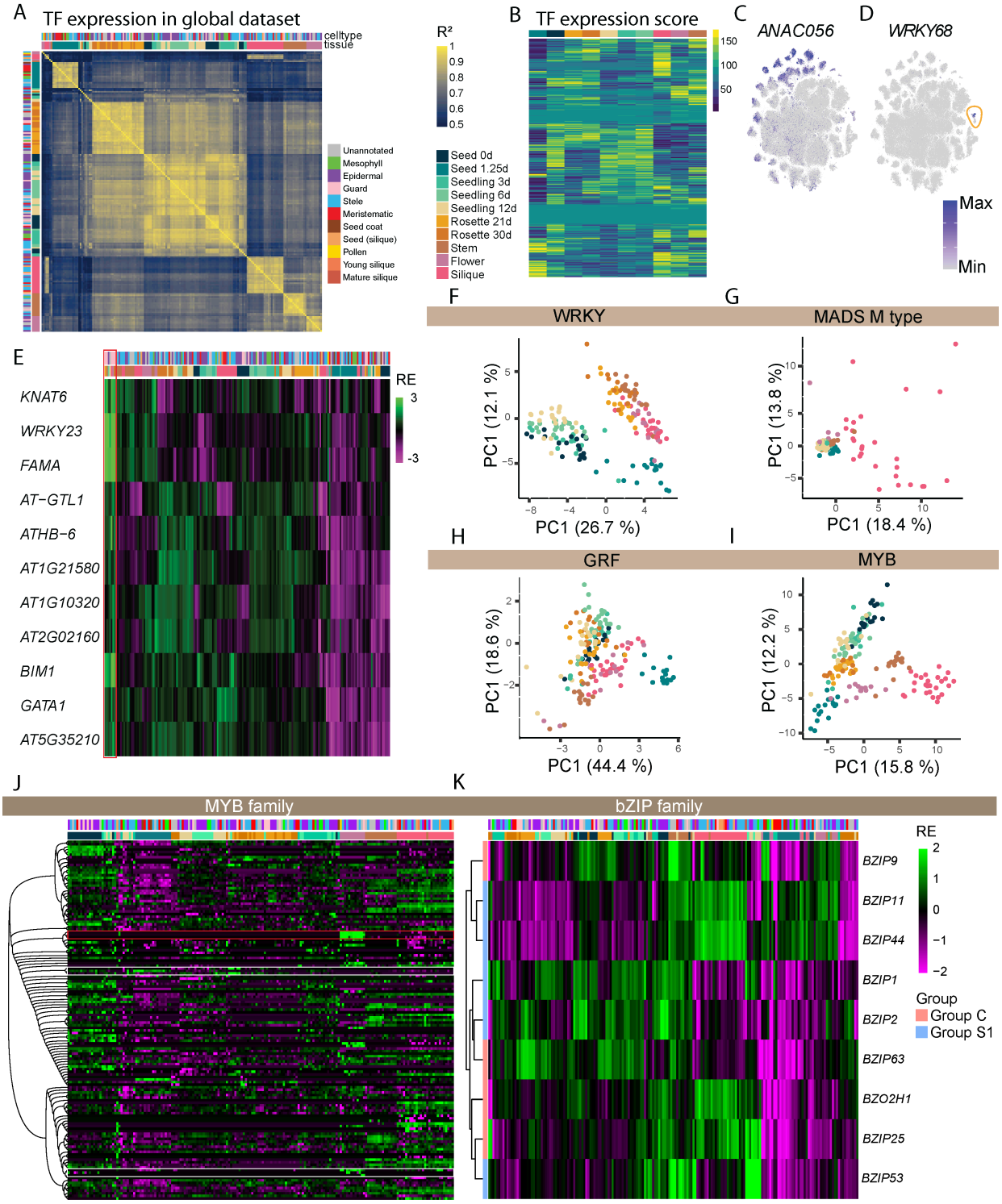
A
TF expression in global dataset
celltype
tissue
Unannotated
Mesophyll
Epidermal
Guard
Stele
Meristematic
Seed coat
Seed (silique)
Pollen
Young silique
Mature silique
R²
1
0.9
0.8
0.7
0.6
0.5
B
TF expression score
150
100
50
Seed 0d
Seed 1.25d
Seedling 3d
Seedling 6d
Seedling 12d
Rosette 21d
Rosette 30d
Stem
Flower
Silique
C
ANAC056
D
WRKY68
Max
Min
E
KNAT6
WRKY23
FAMA
AT−GTL1
ATHB−6
AT1G21580
AT1G10320
AT2G02160
BIM1
GATA1
AT5G35210
RE
3
-3
F
WRKY
PC1 (12.1 %)
PC1 (26.7 %)
G
MADS M type
PC1 (13.8 %)
PC1 (18.4 %)
H
GRF
PC1 (18.6 %)
PC1 (44.4 %)
I
MYB
PC1 (12.2 %)
PC1 (15.8 %)
J
MYB family
K
bZIP family
BZIP9
BZIP11
BZIP44
BZIP1
BZIP2
BZIP63
BZO2H1
BZIP25
BZIP53
RE
2
1
0
−1
−2
Group
Group C
Group S1

Figure 5. Diverse expression patterns of TFs across cell types and organs

(**A**) Heatmap showing pair-wise correlation of pseudo-bulk TF expression between major clusters of individual samples. (**B**) Averaged ranks of TF expression in the clusters derived from each tissue of origin. A higher value indicates more global and dominant expression of a TF in a tissue of origin. (**C** and **D**) Expression of TFs, AT3G15510 (*ANAC056*) and AT3G62340 (*WRKY68*), that are specifically expressed in the silique visualized in a globally clustered tSNE. A cluster with strong gene expression was circled in D. (**E**) Heatmap showing pseudo-bulk expression of marker TFs of global cluster 21 (guard cell cluster; Fig. 4A). Clusters that showed strong expression of these markers were highlighted in red. (**F** to **I**) Principal component analysis of TF family expression in each cluster. Each dot represents a major cluster colored by a tissue of origin. TF family names are indicated in the top bars. (**J**) Heatmap showing expression of MYB family TFs in all major clusters. Genes were ordered by the phylogenetic relationships defined by their amino acid sequences. An example of genes that showed tissue specific expression were highlighted in red. Examples of closely related genes that showed contrasting expression patterns were shown in white. (**K**) Heatmap showing relative expression of bZIP family TFs that belong to Group C and S1. For each heatmap, top bar indicates tissue of origin and/or cell type annotation. (E, J, K) Expression was shown as $\log_2$ relative expression (RE).

within the entire context of development (Fig. 5D). Notably, we identified TFs expressed in specific clusters of each organ. TFs such as *FAMA*, *WRKY23*, and *BIM1* were highly expressed in clusters mostly annotated as guard cells in seven different samples (Fig. 5E; highlighted in red), suggesting that these TFs have conserved cell type-specific roles across various organs and development.

We next asked whether TF clades can explain organ specificity and within-organ heterogeneity of expression patterns. We performed principal component analysis on the expression of 70 TF clades in individual clusters and found both cases where TF family members showed either shared or diverse cell type/organ specificity (Figs. 5F-I and fig. S11D). For instance, expression patterns of WRKY TFs are primarily separated into three organ groups (regardless of cell type): (I) germinating seeds (1.25d), (II) imbibed seeds (0d) and seedlings, and (III) rosette, stem, flower, and silique. Cell type/state heterogeneity was observed within each organ group (Fig. 5F). Some TF families have largely expanded and showed high degrees of heterogeneity (Figs. 5J and K and fig. S11E and F). Interestingly, we found that, in some cases, phylogenetic relationships of genes within TF families can explain their organ/cell type specificity, as clusters associated with cell types and/or organ formed grouped when TF family members were clustered phylogenetically (Fig. 5J; highlighted with red boxes). We also found cases where closely related TFs showed contrasting expression patterns (Fig. 5J; highlighted with white boxes).

Basic leucine zipper (bZIP) TFs recognize various DNA sequences by forming various homo- and hetero-dimers (*24*). A previous study showed that Group C bZIP TFs (bZIP9, bZIP10, bZIP25, bZIP63) do not bind DNA as homodimers but form heterodimers with Group S1 members (bZIP1, bZIP2, bZIP11, bZIP44, bZIP53) to bind unique DNA sequences (*25*). We found that genes in each group showed varying expression patterns (Fig. 5K). Group C TFs were clustered with at least one Group S TF with overlapping expression patterns (Fig. 5K), implying the co-evolution of cell type/organ specificity in these dimerization pairs. Our dataset will

be valuable for analyzing the neo/sub-functionalization of TFs over the course of gene family expansion, which has previously been suggested for some genes (*26*).

Spatial mapping of cell types

Annotating cell types and cell states to clusters identified in snRNA-seq requires *a priori* knowledge of previously validated cell type-specific markers by imaging- and dissection-based analyses. However, such ground truth cell type markers are unavailable for many organs and cell types within the context of the entire Arabidopsis lifecycle. Therefore, to validate our cluster annotations, we utilized two spatial transcriptomics technologies to simultaneously validate cell type-specific expression patterns of several *de novo* identified clusters from two datasets associated with highly dynamic stages of growth; germinating seeds (1.25d seeds) and siliques.

To spatially profile the transcriptome of germinating seeds (1.25d) (Fig. 2A), we utilized a sequencing-based spatial transcriptomics platform with 10 µm spatial resolution to profile cell type and cell layer-specific transcriptomes (*27*). Plotting the transcript detection of spots in spatial coordinates revealed shapes resembling both seed and embryo structures (figs. S12A and B). Dimensionality reduction and *de novo* clustering of the filtered spatial transcriptomics data revealed two major groupings of clusters (Fig. 6A). Spatial mapping revealed that these clusters broadly correspond to the cotyledons (clusters 0 and 5), root tip region (cluster 1), epidermis (cluster 2), seed coat (cluster 3), and the provasculature (cluster 4) (Figs. 6B to D). Mapping the expression of cluster markers corresponding to the cotyledon and epidermal clusters (clusters 0 and 2; fig. S10) onto the matched droplet-based single-nuclei dataset similarly revealed cluster-specific expression patterns to the correspondingly annotated cell types (Fig. 6E), demonstrating the ability to accurately annotate the droplet-based clusters.

As fully elongated siliques simultaneously represent dynamic growth patterns associated with both small (egg sac; embryo) and large tissues (~15 mm silique when fully elongated), we opted for an imaging-based spatial transcriptomics method (MERFISH; see Methods) that can accommodate higher resolution of small cells as well as larger tissue areas to simultaneously visualize the expression of many transcripts on a tissue section at single-molecule resolution. We selected 140 *de novo* identified silique cluster marker genes and spatially mapped them on 10 µm tissue sections of fully expanded green siliques (Fig. 6F; Table S6). Marker genes of clusters predicted to be embryo, and chalazal endosperm (Fig. 2B-C, Table S2) were successfully mapped in expected regions of the tissues (Figs. 6G and H). Among these markers, AT2G44240 has been validated as a chalazal marker by *in situ* hybridization (*17*). We also identified cells in the stomatal lineage based on expression of the guard cell marker gene *FAMA (19)* and validated the *de novo* identified guard cell marker gene AT3G16340 (PDR1) (Figs. 6I and J). These results validate our cell type prediction and, *de novo* identified marker genes of these cell types. Mapping the expression patterns of these highly specific genes revealed that many of these cell type-specific transcripts of the silique are uniquely expressed in this

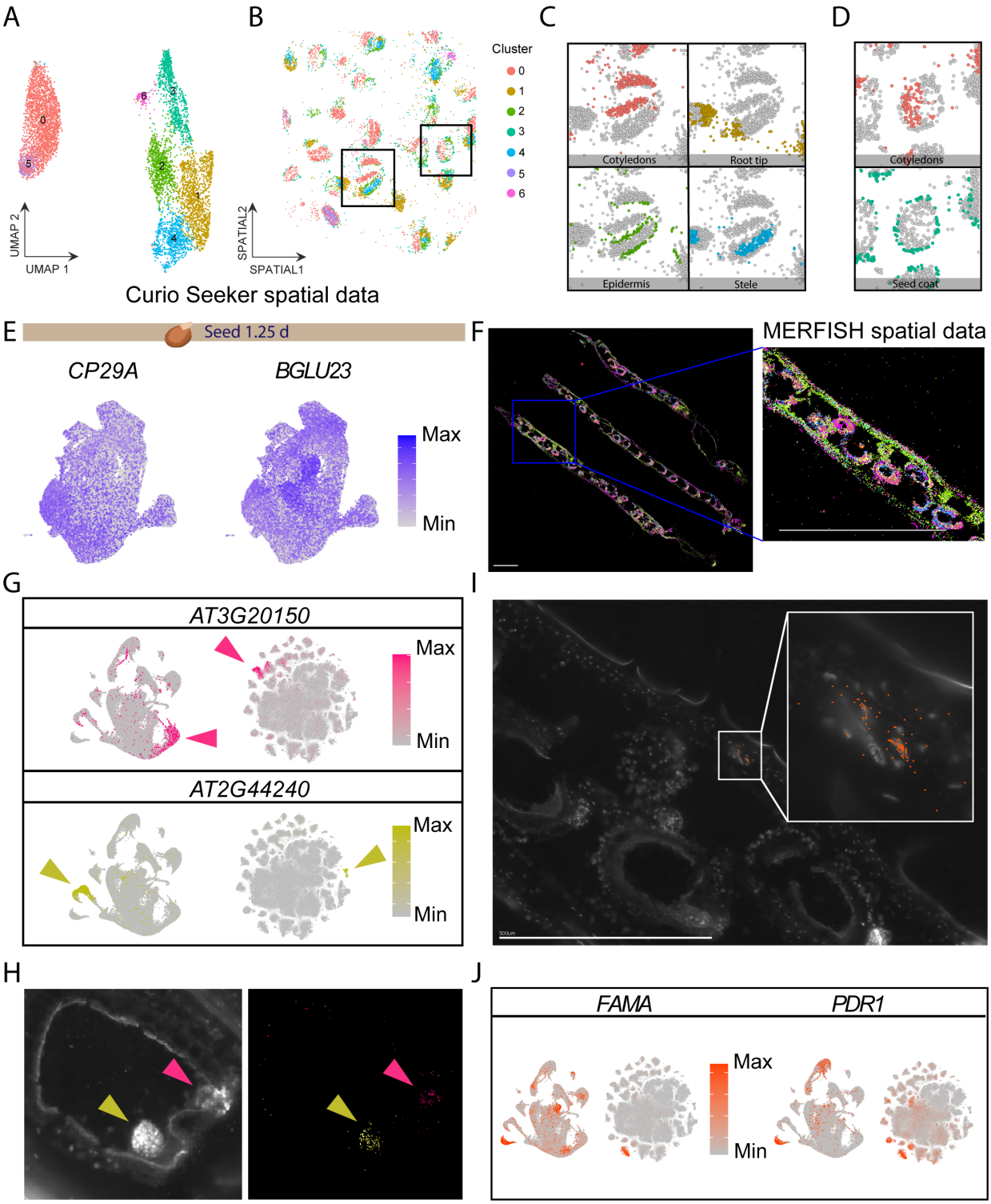
A
B
C
D
UMAP 2
UMAP 1
SPATIAL2
SPATIAL1
Cluster
0
1
2
3
4
5
6
Cotyledons
Root tip
Epidermis
Stele
Cotyledons
Seed coat
Curio Seeker spatial data
E
Seed 1.25 d
CP29A
BGLU23
Max
Min
F
MERFISH spatial data
G
AT3G20150
Max
Min
AT2G44240
Max
Min
I
H
J
FAMA
PDR1
Max
Min

Figure 6. Spatial transcriptomics annotates silique cell types and validates de novo marker genes (**A**) Dimensionality reduction and *de novo* clustering of spatial transcriptomics data of 1.25-day-old germinating seeds. (**B** to **D**) Mapping of clusters to spatial coordinates with a focus on clusters corresponding to the (C) embryonic seedling and (D) seed coat. (**E**) Expression of *de novo* identified spatial cluster markers corresponding to cotyledons (left, *CP29A* [*AT3G53460*]) in the matched single-nucleus droplet based dataset (Figure 2A). (**F**) Visualization of all 140 transcripts targeted with MERFISH in siliques. Scale bar = 1 mm. (**G**) Expression of *de novo* identified marker genes of silique clusters predicted to be the embryo (*AT3G20150*) and chalazal endosperm (*AT2G44240*) in the (left) silique and (right) globally clustered datasets. (**H**) (Left) DAPI staining of an embryo within the silique. (Right) Spatial mapping of predicted marker genes of embryo (yellow) and chalazal (pink). Scale bars = 0.2 mm. (**I**) Spatial mapping of identified guard cell marker genes. The area in the white box was enlarged. Scale bar = 0.5 mm. (**J**) Expression of two *de novo* identified guard cell lineage markers shown in the silique dataset (left) and the global dataset (right).

tissue throughout all stages of Arabidopsis development assayed in this study, highlighting the ability to investigate cell type and developmental-specific expression patterns in our dataset (Figs. 6G and J). In summary, spatial transcriptomics validated *de novo-identified* cluster markers and facilitated cluster annotation in these complex organs.

Discussion

In this study, utilizing single-nucleus sequencing technologies, we sought to characterize a diverse range of tissues and organs along the entire life cycle of an organism. The post-embryonic development and emergence of distinct tissues and organs is a notable aspect of plant development, which is not present in most other systems, and was incorporated into our experimental design. With a broad sampling strategy spanning several well-studied developmental stages, our single-nucleus atlas of Arabidopsis development provides a resource of wide use to the Arabidopsis and plant community for hypothesis generation and reference for future single-cell genomics studies.

We identified tissue of origin transcriptional signatures in like or identical cell types, uncovering additional layers of heterogeneity that holistically contribute to the transcriptional identity of individual cells. Contrastingly, we also identified genes under the combined regulation of cell type and tissue-of-origin specificity, with many of these genes uniquely expressed solely within a single cell type of single organs. Many of these genes uniquely expressed in our dataset are not functionally annotated and thus serve as potential candidates for follow-up studies. Interestingly, cell types of the terminal reproductive tissues containing both germ- and non-germ-line cells were associated with greater quantities of genes uniquely expressed in these organs throughout all cell types and developmental stages assayed, including many TFs, some of which showed phylogenetic conservation, which could explain their organ specificity. Further phylogenetic analysis of large gene families could reveal evolutionary mechanisms underlying the diversification and specification of gene expression across cell types and organs.

While some high-resolution cell-type atlases exist for selected Arabidopsis organs, tissues, and cell types, the spatiotemporal axis of development,

derived from the non-renewing nature of plant cells, introduces additional levels of complexity that can obfuscate the interpretation of dissociation-based single-cell methods. An *a priori* understanding of cell-type specific developmental programs can aid in the understanding and annotation of single-cell datasets (*9*), but such datasets have yet to be generated for the entirety of Arabidopsis tissues and development.

Application of two spatial transcriptomics technologies to the germinating seed and silique tissues revealed that high-resolution spatial studies can disentangle plant organ heterogeneity, as evidenced by the validation of both cluster markers and cluster annotations in these highly dynamic tissues. As *in situ* spatial transcriptomic methodologies require knowledge of up to thousands of transcripts with distinct expression patterns, our datasets also function as a foundation for future spatial transcriptomics experiments

Overall, our dataset highlights the complexity of cell types throughout the entirety of development of an organism.

Our dataset also serves as a foundation for future studies that seek to further characterize these tissues and developmental stages, or atlases investigating stress or stimulus-driven responses, with access to the processed datasets for rapid exploration in a web-based interface available at our web portal (http://arabidopsisdevatlas.salk.edu/).

Acknowledgments

We would like to thank Genevieve Zzyzyx for the illustrations. We would also like to thank James Walker for help with annotation of pollen cells and manuscript input, and Hanqing Liu for reading and input to the manuscript. T.N. was supported by Human Frontiers Science Program (HFSP) Long-term Fellowship (LT000661/2020-L). N.I-E is a research fellow at the George E. Hewitt Foundation for Medical Research and an Awardee of the Weizmann Institute of Science – Israel National Postdoctoral Award Program for Advancing Women in Science.

References

1. J. C. Larkin, M. L. Brown, J. Schiefelbein, How do cells know what they want to be when they grow up? Lessons from epidermal patterning in Arabidopsis. *Annu. Rev. Plant Biol.* **54**, 403–430 (2003).

2. B. De Rybel, A. P. Mähönen, Y. Helariutta, D. Weijers, Plant vascular development: from early specification to differentiation. *Nat. Rev. Mol. Cell Biol.* **17**, 30–40 (2016).

3. B. Cole, D. Bergmann, C. E. Blaby-Haas, I. K. Blaby, K. E. Bouchard, S. M. Brady, D. Ciobanu, D. Coleman-Derr, S. Leiboff, J. C. Mortimer, T. Nobori, S. Y. Rhee, J. Schmutz, B. A. Simmons, A. K. Singh, N. Sinha, J. P. Vogel, R. C. O'Malley, A. Visel, D. E. Dickel, Plant single-cell solutions for energy and the environment. *Commun Biol*. **4**, 962 (2021).

4. I. W. Taylor, O. Rahul Patharkar, C.-W. Hsu, J. Baer, C. E. Niederhuth, U. Ohler, P. N. Benfey, J. C. Walker, Single-cell transcriptomics of the

Arabidopsis floral abscission zone. *bioRxiv* (2022), , doi:10.1101/2022.07.14.500021.

5. T.-Q. Zhang, Y. Chen, J.-W. Wang, A single-cell analysis of the Arabidopsis vegetative shoot apex. *Dev. Cell*. **56**, 1056–1074.e8 (2021).

6. Q. Song, A. Ando, N. Jiang, Y. Ikeda, Z. J. Chen, Single-cell RNA-seq analysis reveals ploidy-dependent and cell-specific transcriptome changes in Arabidopsis female gametophytes. *Genome Biol.* **21**, 178 (2020).

7. F. Apelt, E. Mavrothalassiti, S. Gupta, F. Machin, J. J. Olas, M. G. Annunziata, D. Schindelasch, F. Kragler, Shoot and root single cell sequencing reveals tissue- and daytime-specific transcriptome profiles. *Plant Physiol.* **188**, 861–878 (2022).

8. C. S. Misra, M. R. Santos, M. Rafael-Fernandes, N. P. Martins, M. Monteiro, J. D. Becker, Transcriptomics of Arabidopsis sperm cells at single-cell resolution. *Plant Reprod.* **32**, 29–38 (2019).

9. R. Shahan, C.-W. Hsu, T. M. Nolan, B. J. Cole, I. W. Taylor, L. Greenstreet, S. Zhang, A. Afanassiev, A. H. C. Vlot, G. Schiebinger, P. N. Benfey, U. Ohler, A single-cell Arabidopsis root atlas reveals developmental trajectories in wild-type and cell identity mutants. *Developmental Cell*. **57** (2022), pp. 543–560.

10. C. N. Shulse, B. J. Cole, D. Ciobanu, J. Lin, Y. Yoshinaga, M. Gouran, G. M. Turco, Y. Zhu, R. C. O'Malley, S. M. Brady, D. E. Dickel, High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. *Cell Rep.* **27**, 2241–2247.e4 (2019).

11. K. H. Ryu, L. Huang, H. M. Kang, J. Schiefelbein, Single-Cell RNA Sequencing Resolves Molecular Relationships Among Individual Plant Cells. *Plant Physiol.* **179**, 1444–1456 (2019).

12. D. C. Boyes, A. M. Zayed, R. Ascenzi, A. J. McCaskill, N. E. Hoffman, K. R. Davis, J. Görlach, Growth Stage–Based Phenotypic Analysis of Arabidopsis: A Model for High Throughput Functional Genomics in Plants. *Plant Cell*. **13**, 1499–1510 (2001).

13. C. Mizzotti, L. Rotasperti, M. Moretto, L. Tadini, F. Resentini, B. M. Galliani, M. Galbiati, K. Engelen, P. Pesaresi, S. Masiero, Time-Course Transcriptome Analysis of Arabidopsis Siliques Discloses Genes Essential for Fruit Development and Maturation. *Plant Physiol.* **178**, 1249–1268 (2018).

14. E. R. Alvarez-Buylla, M. Benítez, A. Corvera-Poiré, A. Chaos Cador, S. de Folter, A. Gamboa de Buen, A. Garay-Arroyo, B. García-Ponce, F. Jaimes-Miranda, R. V. Pérez-Ruiz, A. Piñeyro-Nelson, Y. E. Sánchez-Corrales, Flower development. *Arabidopsis Book*. **8**, e0127 (2010).

15. H. Li, J. Janssens, M. De Waegeneer, S. S. Kolluru, K. Davie, V. Gardeux, W. Saelens, F. David, M. Brbić, J. Leskovec, C. N. McLaughlin, Q. Xie, R. C. Jones, K. Brueckner, J. Shim, S. G. Tattikota, F. Schnorrer, K. Rust, T. G. Nystul, Z. Carvalho-Santos, C. Ribeiro, S. Pal, T. M. Przytycka, A. M. Allen, S. F. Goodwin, C. W. Berry, M. T. Fuller, H. White-Cooper, E. L. Matunis, S. DiNardo, A. Galenza, L. E. O'Brien, J. A. T. Dow, H. Jasper, B. Oliver, N. Perrimon, B. Deplancke, S. R. Quake, L. Luo, S. Aerts, FCA Consortium, Fly

Cell Atlas: a single-cell transcriptomic atlas of the adult fruit fly. *Science*. **375**, doi:10.1101/2021.07.04.451050.

16. A. Gómez-Felipe, M. Marconi, E. Branchini, B. Wang, H. Bertrand-Rakusova, T. Stan, J. Burkiewicz, S. de Folter, A.-L. Routier-Kierzkowska, K. Wabnik, D. Kierzkowski, Competing differentiation gradients coordinate fruit morphogenesis. *bioRxiv* (2023), , doi:10.1101/2023.01.19.524793.

17. C. L. Picard, R. A. Povilus, B. P. Williams, M. Gehring, Transcriptional and imprinting complexity in Arabidopsis seeds at single-nucleus resolution. *Nat Plants*. **7**, 730–738 (2021).

18. C. Grierson, E. Nielsen, T. Ketelaarc, J. Schiefelbein, Root hairs. *Arabidopsis Book*. **12**, e0172 (2014).

19. K. Ohashi-Ito, D. C. Bergmann, Arabidopsis FAMA controls the final proliferation/differentiation switch during stomatal development. *Plant Cell*. **18**, 2493–2505 (2006).

20. R. Stadler, N. Sauer, TheArabidopsis thaliana AtSUC2Gene is Specifically Expressed in Companion Cells. *Bot. Acta*. **109**, 299–306 (1996).

21. E. Truernit, N. Sauer, The promoter of the Arabidopsis thaliana SUC2 sucrose-H+ symporter gene directs expression of beta-glucuronidase to the phloem: evidence for phloem loading and unloading by SUC2. *Planta*. **196**, 564–570 (1995).

22. B. Bai, A. Peviani, S. van der Horst, M. Gamm, B. Snel, L. Bentsink, J. Hanson, Extensive translational regulation during seed germination revealed by polysomal profiling. *New Phytol.* **214**, 233–244 (2017).

23. K. S. Browning, J. Bailey-Serres, Mechanism of cytoplasmic mRNA translation. *Arabidopsis Book*. **13**, e0176 (2015).

24. M. Jakoby, B. Weisshaar, W. Dröge-Laser, J. Vicente-Carbajosa, J. Tiedemann, T. Kroj, F. Parcy, bZIP Research Group, bZIP transcription factors in Arabidopsis. *Trends Plant Sci.* **7**, 106–111 (2002).

25. M. Li, W. Lin, W. Hinckley, T. Yao, W. Muchero, J.-G. Chen, S. Carol Huang, Uncovering the “ZIP code” for bZIP dimers reveals novel motifs, regulatory rules and one billion years of cis-element evolution. *bioRxiv* (2022), , doi:10.1101/2022.04.17.488518.

26. R. De Smet, E. Sabaghian, Z. Li, Y. Saeys, Y. Van de Peer, Coordinated Functional Divergence of Genes after Genome Duplication in Arabidopsis thaliana. *Plant Cell*. **29**, 2786–2800 (2017).

27. R. R. Stickels, E. Murray, P. Kumar, J. Li, J. L. Marshall, D. J. Di Bella, P. Arlotta, E. Z. Macosko, F. Chen, Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* **39**, 313–319 (2021).

28. H. Liu, J. Zhou, W. Tian, C. Luo, A. Bartlett, A. Aldridge, J. Lucero, J. K. Osteen, J. R. Nery, H. Chen, A. Rivkin, R. G. Castanon, B. Clock, Y. E. Li, X. Hou, O. B. Poirion, S. Preissl, A. Pinto-Duarte, C. O’Connor, L. Boggeman, C. Fitzpatrick, M. Nunn, E. A. Mukamel, Z. Zhang, E. M. Callaway, B. Ren, J. R. Dixon, M. M. Behrens, J. R. Ecker, DNA methylation atlas of the mouse

brain at single-cell resolution. *Nature*. **598**, 120–128 (2021).

29. D. Kobak, P. Berens, The art of using t-SNE for single-cell transcriptomics. *Nat. Commun.* **10**, 5416 (2019).

30. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P.-R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods*. **16**, 1289–1296 (2019).

31. G. C. Linderman, M. Rachh, J. G. Hoskins, S. Steinerberger, Y. Kluger, Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data. *Nat. Methods*. **16**, 243–245 (2019).

32. K. Street, D. Risso, R. B. Fletcher, D. Das, J. Ngai, N. Yosef, E. Purdom, S. Dudoit, Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*. **19**, 477 (2018).

33. K. Van den Berge, H. Roux de Bézieux, K. Street, W. Saelens, R. Cannoodt, Y. Saeys, S. Dudoit, L. Clement, Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11**, 1201 (2020).

34. M. Krassowski, M. Arts, C. Lagger, Max, *krassowski/complex-upset: v1.3.5* (Zenodo, 2022; https://zenodo.org/record/3700590).

35. M. Alandete-Saez, M. Ron, S. McCormick, GEX3, expressed in the male gametophyte and in the egg cell of Arabidopsis thaliana, is essential for micropylar pollen tube guidance and plays a role during early embryogenesis. *Mol. Plant*. **1**, 586–598 (2008).

36. E. R. Alvarez-Buylla, S. J. Liljegren, S. Pelaz, S. E. Gold, C. Burgeff, G. S. Ditta, F. Vergara-Silva, M. F. Yanofsky, MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. *Plant J.* **24**, 457–466 (2000).

37. E. Andréasson, L. Bolt Jørgensen, A. S. Höglund, L. Rask, J. Meijer, Different myrosinase and idioblast distribution in Arabidopsis and Brassica napus. *Plant Physiol.* **127**, 1750–1763 (2001).

38. S. Baima, M. Possenti, A. Matteucci, E. Wisman, M. M. Altamura, I. Ruberti, G. Morelli, The arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiol.* **126**, 643–655 (2001).

39. C. Barth, G. Jander, Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense. *Plant J.* **46**, 549–562 (2006).

40. F. Borges, R. Gardner, T. Lopes, J. P. Calarco, L. C. Boavida, R. K. Slotkin, R. A. Martienssen, J. D. Becker, FACS-based purification of Arabidopsis microspores, sperm cells and vegetative nuclei. *Plant Methods*. **8**, 44 (2012).

41. Y. Chen, J. Takita, M. Hiwatari, T. Igarashi, R. Hanada, A. Kikuchi, T. Hongo, T. Taki, M. Ogasawara, A. Shimada, Y. Hayashi, Mutations of the PTPN11 and RAS genes in rhabdomyosarcoma and pediatric hematological malignancies. *Genes*

Chromosomes Cancer. **45**, 583–591 (2006).

42. L.-Q. Chen, B.-H. Hou, S. Lalonde, H. Takanaga, M. L. Hartung, X.-Q. Qu, W.-J. Guo, J.-G. Kim, W. Underwood, B. Chaudhuri, D. Chermak, G. Antony, F. F. White, S. C. Somerville, M. B. Mudgett, W. B. Frommer, Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature*. **468**, 527–532 (2010).

43. M. L. Churchman, M. L. Brown, N. Kato, V. Kirik, M. Hülskamp, D. Inzé, L. De Veylder, J. D. Walker, Z. Zheng, D. G. Oppenheimer, T. Gwin, J. Churchman, J. C. Larkin, SIAMESE, a plant-specific cell cycle regulator, controls endoreplication onset in Arabidopsis thaliana. *Plant Cell*. **18**, 3145–3157 (2006).

44. M. Endo, N. Mochizuki, T. Suzuki, A. Nagatani, CRYPTOCHROME2 in vascular bundles regulates flowering in Arabidopsis. *Plant Cell*. **19**, 84–93 (2007).

45. M. Endo, H. Shimizu, M. A. Nohales, T. Araki, S. A. Kay, Tissue-specific clocks in Arabidopsis show asymmetric coupling. *Nature*. **515**, 419–422 (2014).

46. M. L. Engel, R. Holmes-Davis, S. McCormick, Green sperm. Identification of male gamete promoters in Arabidopsis. *Plant Physiol.* **138**, 2124–2133 (2005).

47. L. Gan, K. Xia, J.-G. Chen, S. Wang, Functional characterization of TRICHOMELESS2, a new single-repeat R3 MYB transcription factor in the regulation of trichome patterning in Arabidopsis. *BMC Plant Biol.* **11**, 176 (2011).

48. M. A. Glaring, A. Zygadlo, D. Thorneycroft, A. Schulz, S. M. Smith, A. Blennow, L. Baunsgaard, An extra-plastidial α-glucan, water dikinase from Arabidopsis phosphorylates amylopectin in vitro and is not necessary for transient starch degradation. *J. Exp. Bot.* **58**, 3949–3960 (2007).

49. W.-J. Guo, W. Bundithya, P. B. Goldsbrough, Characterization of the Arabidopsis metallothionein gene family: tissue-specific expression and induction during senescence and in response to copper. *New Phytol.* **159**, 369–381 (2003).

50. P.-K. Hsu, Y.-F. Tsay, Two phloem nitrate transporters, NRT1.11 and NRT1.12, are important for redistributing xylem-borne nitrate to enhance plant growth. *Plant Physiol.* **163**, 844–856 (2013).

51. J. Huang, H. Wang, W. Liang, X. Xie, G. Guo, Developmental expression of Arabidopsis methyltransferase genes MET1, DRM2, and CMT3. *Mol. Biol.* **48**, 681–687 (2014).

52. H. Husebye, S. Chadchawan, P. Winge, O. P. Thangstad, A. M. Bones, Guard cell- and phloem idioblast-specific expression of thioglucoside glucohydrolase 1 (myrosinase) in Arabidopsis. *Plant Physiol.* **128**, 1180–1188 (2002).

53. H. Ilgenfritz, D. Bouyer, A. Schnittger, J. Mathur, V. Kirik, B. Schwab, N.-H. Chua, G. Jürgens, M. Hülskamp, The Arabidopsis STICHEL gene is a regulator of trichome branch number and encodes a novel protein. *Plant Physiol.* **131**, 643–655 (2003).

54. M. J. Jakoby, D. Falkenhan, M. T. Mader, G. Brininstool, E. Wischnitzki, N. Platz, A. Hudson, M. Hülskamp, J. Larkin, A. Schnittger, Transcriptional profiling of mature Arabidopsis trichomes reveals that NOECK encodes the MIXTA-like transcriptional regulator MYB106. *Plant Physiol.* **148**, 1583–1602 (2008).

55. Z. Kevei, M. Baloban, O. Da Ines, H. Tiricz, A. Kroll, K. Regulski, P. Mergaert, E. Kondorosi, Conserved CDC20 cell cycle functions are carried out by two of the five isoforms in Arabidopsis thaliana. *PLoS One*. **6**, e20618 (2011).

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

57. S. Kirschner, H. Woodfield, K. Prusko, M. Koczor, U. Gowik, J. M. Hibberd, P. Westhoff, Expression of SULTR2;2, encoding a low-affinity sulphur transporter, in the Arabidopsis bundle sheath and vein cells is mediated by a positive regulator. *J. Exp. Bot.* **69**, 4897–4906 (2018).

58. T. Kurata, C. Kawabata-Awai, E. Sakuradani, S. Shimizu, K. Okada, T. Wada, The YORE-YORE gene regulates multiple aspects of epidermal cell differentiation in Arabidopsis. *Plant J.* **36**, 55–66 (2003).

59. J. Long, J. Walker, W. She, B. Aldridge, H. Gao, S. Deans, M. Vickers, X. Feng, Nurse cell--derived small RNAs define paternal epigenetic inheritance in Arabidopsis. *Science*. **373**, eabh0556 (2021).

60. C. B. Lopez-Anido, A. Vatén, N. K. Smoot, N. Sharma, V. Guo, Y. Gong, M. X. Anleu Gil, A. K. Weimer, D. C. Bergmann, Single-cell resolution of lineage trajectories in the Arabidopsis stomatal lineage and developing leaf. *Dev. Cell*. **56**, 1043–1055.e4 (2021).

61. L. Lu, Y.-R. J. Lee, R. Pan, J. N. Maloof, B. Liu, An internal motor kinesin is associated with the Golgi apparatus and plays a role in trichome morphogenesis in Arabidopsis. *Mol. Biol. Cell*. **16**, 811–823 (2005).

62. X. Ma, L. Song, Y. Yang, D. Liu, A gain-of-function mutation in the ROC1 gene alters plant architecture in Arabidopsis. *New Phytol.* **197**, 751–762 (2013).

63. J. Mathur, G. Molnár, S. Fujioka, S. Takatsuto, A. Sakurai, T. Yokota, G. Adam, B. Voigt, F. Nagy, C. Maas, J. Schell, C. Koncz, M. Szekeres, Transcription of the Arabidopsis CPD gene, encoding a steroidogenic cytochrome P450, is negatively controlled by brassinosteroids. *Plant J.* **14**, 593–602 (1998).

64. J. Mitchell, N. K. Mukhtar, I. Skinner, G. W. Bassel, Gibberellin response in the embryo epidermis regulates germination uniformity in response to seed priming. *bioRxiv* (2018), , doi:10.1101/436121.

65. S. Müller, S. Han, L. G. Smith, Two Kinesins Are Involved in the Spatial Control of Cytokinesis in Arabidopsis thaliana. *Current Biology*. **16** (2006), pp. 888–894.

66. A. Mustroph, M. E. Zanetti, C. J. H. Jang, H. E. Holtan, P. P. Repetti, D. W. Galbraith, T. Girke, J. Bailey-Serres, Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 18843–18848 (2009).

67. N. Nakayama, J. M. Arroyo, J. Simorowski, B. May, R. Martienssen, V. F. Irish, Gene trap lines define domains of gene regulation in Arabidopsis petals and stamens. *Plant Cell*. **17**, 2486–2506 (2005).

68. Y. Ohashi, A. Oka, I. Ruberti, G. Morelli, T. Aoyama, Entopically additive expression of GLABRA2 alters the frequency and spacing of trichome initiation. *Plant J.* **29**, 359–369 (2002).

69. S. Okumoto, R. Schmidt, M. Tegeder, W. N. Fischer, D. Rentsch, W. B. Frommer, W. Koch, High affinity amino acid transporters specifically expressed in xylem parenchyma and developing seeds of Arabidopsis. *J. Biol. Chem.* **277**, 45338–45346 (2002).

70. G. Pilot, H. Stransky, D. F. Bushey, R. Pratelli, U. Ludewig, V. P. M. Wingate, W. B. Frommer, Overexpression of GLUTAMINE DUMPER1 leads to hypersecretion of glutamine from Hydathodes of Arabidopsis leaves. *Plant Cell*. **16**, 1827–1840 (2004).

71. B. Pommerrenig, K. Feussner, W. Zierer, V. Rabinovych, F. Klebl, I. Feussner, N. Sauer, Phloem-Specific Expression of Yang Cycle Genes and Identification of Novel Yang Cycle Enzymes in*Plantago*and*Arabidopsis*. *The Plant Cell*. **23** (2011), pp. 1904–1919.

72. C. Procko, T. Lee, A. Borsuk, B. O. R. Bargmann, T. Dabi, J. R. Nery, M. Estelle, L. Baird, C. O'Connor, C. Brodersen, J. R. Ecker, J. Chory, Leaf cell-specific and single-cell transcriptional profiling reveals a role for the palisade layer in UV light protection. *Plant Cell*. **34**, 3261–3279 (2022).

73. R. E. Pruitt, J. P. Vielle-Calzada, S. E. Ploense, U. Grossniklaus, S. J. Lolle, *FIDDLEHEAD*, a gene required to suppress epidermal cell interactions in *Arabidopsis*, encodes a putative lipid biosynthetic enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1311–1316 (2000).

74. J.-L. Qiu, R. Jilk, M. D. Marks, D. B. Szymanski, The Arabidopsis SPIKE1 gene is required for normal cell shape control and tissue development. *Plant Cell*. **14**, 101–118 (2002).

75. A. Ranjan, G. Fiene, P. Fackendahl, U. Hoecker, The Arabidopsis repressor of light signaling SPA1 acts in the phloem to regulate seedling de-etiolation, leaf expansion and flowering time. *Development*. **138**, 1851–1862 (2011).

76. C. Rautengarten, B. Ebert, T. Herter, C. J. Petzold, T. Ishii, A. Mukhopadhyay, B. Usadel, H. V. Scheller, The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in Arabidopsis. *Plant Cell*. **23**, 1373–1390 (2011).

77. D. T. Rouse, R. Marotta, R. W. Parish, Promoter and expression studies on an *Arabidopsis thaliana* dehydrin gene. *FEBS Letters*. **381** (1996), pp. 252–256.

78. M. G. Sawchuk, T. J. Donner, P. Head, E. Scarpella, Unique and Overlapping

Expression Patterns among Members of Photosynthesis-Associated Nuclear Gene Families in Arabidopsis. *Plant Physiology*. **148** (2008), pp. 1908–1924.

79. K. R. Siegfried, Y. Eshed, S. F. Baum, D. Otsuga, G. N. Drews, J. L. Bowman, Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. *Development*. **126**, 4117–4128 (1999).

80. M. L. Sistrunk, D. M. Antosiewicz, M. M. Purugganan, J. Braam, Arabidopsis TCH3 encodes a novel Ca2+ binding protein and shows environmentally induced and tissue-specific regulation. *Plant Cell*. **6**, 1553–1565 (1994).

81. D. B. Szymanski, R. A. Jilk, S. M. Pollock, M. D. Marks, Control of GL2 expression in Arabidopsis leaves and trichomes. *Development*. **125**, 1161–1171 (1998).

82. S. Takada, N. Takada, A. Yoshida, ATML1 promotes epidermal cell differentiation in Arabidopsis shoots. *Development*. **140**, 1919–1923 (2013).

83. O. P. Thangstad, B. Gilde, S. Chadchawan, M. Seem, H. Husebye, D. Bradley, A. M. Bones, Cell specific, cross-species expression of myrosinases in Brassica napus, Arabidopsis thaliana and Nicotiana tabacum. *Plant Mol. Biol.* **54**, 597–611 (2004).

84. K. Uemoto, T. Araki, M. Endo, Isolation of Arabidopsis Palisade and Spongy Mesophyll Cells. *Methods in Molecular Biology* (2018), pp. 141–148.

85. T. Werner, V. Motyka, V. Laucou, R. Smets, H. Van Onckelen, T. Schmülling, Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *Plant Cell*. **15**, 2532–2550 (2003).

86. H. Wu, L. Li, J. Du, Y. Yuan, X. Cheng, H.-Q. Ling, Molecular and Biochemical Characterization of the Fe(III) Chelate Reductase Gene Family in Arabidopsis thaliana. *Plant and Cell Physiology*. **46** (2005), pp. 1505–1514.

87. R. K. Yadav, M. Tavakkoli, M. Xie, T. Girke, G. V. Reddy, A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche. *Development*. **141**, 2735–2744 (2014).

88. A. Yephremov, E. Wisman, P. Huijser, C. Huijser, K. Wellesen, H. Saedler, Characterization of the FIDDLEHEAD gene of Arabidopsis reveals a link between adhesion response and cell differentiation in the epidermis. *Plant Cell*. **11**, 2187–2201 (1999).

89. S. I. Zandalinas, R. Mittler, Vascular and nonvascular transmission of systemic reactive oxygen signals during wounding and heat stress. *Plant Physiology*. **186** (2021), pp. 1721–1733.

90. X. Zhang, J. Dyachok, S. Krishnakumar, L. G. Smith, D. G. Oppenheimer, IRREGULAR TRICHOME BRANCH1 in Arabidopsis encodes a plant homolog of the actin-related protein2/3 complex activator Scar/WAVE that regulates actin and microtubule organization. *Plant Cell*. **17**, 2314–2326 (2005).

91. L. Zhang, Q. Tan, R. Lee, A. Trethewy, Y.-H. Lee, M. Tegeder, Altered xylem-phloem transfer of amino acids affects

metabolism and leads to increased seed yield and oil content in Arabidopsis. *Plant Cell*. **22**, 3603–3620 (2010).

92. Y. Zhu, L. Liu, L. Shen, H. Yu, NaKR1 regulates long-distance movement of FLOWERING LOCUS T in Arabidopsis. *Nat Plants*. **2**, 16075 (2016).

93. I. Zimmermann, R. Saedler, M. Mutondo, M. Hulskamp, The Arabidopsis GNARLED gene encodes the NAP125 homolog and controls several actin-based cell shape changes. *Mol. Genet. Genomics*. **272**, 290–296 (20

Funding:

J.R.E. is an Investigator of the Howard Hughes Medical Institute

Author contributions

Conceptualization: TL, TN, JRE

Methodology: TL, TN, NIE

Investigation: TL, TN, NIE, BJ, JRN

Data curation: TL, TN, JRN

Data analysis: TL, TN, JX

Visualization: TL, TN, NIE

Funding acquisition: JRE

Project administration: TL, TN, NIE

Supervision: JRE

Writing – original draft: TL, TN, NIE, JRE

Writing – review & editing: TL, TN, NIE, JRE

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All datasets are available for browsing at arabidopsisdevatlas.salk.edu. Raw data and processed datasets are available at GEO (accession number GSE226097).

Supplementary Materials

Materials and Methods

Figs. S1 to S12

Tables S1 to S6