Single-cell resolution of lineage trajectories in the Arabidopsis stomatal lineage and developing leaf

Lopez-Anido, Camila B.

2021-04-05

Lopez-Anido , C B , Vaten , A , Smoot , N K , Sharma , N , Guo , V , Gong , Y , Gil , M X A , Weimer , A K & Bergmann , D C 2021 , ' Single-cell resolution of lineage trajectories in the Arabidopsis stomatal lineage and developing leaf ' , Developmental Cell , vol. 56 , no. 7 , pp. 1043-+ . https://doi.org/10.1016/j.devcel.2021.03.014

http://hdl.handle.net/10138/342383
https://doi.org/10.1016/j.devcel.2021.03.014

HHS Public Access

Author manuscript

Dev Cell. Author manuscript.

Single-Cell Resolution of Lineage Trajectories in the Arabidopsis Stomatal Lineage and Developing Leaf

Camila B. Lopez-Anido[1,2], **Anne Vatén**[1,3], **Nicole K. Smoot**[2], **Nidhi Sharma**[2], **Victoria Guo**[2,4], **Yan Gong**[1], **M. Ximena Anleu Gil**[2,5], **Annika K. Weimer**[1,6], **Dominique C. Bergmann**[1,2,7,*]

[1]Department of Biology, Stanford University, Stanford, CA 94305-5020, USA [2]Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305-5020, USA [3]Present address: Organismal and Evolutionary Biology Research Programme, Biological and Environmental Sciences, University of Helsinki, Helsinki 00014, Finland [4]Present address: Genentech Inc, 1 DNA Way, South San Francisco, CA 94080 [5]Present address: Department of Plant Biology, 2203 Life Sciences Building, University of California, Davis, CA 95616, USA [6]Present address: Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA [7]Lead Contact

SUMMARY

Dynamic cell identities underlie flexible developmental programs. The stomatal lineages in the Arabidopsis leaf epidermis feature asynchronous and indeterminate divisions that can be modulated by environmental cues. The products of these lineages, stomatal guard cells and pavement cells, regulate plant-atmosphere exchanges, and the epidermis as a whole influences overall leaf growth. How flexibility is encoded in development of the stomatal lineage, and how cell fates are coordinated in the leaf are open questions. Here, by leveraging single-cell transcriptomics and molecular genetics, we uncovered models of cell differentiation within Arabidopsis leaf tissue. Profiles across leaf tissues identified points of regulatory congruence. In the stomatal lineage, single-cell resolution resolved underlying cell heterogeneity within early stages and provided a fine-grained profile of guard cell differentiation. Through integration of genome-scale datasets and spatiotemporally precise functional manipulations, we also identified

*Correspondence: dbergmann@stanford.edu.

AUTHOR CONTRIBUTIONS

C.L.A. conceived, designed, performed all experiments (except as noted below), analyzed results, performed statistical analysis, and wrote the manuscript. A.V. created the *MUTEp::amiSPCH* and *SPCH* rescue lines. M.X.A.G. contributed to imaging *SPCH* rescue lines. N.K.S., N.S., and V.G. contributed to creation and imaging of reporter lines. Y.G. performed time-lapse imaging. A.K.W. provided cell cycle insight. D.C.B. analyzed results and wrote the manuscript.

DECLARATION OF INTERESTS

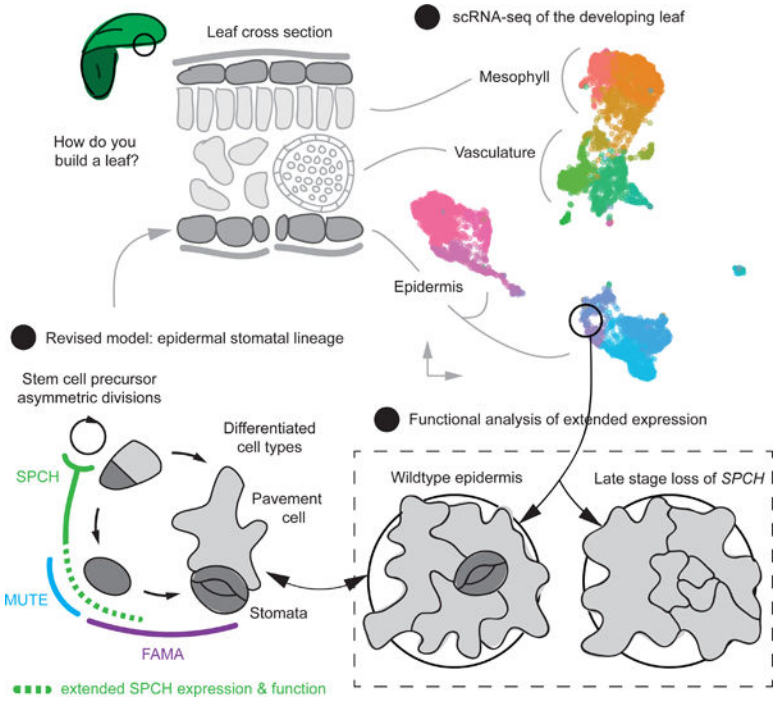
The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

an extended role for the transcriptional regulator SPEECHLESS in reinforcing cell fate commitment.

Graphical Abstract



eTOC blurb

How do cells build developmentally flexible organs? Lopez-Anido et al. employed single-cell transcriptomics to uncover distinct models of cell differentiation within Arabidopsis leaf tissue. Along with revealing dynamics of cellular programs, they identified underlying cell heterogeneity within the epidermal lineage and interrogated new roles for a core transcriptional regulator.

INTRODUCTION

Cell fate specification and differentiation are core features of developmental programs in multicellular organisms. Molecular genetics established a classical view of cell fate commitments as discrete and sequential stages. More recent studies, informed by technological advances with single-cell transcriptomics, however, have revealed ways in which cell fate decisions are continuous and heterogeneous. Remarkable progress has been made to elucidate cell lineage trajectories within a range of tissue and animal systems (Karaiskos et al., 2017; Tintori et al., 2016; Wagner et al., 2018). Much is to be learned from lineage dynamics in plants, which display extreme developmental flexibility and regenerative capacity. Moreover, given that multicellularity arose independently in plants and animals, comparing developmental strategies across kingdoms sheds light on what are universal rules, and what may be alternatives to build complex functional tissues.

An ever-growing number of single-cell transcriptomics studies have created tissue “atlases” and refined largely unidirectional differentiation processes in plant tissues (Nelms and Walbot, 2019; Ryu et al., 2019; Shulse et al., 2019; Xu et al., 2021). It remains unclear, however, how heterogenous cell identities may define flexible and indeterminate features of plant development. One exemplary model for flexible development is the stomatal lineage in the Arabidopsis leaf epidermis. Spatially dispersed stomatal lineage founder cells produce multipotent cells that ultimately differentiate into one of two fates: pavement cells or stomatal guard cells (Lee and Bergmann, 2019). Guard cells regulate the stomatal pore between them to enable plant-atmosphere gas exchange and photosynthesis. Stomatal lineages generate the majority of cells in the leaf epidermis, and developmental flexibility in response to the environment can result in different leaf sizes and cellular compositions (Hetherington and Woodward, 2003).

The stomatal lineage begins with asynchronous, indeterminate and self-renewing asymmetric divisions. These divisions are modulated by extrinsic environmental cues and cell-cell signaling cascades operating locally and systemically. Signaling inputs ultimately converge on the expression and behavior of lineage-specific transcription factors including the bHLH transcription factor SPEECHLESS (SPCH) (Lampard et al., 2008; Lau et al., 2014; Vatén et al., 2018). SPCH is expressed in asymmetrically dividing early stomatal lineage cells and is required for their self-renewing divisions (MacAlister et al., 2007; Pillitteri et al., 2007). Previous bulk RNA-seq studies of the stomatal lineage deepened our understanding of stomatal guard cell specification and maturation, but collapsed the complexity and diversity of the early lineage (Adrian et al., 2015). For example, bulk RNA-seq could not distinguish cells that had just entered the lineage from those that had undergone multiple rounds of asymmetric division, nor could it distinguish whether a multipotent early lineage cell was beginning the journey toward guard cell or toward pavement cell identity.

Here, we leveraged single-cell transcriptomics and molecular genetics to examine the dynamic development of the stomatal lineage in the context of the growing Arabidopsis leaf. A broader organ-level perspective provided insight into coordinated development, and enabled us to identify congruous models of cell differentiation within the vasculature, mesophyll, and epidermis. This set the context for a detailed examination of the stomatal lineage that included reconstructed trajectories derived from an additional >13,000 stomatal lineage cells. With fine-grained resolution of stomatal lineage cell fate commitment and differentiation, we correlated a global increase in transcriptional diversity with a developmental decision point and identified regulators that antagonistically regulate cell size at this critical junction. We found that the early flexible cell identities appear to exist along a continuum, with progressive cell specification, thus indicating that obligate core signaling cascades and transcriptional regulators control a range of cellular programs within lineage continuums. Collectively, our findings have refined the canonical stomatal lineage paradigm and uncovered progressive cell fate decisions that confer tissue flexibility.

RESULTS

Complementary single-cell transcriptomics approaches to define underlying cell heterogeneity

To capture the complexity of multiple cell populations from the developing leaf, we applied complementary single-cell RNA sequencing (scRNA-seq) techniques (Table 1). With this organ-level context, we can examine commonalities and distinctions between developmental lineages in different tissues. Because mesophyll tissue is highly abundant in leaves and recent leaf scRNA-seq datasets have mostly captured mesophyll transcriptomes (Kim et al., 2021; Liu et al., 2020), we included a fluorescence-activated cell sorting (FACS) step to obtain a more balanced representation of all cell types. Reporters driven by the *Arabidopsis thaliana MERISTEM LAYER 1 (ATML1)* promoter can be used for FACS enrichment of epidermal cells, without entirely eliminating internal cell types (Adrian et al., 2015; Tian et al., 2019; Yadav et al., 2014). In addition, we created scRNA-seq transcriptomes from just the stomatal lineage using a reporter for *TOO MANY MOUTHS (TMM)* (Nadeau and Sack, 2002). Our comprehensive approach with two targeted populations and plate-based Smart-seq2 (Picelli et al., 2014) and droplet-based 10X Genomics (Zheng et al., 2017) technologies thus enabled us to analyze a range of cytometry data, gene coverage, and cell identities. Importantly, this includes cell identities that are relatively rare and transient within a lineage. Here, we define cell states and identities as instances of putative cell transitions or types (e.g. specific precursor or mature cell), encompassing the range of possibilities for what it means to be a cell along a lineage trajectory (Morris, 2019).

Distinct models of cell differentiation and lineage trajectories within leaf tissue

As the behavior of stomatal lineage cells is defined, in part, by their organ context, we first aimed to generate datasets that would represent multiple leaf tissues and allow us to computationally reconstruct lineages. Following FACS enrichment, *ATML1*-expressing cells were captured by 10X Genomics microfluidics (Zheng et al., 2017) for downstream library preparation and sequencing. With the Seurat v3 analysis pipeline (Butler et al., 2018), we validated our dataset though a variety of complementary approaches (methods, Table 1, Figure S1–2). Graph-based unsupervised clustering and differential expression analysis provided insight into vasculature, mesophyll, and epidermal tissues (Figure S1A, Table S3). Clusters of cells were visualized on a Uniform Manifold Approximation and Projection (UMAP) plot, wherein local and global graph structure is preserved to represent cell similarity within and between clusters (Becht et al., 2018; McInnes et al., 2018). By removing confounding effects of the cell cycle (Figure S1B), we found that vasculature and mesophyll cell identities laid adjacent to each other within our UMAP, while younger epidermal cells were separated from differentiated guard cells. To further validate our dataset, we generated and characterized reporters of a selection of differentially expressed genes (Figure S1E), which provided additional *in planta* support for our epidermal cluster assignments. To elucidate transient identities and fate decisions within leaf tissues, we inferred transcriptional dynamics with a steady-state deterministic model using scVelo (Figure 1A) (Bergen et al., 2020; La Manno et al., 2018). scVelo predicts future cell states by leveraging ratios of nascent (unspliced) and mature (spliced) RNA transcripts. Thus, we identified putative, directed transitions between cell identities (Figure 1B–E), which were

corroborated by previously described gene expression profiles *in vivo*. Our approach also revealed enriched components of signaling pathways and transcriptional control that define these transitions in the vasculature, mesophyll, and epidermis (Table S3–S4).

Genetic programs that drive differentiation of vascular tissues are well defined in the stem and root (Hellmann et al., 2018), and thus provide a useful reference to interrogate the scVelo inference trajectory. Using known markers, we identified procambial meristematic cells, vasculature precursor cells, and maturing phloem and xylem cells (Figure 1B). The leaf vasculature exhibited similar lineage trajectories as has been described in other organs. Expression of *SUPPRESSOR OF MAX2-LIKE* 5 (*SMXL5*) (Wallner et al., 2017) and *TARGET OF MONOPTEROS 6* (*TMO6/DOF5.3*) (Schlereth et al., 2010) demarcated differentiating procambium and phloem, while *SWEET11* (Chen et al., 2012) defined potential companion and parenchyma cells. *ALTERED PHLOEM DEVELOPMENT (APL*) (Bonke et al., 2003) was expressed in mature phloem cells, downstream of the predicted onset of *SMXL5* and *SWEET11*. *TMO5/BHLH32* (De Rybel et al., 2013) was expressed in the procambium and overlapped with *SMXL5* expression. *TMO5* was also expressed in the xylem, along with *PHLOEM INTERCALATED WITH XYLEM (PXY*) (Fisher and Turner, 2007). Downstream of *TMO5* and *PXY* expression in our lineage inference, *VASCULAR RELATED NAC-DOMAIN PROTEIN 7 (VND7*) (Kubo et al., 2005) was restricted to differentiating xylem cells with activated secondary wall biosynthesis programs. Our differential expression analysis also identified several putative transcription factors with unappreciated roles in the vasculature (e.g. *BZIP9, AGL15, and DOF2)* (Table S3C,G). A comparison with root scRNA-seq data (Ryu et al., 2019) indicated similar restriction of expression to phloem or xylem cell types. Further analysis is necessary to define the function of these genes. However, as a general strategy for prioritizing potential lineage regulators, it is notable to detect the consistent association of a gene with common cell types from different datasets. Accordingly, as this work was in review, leaf vasculature expression of *BZIP9* was confirmed (Kim et al., 2021).

Given that we were able to elucidate vasculature cell fate decisions, we next asked whether our scRNA-seq dataset could provide insight into developmental events unique to leaves. Notably, vascular and mesophyll tissue share a common origin, the ground meristem. It has been previously challenging to profile such cells, but our trajectory inference delineated distinct cell fate decisions that distinguish the mesophyll and vasculature (Figure 1C). Mesophyll and vasculature clusters appear bridged by putative ground meristem and procambial cells that give rise to differentiating cells within each tissue. For instance, expression onset of *PHOTOSYSTEM I LIGHT HARVESTING COMPLEX GENE 6* (*LHCA6*) and *ARABIDOPSIS HOMEOBOX GENE 8* (*ATHB8),* previously characterized progenitor markers (Sawchuk et al., 2008; Scarpella et al., 2004), was detected in ground cells that are committed to mesophyll and vasculature development, respectively. We thus leveraged differentially expressed genes (Table S3C,G) to identify potential regulators of the ground meristem, which yielded the transcription factor genes *MYB15* and *HAG1*/*MYB28*. Further analysis of downstream putative regulators of mesophyll differentiation identified the transcription factor gene *AT5G50915*. *AT5G50915* expression overlapped with that of *STOMAGEN/EPFL9*, a gene encoding a secreted peptide that mediates epidermal-mesophyll tissue coordination to control stomatal density (Hunt et al., 2010; Kondo et al.,

2010; Sugano et al., 2010). Thus, our scRNA-seq dataset of the developing leaf provides models that define both the common origin and distinct developmental programs of inner tissues.

We then focused on directed developmental transitions within the epidermis, represented by two main clusters in our UMAP of the developing leaf. One cluster featured expression of the bHLH transcription factor gene *FAMA* (Ohashi-Ito and Bergmann, 2006), which indicated that it represents differentiating stomatal guard cells (Figure 1D), from which we also identified adaxial and abaxial guard cells (Figure S3, Table S4A). Within this cluster, the transcription factor *FOUR LIPS (FLP)/MYB124* exhibited much more restricted expression, consistent with reports linking *FLP* to the earliest stages of guard cell differentiation (Lai et al., 2005). On the other end of the cluster and guided by scVelo arrows, mature guard cells expressed *STOMATAL CARPENTER 1* (*SCAP1)* (Negi et al., 2013). *FAMA* gene expression was also detected in these mature guard cells, consistent with the possibility that FAMA plays critical roles in guard cell maturation and fate maintenance (Matos et al., 2014). With *FAMA*, *FLP*, and *SCAP1* gene profiles as guides of differentiation states, differential gene expression analysis identified additional putative regulators, such as the transcription factors *GATA2* and *HAT5* (Figure 1D) that respectively mark the beginning and end of guard cell differentiation.

The second epidermal cluster represented young epidermal and stomatal lineage cells (Figure 1E). A notable feature of this cluster was its diverging trajectory inference (depicted as deviating arrows on the UMAP, Figure 1A) modeling multiple cell fate decisions. A finer examination of these states derived from the *TMM*-enriched scRNA-seq data follows below (Figure 3), but at this initial resolution, we observed relatively broad expression of *SPCH* in early stomatal lineage stem cells. Conversely, expression of *SPCH's* paralogue *MUTE* (MacAlister et al., 2007; Pillitteri et al., 2007) was restricted to cells committing to guard cell differentiation. These committed cells in the young epidermal cluster were oriented toward the guard cell cluster, with corresponding lineage inference predictions (Figure 1A). Diverging trajectories towards the opposite end of the young epidermal cluster were correlated with expression onset of the transcription factor gene *MYB7*. MYB7 has been associated with production of secondary metabolite phenylpropanoids (Albert, 2015) indicating progression toward the alternative fate of epidermal pavement cell. Mature pavement cells are not captured in our datasets here due to cell size selection. Genes encoding cuticular wax biosynthesis and defense-related factors were co-expressed with *MYB7* (Table S3A,E), and overall there were a substantial number of putative regulators dynamically expressed in patterns traversing the young epidermal clusters. In contrast to previous bulk RNA-seq studies of cells expressing specific markers (e.g. *SPCH*, *MUTE*, and *FAMA*) (Adrian et al., 2015), employing the broadly expressed *ATML1* gene promoter with scRNA-seq enabled us to capture cells that take alternative trajectories to generate either stomata or pavement cells.

Cell cycle regulators exhibit distinct, yet overlapping expression profiles

We next asked whether comparisons between leaf tissues can serve to define distinct and overlapping features of common regulatory programs. We hypothesized that common

programs integrate with, and are controlled by, different tissue-specific contexts. Given that we were able to disentangle cell cycle signatures from cell lineage identities (Figure S1B), we were poised to address the foundational question of how the cell cycle may be coordinated with cell fate. As plant genomes encode several core cyclin-dependent kinase and cyclin regulators, which can exhibit lineage-specific expression and function (Adrian et al., 2015; Gutierrez, 2016; Han et al., 2018; Sozzani et al., 2010; Weimer et al., 2018), the resolution of single-cell transcriptomics could elucidate unique or repeated regulatory programs.

We found that core regulators exhibit distinct, yet overlapping, expression (Figure 2A, S3). G1/S-phase transition genes *CYCLIN-DEPENDENT KINASE A;1 (CDKA;1), CYCLIN D3;2 (CYCD3;2), and RETINOBLASTOMA-RELATED PROTEIN 1 (RBR1)* were broadly expressed (Figure 2B), while members of the G2/M-phase *CDKB* and *CYCB* gene families were expressed in distinct clusters (Figure 2C). *CYCA* gene family members, expressed at G1/S and G2/M, displayed a range of profiles within the dataset (Figure 2C). For instance, *CYCA2;3* was widely expressed, but expression of the G1/S-phase gene *CYCA3;1* (Takahashi et al., 2010) was more restricted. We also found both broad and tissue-specific expression profiles among G1/S-phase *CYCD* gene family members (Figure 2D–E). *CYCD3;2* was widely expressed, but *CYCD3;3* and *CYCD2;1* were selectively expressed in distinct mature phloem clusters and *CYCD6;1* was expressed in the procambium/xylem. The expression patterns of CYCDs connected to stomatal development appeared strikingly divergent (Han et al., 2018; Weimer et al., 2018), with *CYCD5;1* expressed in few cells in many tissues, and *CYCD7;1* exclusively overlapping with *MUTE* in the epidermis. Interestingly, the kinase *WEE1*, classically described in yeast as a cell size regulator that inhibits the G2/M transition (Nurse and Thuriaux, 1980), is expressed in the vascular and epidermal clusters where restricted G1/S-phase *CYCDs* converge (Figure 2E).

Integrated analysis elucidates flexible stomatal lineage cell states along a continuum

As distinct models of cell differentiation were observed in our leaf scRNA-seq dataset (Figure 1), we next pursued a more comprehensive, integrated approach focused solely on stomatal lineage cells. Briefly, the early stomatal lineage features asynchronous, asymmetric divisions of self-renewing stem cells – smaller meristemoids and larger stomatal lineage ground cells (SLGCs) – either of which can divide to yield another meristemoid/SLGC pair or differentiate into distinct epidermal cell types (Nadeau and Sack, 2002) (Figure 3A). After an unspecified number of asymmetric divisions, meristemoids differentiate into guard mother cells (GMCs), which divide symmetrically to yield a pair of stomatal guard cells, and SLGCs eventually endoreduplicate and differentiate into pavement cells.

We leveraged 10X Genomics and Smart-seq2 technologies to profile stomatal lineage cells FACS-isolated with the *TMM*-driven reporter, and we integrated technical replicates from independent pools of tissue (detailed in the methods, Table 1 and Figure S2, S4) with Seurat v3 (Butler et al., 2018). Potential cell states were identified through graph-based clustering and differential gene expression analysis (Figure S4, 3B and Table S5A). To reconstruct a stomatal lineage model that reflects transitions between cell identities and fate decisions, we generated a pseudotime axis with simultaneous principle curves using slingshot (Street et al.,

2018) (Figure 3B). Lineage trajectories were annotated using a similar approach as for the leaf, assessing expression profiles of genes that encode known players in the lineage, along with genes that display differential profiles across the clusters (Figure 3C–D).

The integrated stomatal lineage scRNA-seq dataset corroborated and extended our understanding from the single-cell resolution map of the developing leaf. A core of continuous states (m1, m2, m3) was defined by peak expression of bHLH transcription factors *SPCH* and *INDUCER OF CBF EXPRESSION 1/SCREAM* (Kanaoka et al., 2008) and represented flexible cell states early in the lineage existing along a pseudotime continuum (Figure 3B–D). Spatially distant clusters (sf1, sf2, sf3, sf4, sf5) represented distinct stages through deterministic guard cell fate commitment, with the expression of *MUTE* followed by *FAMA* (Figure 3B–D). Extending from the m1-m3 continuum in the opposite direction was the alternative fate trajectory representing SLGC differentiation towards pavement cell identity (clusters af1, af2, af3). While there are no well-characterized, specific markers that demark SLGC identity or differentiation activities ascribed to pavement cell formation and function were enriched in clusters af2 and af3. For example, the production of specialized metabolites has been ascribed to the activities of transcription factors MYB7 and MYB4, and the acquisition of polarized and lobed morphologies is related to the activities of the auxin regulator MACCHI-BOU4 (Furutani et al., 2014) and CASPARIAN STRIP MEMBRANE DOMAIN proteins (Roppolo et al., 2014) (Table S5D). We further validated our reconstructed lineage model of cell fate dynamics *in vivo* by generating reporters of genes with no known lineage function. Genes selected to represent broad and restricted expression patterns in pseudotime recapitulated those patterns *in planta* (Figure 3D–E).

The clusters m1, m2, and m3 are contiguous, yet across replicates and scRNA-seq technologies, the UMAP placement of these three clusters, and the number of cells that comprise them, are quite consistent (Figure S4A–B). Sorting with the TMM reporter enriches for these cells relative to the epidermal fraction of the *ATML1* promoter-sorted sample (Figure S2), as would be expected from reporter expression patterns. Clusters m1-m3, therefore, likely represent subpopulations of asymmetrically dividing precursor cells (meristemoids and SLGCs). Analysis of genes differentially expressed among these clusters and relative to other populations (Figure 3B–C, Table S5A) revealed an enrichment in metabolic and hormone pathway processes and processes involved in stress and environmental responses (Table S5D). The position of clusters m1, m2, and m3 relative to cluster sf1 (initiation of GMC fate commitment) in the pseudotime trajectory suggests that there is not a single linear path towards stomatal differentiation (Figure 3B–C). Cluster m1 shares enriched gene expression profiles with af2, and it is plausible that m1 represents multipotent SLGCs. Ribosomal gene expression is enriched in cluster m3, a feature it shares with sf1 and sf2. In sf1 and sf2, enriched ribosomal gene expression portends entry into G1/S-phase. It is not yet clear whether cluster m3 has committed to a GMC G1/S-phase program, or whether m3 cells could cycle back to m2 and m1 cell states.

Single-cell resolution defines cell fate commitment to yield stomata

The large number of stomatal linage cells in our reconstruction led to an extraordinarily fine-scale dissection of GMC fate commitment. This unidirectional differentiation process coincides with the onset of *MUTE* expression, involves a single symmetric cell division, and results in the formation of two sister guard cells. After initiation of fate commitment (cluster sf1, defined by *MUTE* gene expression), transition through G1/S and G2/M phases of the cell cycle can be seen in clusters sf2, sf3, sf4, and sf5. Cell cycle phases also correlated with a discrete border that distinguishes *MUTE* and *FAMA* expression domains (between clusters sf2 and sf3, Figure 3D).

We observed increased transcriptional diversity during GMC fate commitment, followed by low diversity in mature guard cells (clusters sf1, sf2, sf3, sf4, sf5; Figure 3F). This feature of lineage progression has been described in animal and human systems, where high diversity is typically anticorrelated with differentiation (Gulati et al., 2020). What is unusual about the stomatal lineage, however, is that diversity peaks in late GMCs, not during the early flexible stages (cluster sf3, Figure 3F). The same trend appeared in plate-based Smart-seq2 datasets and using FACS forward scatter measurements (FSC) as a relative estimate of cell size, we found that cell size positively correlated with transcriptional diversity in differentiating GMCs (clusters sf1, sf2, sf3; Figure 3G). A drop-off in both diversity and size was detected after the inferred symmetric division (clusters sf4, sf5), which was preceded by the onset of *FAMA* expression (cluster sf3; Figure 3D). In contrast, cell sizes and transcriptional diversity showed no clear correlation in the early lineage clusters (m1, m2, m3; Figure 3F–G).

Broad, concomitant changes in gene expression often have origins in the remodeling of the chromatin landscape. We therefore surveyed expression of specific transcription factors, chromatin remodelers, and cell-cycle regulators for differential expression along the pseudotime axis (Table S5C). Members of all these functional classes were differentially expressed, but we were especially struck by the expression profiles of cell cycle regulators (Figure 3D). *MUTE* was co-expressed with *CYCD7;1* and *WEE1* (clusters sf1, sf2), prior to the peak of transcriptional diversity and increased cell size in GMCs (cluster sf3; Figure 3F–G). *CYCD7;1* had already been shown to regulate GMC size and progression through G1/S (Weimer et al., 2018), and we again observed larger GMCs in 4 dpg *cycd7;1*−/− cotyledons (Figure 3H). We found, however, that GMCs from 4 dpg *wee1*−/− cotyledons were smaller than wildtype. The combined loss of *CYCD7;1* and *WEE1* yielded an intermediate phenotype (Figure 3H), which suggests that the two may be antagonistic inputs into GMC fate progression during the G1/S phase transition.

Broadened functional roles of an early lineage core regulator

A paradigm of stomatal lineage progression is that sequential function of SPCH, MUTE, and FAMA defines corresponding discrete stages (MacAlister et al., 2007; Ohashi-Ito and Bergmann, 2006; Pillitteri et al., 2007). While we found a distinct division between *MUTE* and *FAMA* gene expression domains from pseudotime reconstruction (Figure 2E, 3D), expression overlays of *SPCH*, *MUTE* and *FAMA* revealed that *SPCH* persisted beyond early stages of the lineage and was co-expressed with either *MUTE* or *FAMA* (Figure 3D, 4A,

S4D). Given the potential conflict of these findings with the sequential model of discrete functions, we reexamined *SPCH* expression *in vivo*. We monitored a published reporter (coding sequence, *cdsSPCH*) (Davies and Bergmann, 2014; Vatén et al., 2018) and a new reporter that contained more of the *SPCH* genomic context (genomic sequence, *gSPCH*), which may more accurately recapitulate regulatory control at the endogenous locus. *SPCH* expression could be detected in dividing GMCs by time-lapse imaging (Figure 4B). With the *gSPCH* reporter, we also detected signal in young guard cells (5.5% or 6/108 total guard cells from 4 dpg cotyledons, representative image in Figure S4E). Thus, extended *SPCH* expression detected by scRNA-seq could be corroborated through different assays and suggests that this core regulator may serve unappreciated roles in maintaining later cell fates.

We integrated genome-wide SPCH ChIP-seq and related RNA-seq datasets (Lau et al., 2014; Lee et al., 2019) with our scRNA-seq dataset to identify putative SPCH targets throughout the lineage. Many targets were restricted in their expression to only some clusters and cell states (Table S5A). They also represented a diverse range of enriched gene ontologies, from stimulus and immune responses to regulation of growth and division processes (Table S5D). As SPCH function likely shifts within different cellular states, it is interesting to consider the potential new role of SPCH in GMCs, where it appears to sequentially overlap with both *MUTE* and *FAMA* expression and targets genes implicated in cell fate and cell cycle control including *STOMATAL DENSITY AND DISTRIBUTION 1* (*SDD1* (von Groll et al., 2002) and *KNOLLE* (Lauber et al., 1997; Lukowitz et al., 1996) (Figure 4C). Genes that negatively correlate with *SPCH* expression were detected in the mature guard cell clusters (Table S5A).

A plausible role for the later-stage SPCH expression, given the identified targets, is to promote completion of the guard cell fate trajectory. The first hint that SPCH contributes to guard cell differentiation came from examination of mature tissue (15 dpg cotyledons) in lines where the *spch-3* null mutant (MacAlister et al., 2007) was rescued by either *cdsSPCH* or *gSPCH* constructs. *gSPCH*-rescued lines were nearly indistinguishable from wildtype, but there were significantly fewer stomata in *cdsSPCH*-rescued lines (Figure 4Di–iv). Moreover, *cdsSPCH* tissue also displayed an unusual phenotype, which consisted of seemingly “diverted” lineage cells that have undergone pavement cell-like lobing, rather than producing the expected meristemoid/SLGC or guard cell pair (Figure 4Dii, white asterisk). This fate defect could be due to disrupted levels and timing of *SPCH* expression, given that *cdsSPCH* transgene expression was fainter at later stages relative to *gSPCH* (Figure 4B). We therefore examined whether SPCH is specifically required at this later developmental transition by selectively depleting late-stage expression using artificial microRNAs driven by the *MUTE* gene promoter (*MUTEp::amiSPCH).* We observed additional rounds of early lineage divisions before fate commitment (Figure 4Dv–vii, S4F), as well as diverted cell identities, similar to what we found in the *cdsSPCH* rescue lines (Figure 4Dvi, white asterisks). Such outcomes were not observed in control seedlings. Overall, along with identifying potential SPCH-controlled programs that may be integrated for early lineage progression, we find that fine-tuned *SPCH* expression is required to maintain GMC fate commitment.

DISCUSSION

Plants display flexible and adaptive growth and patterning that offer insight into exceptional developmental strategies. Here, we uncovered distinct models of cell differentiation within Arabidopsis leaf tissue by leveraging single-cell transcriptomics and functional molecular genetics. A comparative analysis of cell identities and trajectories in the vasculature, mesophyll, and epidermal clusters revealed tissue-specific cellular programs (Figure 1–2, Table S3), along with adaxial/abaxial tissue polarity signatures (Figure S3, Table S4). Profiling an additional ~13,000 cells derived solely from stomatal lineages in the leaf epidermis resolved differentiation trajectories towards either stomatal fate or an alternative fate previously characterized only by cell morphology. Clusters of potentially uncommitted cells (m1, m2, and m3 in the epidermis), although distinct and reproducible, were characterized by a modest enrichment of many genes (Figure 3, Table S5). Pseudotime trajectories suggested multiple paths towards stomatal differentiation paralleling time-lapse data showing that GMCs originate from morphologically heterogenous cells (Davies and Bergmann, 2014). We hypothesize that the propensity of a given cell to commit to stomatal or alternative fates may be due to rapid, local, or even stochastic events rather than a gradual accumulation of potential.

Unexpected detection of the early stage regulator *SPCH* in later cell clusters forced us to revisit a previous model of sequential, discrete control of stomatal development by SPCH, MUTE, and FAMA (Lee and Bergmann, 2019; MacAlister et al., 2007; Ohashi-Ito and Bergmann, 2006; Pillitteri et al., 2007). Through cell-type specific functional manipulations, we showed that late *SPCH* depletion delayed lineage progression and diverted cells from guard cell fate. The observed diverted cell fates have not been described for other loss-of-function mutations, but do result from ectopic overexpression of the signaling ligand EPF1 (Hara et al., 2007, 2009; Triviño et al., 2013). Consistent with this, EPF1 represses SPCH function through kinase signaling cascades (Lampard et al., 2008). Thus, in our new model of the stomatal lineage, we postulate that SPCH serves dynamic roles during early flexible cell states, and also complements MUTE and FAMA at later stages to drive cell fate commitment and differentiation of guard cells. Signaling and regulatory modalities that selectively target SPCH may be recruited into these later stages. For example, when Arabidopsis is subjected to environmental stresses (e.g. heat or drought), the stomatal lineage response has centered on transcriptional and post-translational control of SPCH (Han et al., 2020; Kumari et al., 2014; Lau et al., 2014). The preference for SPCH was often ascribed to its role in early, flexible development. However, SPCH protein also has unique domains that render it a partner or target of regulatory proteins that MUTE and FAMA cannot engage (Lampard et al., 2008; Putarjunan et al., 2019), and thus it may be that SPCH is necessary to transmit such environmental information to later decisions.

Our work also highlighted distinct and overlapping features of common regulatory programs within a developing leaf. We found that cell cycle regulators integrated with tissue-specific regulatory contexts and may coordinate cell fate. We see parallels to the stomatal lineage in the vasculature cluster, two places where cell states are dependent on complexes containing RBR1 and a lineage-specific transcription factor (FAMA and SCARECROW (SCR), respectively) (Cruz-Ramírez et al., 2012; Matos et al., 2014). We found that there may be

shared regulatory logic between CYCD5;1, WEE1, and CYCD6;1/CYCD7;1 in these decisions as well. CYCD6;1 promotes asymmetric divisions of the root cortex-endodermal initials and is a direct target of a complex containing SCR and SHORTROOT (SHR) (Sozzani et al., 2010). The rapid upregulation of the *CYCD6;1* gene in response to SHR induction in roots is well characterized, though less appreciated may be the concomitant upregulation of *WEE1* and *CYCD5;1* (Sozzani et al., 2010). Moreover, *SHR* is required for leaf growth in part through upregulation of genes that promote the G1/S transition, including *CYCD5;1* (Dhondt et al., 2010). In the stomatal lineage, MUTE targets *CYCD5;1*, but not *CYCD7;1* and *WEE1* is also upregulated upon induced *MUTE* expression (Han et al., 2018). We showed that, like *CYCD5;1* and *CYCD7;1*, *WEE1* was required for GMC size (Figure 3H), providing the first report of a function for the kinase WEE1 during normal plant development. Why might expression of these cell cycle regulators converge at specific developmental transitions? Connections between cell cycle regulators, chromatin modification, and differentiation have been explored in other contexts (Ma et al., 2015). An emerging theme is a differential modulation of G1 and S phases upon cell fate commitment that facilitates remodeling of chromatin and the creation of a stable transcriptional state.

It will be interesting to leverage our developing leaf dataset to examine expression trends among components of transcriptional regulatory networks, signaling cascades and hormone responses. Integration of scRNA-seq data with SPCH ChIP-seq and RNA-seq was useful to define genes downstream of SPCH at various stages. Similar integration could be performed with existing genome-wide data on targets of the transcription factors that guide development of internal tissues. Single-cell transcriptomics ultimately provides a framework to compare molecular mechanisms of specific cell fate transitions in plant development between organs, and species, as well as with animal development. It will be exciting to consider evolutionarily divergent and convergent deployment of specific molecular regulators and developmental strategies.

Limitations of the study

The core limitation of the study is that we assume gene expression profiles largely capture the cellular programs and developmental features that define cell identity. Complementary genomic and proteomic approaches will be necessary to build more complex gene regulatory networks and to understand the contribution of protein-protein interactions and post-transcriptional regulatory mechanisms to flexible development. In addition, the isolation of cells for scRNA-seq necessarily removes them from their native context, and physical enrichment for rare cell types may introduce expression changes that obscure interesting biological signals.

STAR METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dominique Bergmann (dbergmann@stanford.edu).

Material availability—There are no restrictions on materials generated for this manuscript.

Data and code availability—The scRNA-seq datasets generated in this study are available at the Gene Expression Omnibus (GEO) under accession number GSE167135. To analyze the datasets, computational methods and code used were from published software described in the Methods details / Quantification and statistical analysis sections and the Key Resources Table.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plant material and growth conditions—All transgenic lines were examined in the *Arabidopsis thaliana* Columbia-0 (Col) ecotype, which served as the wildtype control in all experiments. Previously described transgenics used included the following: *ATML1p::YFP-RCI2A* (Roeder et al., 2010), *TMMp::TMM-YFP* (Nadeau and Sack, 2002), *SPCHp::cdsSPCH-YFP* and *SPCHp::gSPCH-YFP* (Vatén et al., 2018). Seedlings were grown on half-strength (½) Murashige and Skoog (MS) medium (Caisson Labs MSP01) at 22°C under long-day conditions (16 hr light and 8 hr dark cycles) and were used at the indicated day post-germination (dpg) for respective experiments.

METHOD DETAILS

Vector construction and plant transformation—Transgenic promoter reporters fused to a nuclear localized version of YFP (*YFP-NLS*) were generated using the Gateway system (Invitrogen). In brief, promoters were cloned into pENTR vectors and recombined with vectors bearing *YFP-NLS*, employing the binary R4pGWB destination vector system (Nakagawa et al., 2008; Tanaka et al., 2011). To generate the *MUTEp::amiSPCH* transgene, the *artificial microRNA* (*amiRNA*) sequence was designed with the Web MicroRNA Designer platform (http://wmd3.weigelworld.org), engineered using the pRS300 plasmid as template (Schwab et al., 2006). Together with the *MUTE* promoter, the engineered plasmid was cloned into the destination vector pGWB. Primer sequences used to generate or verify transgenics are provided in Table S1, and all referenced genes are included in Table S2. Transgenic plants were generated by Agrobacterium-mediated transformation (Clough, 2005; Clough and Bent, 1998), and seedlings were selected on ½ MS plates supplemented with respective antibiotics.

Analysis of *in vivo* expression and phenotypes—Seedlings from at least two independent lines of transcriptional and translational reporters were collected and analyzed. Fluorescence images were taken with a Leica TCS SP5 microscope (Leica Microsystems) and processed using Fiji (Schindelin et al., 2012). Cell outlines were visualized by either the plasma membrane marker *ATML1p::mCherry-RCI2A* or with propidium iodide (Molecular Probes, P3566). To quantify developmental phenotypes, we employed previously described homozygous T-DNA lines of *cycd7;1–1* FLAG_369E02 (Weimer et al., 2018) and *wee1–1* GABI_270E05 (Kleinboelting et al., 2012; De Schutter et al., 2007), and crossed them to generate double nulls (genotyping primers provided in Table S1). In addition, two independent, segregating lines of *MUTEp::amiSPCH*, generated as described above, were collected for quantification. In brief, seedlings were cleared in 7:1 ethanol:acetic acid,

treated for 30 min with 1 N potassium hydroxide, rinsed in water, and mounted in Hoyer's medium. Differential contrast interference (DIC) images of the abaxial epidermis from cotyledon middle regions were obtained with a Leica DM2500 microscope (Leica Microsystems).

Single cell isolation for scRNA-seq—Single cells (i.e. protoplasts, plant cells released from their cell walls) from respective transgenic reporter lines were isolated as described previously (Adrian et al., 2015; Bargmann and Birnbaum, 2010). In brief, whole aerial tissue or first true leaves harvested from ~180–260 total 10 dpg seedlings were placed into 15 mL protoplasting solution: pH 5.7 with 1 M Tris-Cl pH 7.5, contains 1.25% [w/v] cellulase R-10 (Yakult), 0.3% [w/v] macerozyme R-10 (Yakult), 0.4 M mannitol (Sigma M1902), 20 mM MES (Caisson Labs M009), 20 mM KCl, 0.1% BSA (Fisher BP1605), and 10 mM $CaCl_2$. After 2 hours of gentle shaking, the protoplast solution was filtered through a 40 μm cell strainer (BD Falcon 352340) and centrifuged for 5 min at 500 x g. Protoplast pellets were resuspended in ~1–3 mL of protoplasting solution and sorted using a FACSAria II (BD Biosciences) instrument fitted with a 100 μm nozzle.

Droplet-based 10X Genomics—300,000 single-cell positive events were sorted into ~1–2mL of protoplasting solution. Kept at room temperature, cells were lightly centrifuged and resuspended in protoplasting solution to yield a sample with ~2,000 cells/μL in a total volume of ~20–30 μL. As our focus was on the stomatal lineage and our aim was to obtain overlapping and complementary datasets (Table 1), we generated one technical replicate of ~5,000 cells with the *ATML1p::YFP-RCI2A* genotype, which would be corroborated with a corresponding Smart-seq2 dataset described below. We compiled a larger *TMMp::TMM-YFP* dataset generated from technical triplicates, of three independent pools of tissue, to yield an estimate of ~10–15,000 total cells. Cell concentration was determined with a hemocytometer, and limited cell lysis was detected by staining with 1:1 trypan blue (Thermo Scientific 12250061). Cells were loaded into a microfluidic device (10X Genomics) with 3' v2 or v3 chemistry to capture ~5,000 cells per sample (Zheng et al., 2017). mRNA was reverse transcribed and cDNA quality was assessed using High Sensitivity NGS Fragment Analyzer (Agilent). Illumina libraries were constructed with Gene Expression v2 and v3 kits (10X Genomics), and sequencing of paired-end 75 bp reads was performed on a HiSeq4000 (Illumina), according to the manufacturer's instructions. Illumina BCL files were processed via Cell Ranger v2.1.1 and v3.1.0 pipelines, with reads aligned to the Arabidopsis TAIR10 genome assembly (Lamesch et al., 2012). Subsequently, spliced and unspliced transcript matrices were generated using samtools (Li et al., 2009) and velocyto (La Manno et al., 2018).

Plate-based Smart-seq2—Representing technical replicates (two independent pools of tissues for each *ATML1p::YFP-RCI2A* and *TMMp::TMM-YFP* genotypes, Table 1), single-cell positive events were indexed and sorted directly into single wells of a 96-well plate, which contained 4 μL of lysis buffer with 1 U/μL of Recombinant RNase Inhibitor (RRI, Clontech 2313B), 0.1% [w/v] Triton X-100 (Thermo Scientific 85111), 2.5 mM dNTP (Thermo Scientific 10297018), and 2.5 μM oligodT30VN (5′AAGCAGTGGTATCAACGCAGAGTACT30VN-3′, IDT). Once sorted, cells were

immediately spun down and frozen at −80°C for temporary storage. cDNA synthesis was performed as previously described using the Smart-seq2 protocol (Picelli et al., 2014), with SMARTScribe (Clontech 639538) reverse transcriptase and KAPA HiFi DNA polymerase (Kapa Biosystems KK2602). cDNA quality was assessed using High Sensitivity NGS Fragment Analyzer Kit (Advanced Analytical DNF-474), and barcoded Illumina libraries were made using the miniaturized Nextera XT protocol (Mora-Castilla et al., 2016). Sequencing of single-end 75 bp reads was performed on a NextSeq500 (Illumina) high output flow cell, according to the manufacturer's instructions. FASTQ files were generated from Illumina BCL files and low-quality sequences were trimmed by cutadapt (Marcel, 2011). Reads were aligned to the TAIR10 genome assembly (Lamesch et al., 2012) with STAR (Dobin et al., 2012), and a gene-by-cell raw count matrix was generated by HTSeq (Anders et al., 2015).

QUANTIFICATION AND STATISTICAL ANALYSIS

scRNA-seq dataset analyses—Transcripts Per kilobase Million (TPM) values (Smart-seq2) or gene-barcode matrices (10X Genomics) from single cells were analyzed with Seurat v3 (Butler et al., 2018). Data processing and plotting was done with R/RStudio (R Core Team, 2019; RStudio Team, 2019), except for as noted otherwise. Cells were pre-processed to filter out poor-quality cells with relatively few detected transcripts (i.e. less than 100–500 UMIs/cell), as well as potential cell doublets (i.e. two cells with a shared cell barcode identifier) with relatively high numbers of detected genes. Cells with substantial detected mitochondria and chloroplast transcripts (i.e. more than 25% of UMIs attributed to respective transcripts) were also filtered out, under the assumption that such profiles may represent plastids released from cells.

After pre-processing, the data were log-normalized by NormalizeData and variable genes were identified with FindVariableGenes, with parameters selection.method = "vst" and nfeatures = 2000. Batch effects were not detected within the technical replicates. Nonetheless, for the *TMMp::TMM-YFP* dataset, the technical triplicates from 10X Genomics were integrated before subsequent integration with the Smart-seq2 dataset, using FindIntegrationAnchors and IntegrateData. Unwanted sources of variation derived from cell cycle stage across different tissues were regressed out of the *ATML1p::YFP-RCI2A* dataset to avoid clustering based on cell cycle (Tirosh et al., 2016). Cell cycle stage genes were defined by known homology and previous time-course experiments (Menges et al., 2003, 2005; Vandepoele et al., 2002). Representative genes induced during S and G2/M phases (S phase: *AT1G02970/WEE1, AT3G24810/ICK3, AT5G22220/E2F1, AT5G43080/CYCA3;1, AT1G47210/CYCA3;2, AT5G65420/CYCD4;1*; G2/M phase: *AT1G76540/CDKB2;1, AT1G20930/CDKB2;2, AT1G44110/CYCA1;1, AT5G25380/CYCA2;1, AT5G11300/CYCA2;2, AT5G06150/CYCB1;2, AT3G11520/CYCB1;3, AT2G26760/CYCB1;4, AT2G17620/CYCB2;1, AT4G35620/CYCB2;2, AT1G76310/CYCB2;4, AT1G16330/CYCB3;1*) were used to assign 'cell cycle scores' with CellCycleScoring for downstream cell cycle regression. The data were then scaled with ScaleData. Principle component analysis (PCA) dimensionality reduction was performed with RunPCA to calculate 50 principal components. A Shared Nearest Neighbor (SNN) graph was generated by FindNeighbors, and graph-based cell clustering based on the SNN was performed by

FindClusters. The clustree package (Zappia and Oshlack, 2018) was implemented to visualize how clusters breakdown from one resolution to another. Uniform Manifold Approximation and Projection (UMAP) embedding was performed by RunUMAP, using all 50 principal components with parameters n_neighbors = 30, min_dist = 0.3, umap.method = “uwot”, and metric = “cosine”. Data visualization was performed with ggplot2 (Wickham, 2016).

Differentially expressed genes in each cluster were identified by the non-parametric Wilcoxon rank sum test (default) using Seurat v3 (Butler et al., 2018) FindAllMarkers, with parameters only.pos = “TRUE”, min.pct = 0.25 and logfc.threshold = 0.223 (1.25-fold change). The min.pct argument requires that genes are detected within a minimum percentage (i.e. 25%) of cells of a given cluster, and the logfc.threshold argument requires a minimum fold change (i.e. 1.25-fold) for a gene to be identified as differentially expressed. Enriched GO terms associated with the top 100 differentially expressed genes were identified by clusterProfiler (Yu et al., 2012) compareCluster and gofilter (level = 3) with the following parameters fun = “enrichGO”, OrgDb = “org.At.tair.db”, ont = “BP”, pAdjustMethod = “BH”, pvalueCutoff=0.05, qvalueCutoff=0.05.

Transcriptional dynamics (i.e. predicting future cell states by leveraging unspliced and spliced transcripts) in the *ATML1p::YFP-RCI2A* 10X Genomics leaf dataset was assessed by a steady-state deterministic model using scVelo (Bergen et al., 2020), with scv.pp.filter_and_normalize parameters min_shared_counts = 100 and min_counts_u = 500, as the program can predict multiple diverging lineages. We were able to detect reads of unspliced transcripts from secondary priming events during 10X Genomics library construction (5% of total reads: ~3,500 unspliced reads/cell with ~70,000 total reads per cell). scVelo was implemented with python using Jupyter/IPython Notebook (Perez and Granger, 2007; Van Rossum and Drake, 2009). To focus on the stomatal lineage and employ a more conservative approach, pseudotime lineage inference was performed with the integrated *TMMp::TMM-YFP* datasets in R/RStudio using Slingshot (Street et al., 2018), which constructs a minimum spanning tree and fits simultaneous principle curves through the tree, with the parameter reduceDim = “UMAP”.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We are grateful to members of the Bergmann lab and the local scientific community for helpful guidance, discussions, and feedback. We acknowledge that this research was performed on the traditional land of the Muwekma Ohlone Tribe. Cell sorting/flow cytometry analysis was done at the Stanford Shared FACS Facility. The Stanford Functional Genomics Facility was supported by NIH Instrumentation Grants (S10OD018220 and 1S10OD021763). The Stanford Genomics Center provided computational power and assistance. C.L.A. was supported by an NIH-NIGMS Ruth L. Kirschstein-NRSA postdoctoral fellowship (F32GM129918), A.V. was supported by an EMBO postdoctoral fellowship (ALTF 707–2012), and A.K.W. was supported by a postdoctoral fellowship from Deutsche Forschungsgemeinschaft. D.C.B. is an HHMI investigator.

REFERENCES

Adrian J, Chang J, Ballenger CE, Bargmann BOR, Alassimone J, Davies KA, Lau OS, Matos JL, Hachez C, Lanctot A, et al. (2015). Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. Dev. Cell 33, 107–118.

Albert NW (2015). Subspecialization of R2R3-MYB Repressors for Anthocyanin and Proanthocyanidin Regulation in Forage Legumes. Front. Plant Sci 6, 1165.

Anders S, Pyl PT, and Huber W (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169.

Bargmann BOR, and Birnbaum KD (2010). Fluorescence activated cell sorting of plant protoplasts. J. Vis. Exp

Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, and Newell EW (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol

Bergen V, Lange M, Peidli S, Wolf FA, and Theis FJ (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol

Bonke M, Thitamadee S, Mähönen AP, Hauser M-T, and Helariutta Y (2003). APL regulates vascular tissue identity in Arabidopsis. Nature 426, 181–186.

Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol 36, 411–420.

Carlson M (2019). org.At.tair.db: Genome wide annotation for Arabidopsis

Chen L-Q, Qu X-Q, Hou B-H, Sosso D, Osorio S, Fernie AR, and Frommer WB (2012). Sucrose Efflux Mediated by SWEET Proteins as a Key Step for Phloem Transport. Science (80-.) 335, 207 LP–211.

Clough SJ (2005). Floral dip: agrobacterium-mediated germ line transformation. Methods Mol. Biol 286, 91–102.

Clough SJ, and Bent AF (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16, 735–743.

Cruz-Ramírez A, Díaz-Triviño S, Blilou I, Grieneisen VA, Sozzani R, Zamioudis C, Miskolczi P, Nieuwland J, Benjamins R, Dhonukshe P, et al. (2012). A bistable circuit involving SCARECROW-RETINOBLASTOMA integrates cues to inform asymmetric stem cell division. Cell 150, 1002–1015.

Davies KA, and Bergmann DC (2014). Functional specialization of stomatal bHLHs through modification of DNA-binding and phosphoregulation potential. Proc. Natl. Acad. Sci. U. S. A 111, 15585–15590.

Dhondt S, Coppens F, De Winter F, Swarup K, Merks RMH, Inzé D, Bennett MJ, and Beemster GTS (2010). SHORT-ROOT and SCARECROW Regulate Leaf Growth in Arabidopsis by Stimulating S-Phase Progression of the Cell Cycle. Plant Physiol 154, 1183 LP–1195.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2012). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, and Huber W (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics 21, 3439–3440.

Durinck S, Spellman PT, Birney E, and Huber W (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc 4, 1184–1191.

Fisher K, and Turner S (2007). PXY, a receptor-like kinase essential for maintaining polarity during plant vascular-tissue development. Curr. Biol 17, 1061–1066.

Furutani M, Nakano Y, and Tasaka M (2014). MAB4-induced auxin sink generates local auxin gradients in Arabidopsis organ formation. Proc. Natl. Acad. Sci 111, 1198 LP–1203.

von Groll U, Berger D, and Altmann T (2002). The Subtilisin-Like Serine Protease SDD1 Mediates Cell-to-Cell Signaling during Arabidopsis Stomatal Development. Plant Cell 14, 1527 LP–1539.

Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, Ilagan F, Kuo AH, Hsieh RW, Cai S, et al. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. Science (80-.) 367, 405 LP–411.

Gutierrez C (2016). 25 Years of Cell Cycle Research: What's Ahead? Trends Plant Sci 21, 823–833.

Han C, Liu Y, Shi W, Qiao Y, Wang L, Tian Y, Fan M, Deng Z, Lau OS, De Jaeger G, et al. (2020). KIN10 promotes stomatal development through stabilization of the SPEECHLESS transcription factor. Nat. Commun 11, 4214.

Han S-K, Qi X, Sugihara K, Dang JH, Endo TA, Miller KL, Kim E-D, Miura T, and Torii KU (2018). MUTE Directly Orchestrates Cell-State Switch and the Single Symmetric Division to Create Stomata. Dev. Cell 45, 303–315.e5.

Hara K, Kajita R, Torii KU, Bergmann DC, and Kakimoto T (2007). The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. Genes Dev 21, 1720–1725.

Hara K, Yokoo T, Kajita R, Onishi T, Yahata S, Peterson KM, Torii KU, and Kakimoto T (2009). Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in Arabidopsis leaves. Plant Cell Physiol 50, 1019–1031.

Hellmann E, Ko D, Ruonala R, and Helariutta Y (2018). Plant Vascular Tissues-Connecting Tissue Comes in All Shapes. Plants (Basel, Switzerland) 7, 109.

Hetherington AM, and Woodward FI (2003). The role of stomata in sensing and driving environmental change. Nature 424, 901–908.

Hunt L, Bailey KJ, and Gray JE (2010). The signalling peptide EPFL9 is a positive regulator of stomatal development. New Phytol 186, 609–614.

Kanaoka MM, Pillitteri LJ, Fujii H, Yoshida Y, Bogenschutz NL, Takabayashi J, Zhu J-K, and Torii KU (2008). SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to arabidopsis stomatal differentiation. Plant Cell 20, 1775–1785.

Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C, Kocks C, Rajewsky N, and Zinzen RP (2017). The Drosophila embryo at single-cell transcriptome resolution. Science (80-.) 358, 194 LP–199.

Kim J-Y, Symeonidi E, Pang TY, Denyer T, Weidauer D, Bezrutczyk M, Miras M, Zöllner N, Hartwig T, Wudick MM, et al. (2021). Distinct identities of leaf phloem cells revealed by single cell transcriptomics. Plant Cell

Kleinboelting N, Huep G, Kloetgen A, Viehoever P, and Weisshaar B (2012). GABI-Kat SimpleSearch: new features of the Arabidopsis thaliana T-DNA mutant database. Nucleic Acids Res 40, D1211–5.

Kondo T, Kajita R, Miyazaki A, Hokoyama M, Nakamura-Miura T, Mizuno S, Masuda Y, Irie K, Tanaka Y, Takada S, et al. (2010). Stomatal density is controlled by a mesophyll-derived signaling molecule. Plant Cell Physiol 51, 1–8.

Kubo M, Udagawa M, Nishikubo N, Horiguchi G, Yamaguchi M, Ito J, Mimura T, Fukuda H, and Demura T (2005). Transcription switches for protoxylem and metaxylem vessel formation. Genes Dev 19, 1855–1860.

Kumari A, Jewaria PK, Bergmann DC, and Kakimoto T (2014). Arabidopsis reduces growth under osmotic stress by decreasing SPEECHLESS protein. Plant Cell Physiol 55, 2037–2046.

Lai LB, Nadeau JA, Lucas J, Lee E-K, Nakagawa T, Zhao L, Geisler M, and Sack FD (2005). The Arabidopsis R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. Plant Cell 17, 2754–2767.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res 40, D1202–D1210.

Lampard GR, MacAlister CA, and Bergmann DC (2008). Arabidopsis Stomatal Initiation Is Controlled by MAPK-Mediated Regulation of the bHLH SPEECHLESS. Science (80-.) 322, 1113 LP–1116.

Lau OS, Davies KA, Chang J, Adrian J, Rowe MH, Ballenger CE, and Bergmann DC (2014). Direct roles of SPEECHLESS in the specification of stomatal self-renewing cells. Science 345, 1605–1609.

Lauber MH, Waizenegger I, Steinmann T, Schwarz H, Mayer U, Hwang I, Lukowitz W, and Jürgens G (1997). The Arabidopsis KNOLLE protein is a cytokinesis-specific syntaxin. J. Cell Biol 139, 1485–1493.

Lee LR, and Bergmann DC (2019). The plant stomatal lineage at a glance. J. Cell Sci 132.

Lee LR, Wengier DL, and Bergmann DC (2019). Cell-type-specific transcriptome and histone modification dynamics during cellular reprogramming in the Arabidopsis stomatal lineage. Proc. Natl. Acad. Sci. U. S. A 116, 21914–21924.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, and Durbin R (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Liu Z, Zhou Y, Guo J, Li J, Tian Z, Zhu Z, Wang J, Wu R, Zhang B, Hu Y, et al. (2020). Global Dynamic Molecular Profiling of Stomatal Lineage Cell Development by Single-Cell RNA Sequencing. Mol. Plant 13, 1178–1193.

Lukowitz W, Mayer U, and Jürgens G (1996). Cy tokinesis in the Arabidopsis embryo involves the syntaxin-related KNOLLE gene product. Cell 84, 61–71.

Ma Y, Kanakousaki K, and Buttitta L (2015). How the cell cycle impacts chromatin architecture and influences cell fate. Front. Genet 6, 19.

MacAlister CA, Ohashi-Ito K, and Bergmann DC (2007). Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. Nature 445, 537–540.

La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, et al. (2018). RNA velo city of single cells. Nature 560, 494–498.

Marcel M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J 17, 10–12.

Matos JL, Lau OS, Hachez C, Cruz-Ramírez A, Scheres B, and Bergmann DC (2014). Irreversible fate commitment in the Arabidopsis stomatal lineage requires a FAMA and RETINOBLASTOMA-RELATED module. Elife 3, 1–15.

McInnes L, Healy J, Saul N, and Grossberger L (2018). UMAP: Uniform Manifold Approximation and Projection. J. Open Source Softw 3, 861.

Menges M, Hennig L, Gruissem W, and Murray JAH (2003). Genome-wide gene expression in an Arabidopsis cell suspension. Plant Mol. Biol 53, 423–442.

Menges M, de Jager SM, Gruissem W, and Murray JAH (2005). Global analysis of the core cell cycle regulators of Arabidopsis identifies novel genes, reveals multiple and highly specific profiles of expression and provides a coherent model for plant cell cycle control. Plant J 41, 546–566.

Mora-Castilla S, To C, Vaezeslami S, Morey R, Srinivasan S, Dumdie JN, Cook-Andersen H, Jenkins J, and Laurent LC (2016). Miniaturization Technologies for Efficient Single-Cell Library Preparation for Next-Generation Sequencing. J. Lab. Autom 21, 557–567.

Morris SA (2019). The evolving concept of cell identity in the single cell era. Development 146, dev169748.

Nadeau JA, and Sack FD (2002). Control of stomatal distribution on the Arabidopsis leaf surface. Science 296, 1697–1700.

Nakagawa T, Nakamura S, Tanaka K, Kawamukai M, Suzuki T, Nakamura K, Kimura T, and Ishiguro S (2008). Development of R4 gateway binary vectors (R4pGWB) enabling high-throughput promoter swapping for plant research. Biosci. Biotechnol. Biochem 72, 624–629.

Negi J, Moriwaki K, Konishi M, Yokoyama R, Nakano T, Kusumi K, Hashimoto-Sugimoto M, Schroeder JI, Nishitani K, Yanagisawa S, et al. (2013). A Dof transcription factor, SCAP1, is essential for the development of functional stomata in Arabidopsis. Curr. Biol 23, 479–484.

Nelms B, and Walbot V (2019). Defining the developmental program leading to meiosis in maize. Science 364, 52–56.

Nurse P, and Thuriaux P (1980). Regulatory genes controlling mitosis in the fission yeast Schizosaccharomyces pombe. Genetics 96, 627–637.

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

Perez F, and Granger BE (2007). IPython: A System for Interactive Scientific Computing. Comput. Sci. Engg 9, 21–29.

Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, and Sandberg R (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc 9, 171–181.

Pillitteri LJ, Sloan DB, Bogenschutz NL, and Torii KU (2007). Termination of asymmetric cell division and differentiation of stomata. Nature 445, 501–505.

Putarjunan A, Ruble J, Srivastava A, Zhao C, Rychel AL, Hofstetter AK, Tang X, Zhu J-K, Tama F, Zheng N, et al. (2019). Bipartite anchoring of SCREAM enforces stomatal initiation by coupling MAP kinases to SPEECHLESS. Nat. Plants 5, 742–754.

R Core Team (2019). R: A Language and Environment for Statistical Computing

Roeder AHK, Chickarmane V, Cunha A, Obara B, Manjunath BS, and Meyerowitz EM (2010). Variability in the control of cell division underlies sepal epidermal patterning in Arabidopsis thaliana. PLoS Biol 8, e1000367.

Roppolo D, Boeckmann B, Pfister A, Boutet E, Rubio MC, Dénervaud-Tendon V, Vermeer JEM, Gheyselinck J, Xenarios I, and Geldner N (2014). Functional and Evolutionary Analysis of the CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN Family. Plant Physiol 165, 1709 LP–1722.

Van Rossum G, and Drake FL (2009). Python 3 Reference Manual (Scotts Valley, CA: CreateSpace).

RStudio Team (2019). RStudio: Integrated Development for R

De Rybel B, Möller B, Yoshida S, Grabowicz I, Barbier de Reuille P, Boeren S, Smith RS, Borst JW, and Weijers D (2013). A bHLH complex controls embryonic vascular tissue establishment and indeterminate growth in Arabidopsis. Dev. Cell 24, 426–437.

Ryu KH, Huang L, Kang HM, and Schiefelbein J (2019). Single-Cell RNA Sequencing Resolves Molecular Relationships Among Individual Plant Cells. Plant Physiol 179, 1444–1456.

Sawchuk MG, Donner TJ, Head P, and Scarpella E (2008). Unique and Overlapping Expression Patterns among Members of Photosynthesis-Associated Nuclear Gene Families in Arabidopsis. Plant Physiol 148, 1908–1924.

Scarpella E, Francis P, and Berleth T (2004). Stage-specific markers define early steps of procambium development in Arabidopsis leaves and correlate termination of vein formation with mesophyll differentiation. Development 131, 3445–3455.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682.

Schlereth A, Möller B, Liu W, Kientz M, Flipse J, Rademacher EH, Schmid M, Jürgens G, and Weijers D (2010). MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. Nature 464, 913–916.

De Schutter K, Joubès J, Cools T, Verkest A, Corellou F, Babiychuk E, Van Der Schueren E, Beeckman T, Kushnir S, Inzé D, et al. (2007). Arabidopsis WEE1 kinase controls cell cycle arrest in response to activation of the DNA integrity checkpoint. Plant Cell 19, 211–225.

Schwab R, Ossowski S, Riester M, Warthmann N, and Weigel D (2006). Highly Specific Gene Silencing by Artificial MicroRNAs in Arabidopsis. Plant Cell 18, 1121 LP–1133.

Shulse CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, Gouran M, Turco GM, Zhu Y, O'Malley RC, Brady SM, et al. (2019). High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. Cell Rep 27, 2241–2247.e4.

Sozzani R, Cui H, Moreno-Risueno MA, Busch W, Van Norman JM, Vernoux T, Brady SM, Dewitte W, Murray JAH, and Benfey PN (2010). Spatiotemporal regulation of cell-cycle genes by SHORTROOT links patterning and growth. Nature 466, 128–132.

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

Sugano SS, Shimada T, Imai Y, Okawa K, Tamai A, Mori M, and Hara-Nishimura I (2010). Stomagen positively regulates stomatal density in Arabidopsis. Nature 463, 241–244.

Takahashi I, Kojima S, Sakaguchi N, Umeda-Hara C, and Umeda M (2010). Two Arabidopsis cyclin A3s possess G1 cyclin-like features. Plant Cell Rep 29, 307–315.

Tanaka Y, Nakamura S, Kawamukai M, Koizumi N, and Nakagawa T (2011). Development of a series of gateway binary vectors possessing a tunicamycin resistance gene as a marker for the transformation of Arabidopsis thaliana. Biosci. Biotechnol. Biochem 75, 804–807.

Tian C, Wang Y, Yu H, He J, Wang J, Shi B, Du Q, Provart NJ, Meyerowitz EM, and Jiao Y (2019). A gene expression map of shoot domains reveals regulatory mechanisms. Nat. Commun 10, 141.

Tintori SC, Osborne Nishimura E, Golden P, Lieb JD, and Goldstein B (2016). A Transcriptional Lineage of the Early C. elegans Embryo. Dev. Cell 38, 430–444.

Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science (80-.) 352, 189 LP–196.

Triviño M, Martín-Trillo M, Ballesteros I, Delgado D, de Marcos A, Desvoyes B, Gutiérrez C, Mena M, and Fenoll C (2013). Timely expression of the Arabidopsis stoma-fate master regulator MUTE is required for specification of other epidermal cell types. Plant J 75, 808–822.

Vandepoele K, Raes J, De Veylder L, Rouzé P, Rombauts S, and Inzé D (2002). Genome-wide analysis of core cell cycle genes in Arabidopsis. Plant Cell 14, 903–916.

Vatén A, Soyars CL, Tarr PT, Nimchuk ZL, and Bergmann DC (2018). Modulation of Asymmetric Division Diversity through Cytokinin and SPEECHLESS Regulatory Interactions in the Arabidopsis Stomatal Lineage. Dev. Cell 47, 53–66.e5.

Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, and Klein AM (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science (80-.) 360, 981 LP–987.

Wallner E-S, López-Salmerón V, Belevich I, Poschet G, Jung I, Grünwald K, Sevilem I, Jokitalo E, Hell R, Helariutta Y, et al. (2017). Strigolactone- and Karrikin-Independent SMXL Proteins Are Central Regulators of Phloem Formation. Curr. Biol 27, 1241–1247.

Weimer AK, Matos JL, Sharma N, Patell F, Murray JAH, Dewitte W, and Bergmann DC (2018). Lineage-and stage-specific expressed CYCD7;1 coordinates the single symmetric division that creates stomatal guard cells. Development 145, dev160671.

Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).

Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu Z, Wang L, Fox N, et al. (2021). Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. Dev. Cell

Yadav RK, Tavakkoli M, Xie M, Girke T, and Reddy GV (2014). A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche. Development 141, 2735–2744.

Yu G, Wang L-G, Han Y, and He Q-Y (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16, 284–287.

Zappia L, and Oshlack A (2018). Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience 7.

Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun 8, 14049.

Highlights

- Distinct models of cell differentiation and lineage trajectories define leaf tissue
- Cell cycle regulators exhibit distinct, yet overlapping expression profiles
- Flexible stomatal lineage cell states exist along a continuum
- Single-cell resolution refines cell fate commitment decisions that yield stomata

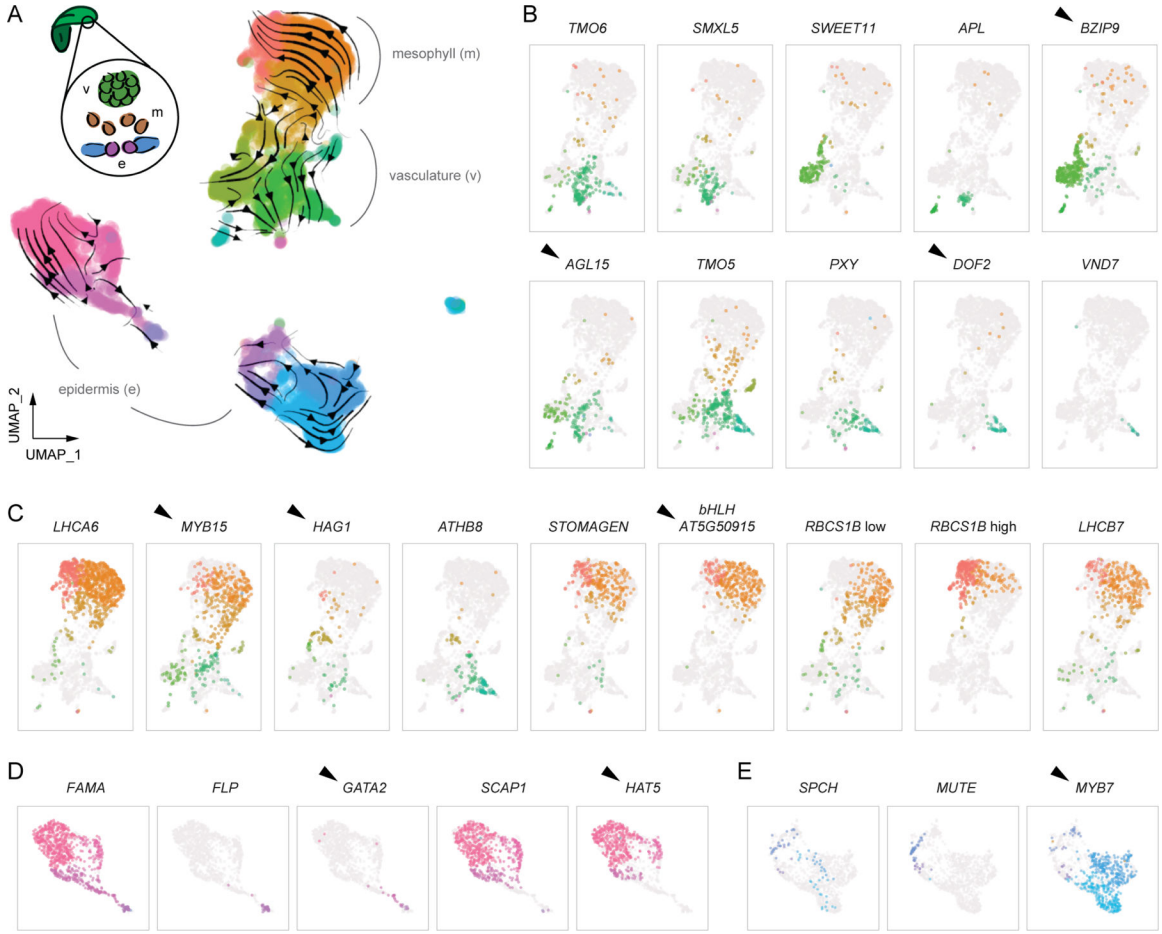


Figure 1. Models of cell state differentiation within leaf tissue.

(A) Lineage inference models by scVelo layered onto the developing leaf scRNA-seq UMAP (10X Genomics v2, 5,021 cells displayed) from 10 days post-germination (dpg) *ATML1p::YFP-RCI2A* seedlings. Size of arrows depicts relative degree of gene expression changes within predicted, directed transitions from one cell state to another. Colors represent graph-based clustering with resolution 1, and grey brackets demarcate major clusters from the mesophyll, vasculature, and epidermis. (B-E) UMAP insets of the major clusters that correspond to different tissues, which include (B) vasculature, (C) mesophyll, (D) stomatal guard cells, and (E) young epidermal cells, showing expression profiles (color: resolution 1) of exemplary lineage marker genes. Two insets depicting *RBCS1B* expression distinguish 'low' versus 'high' gene expression within mesophyll cells, respectively selected from the bottom and top 25% of the expression range, which demark normalized expression values less than 1.5 or greater than 4.3. Gene names of previously uncharacterized, potential regulators are designated by arrowheads. See also Figure S1–S3, and Table S3–S4.

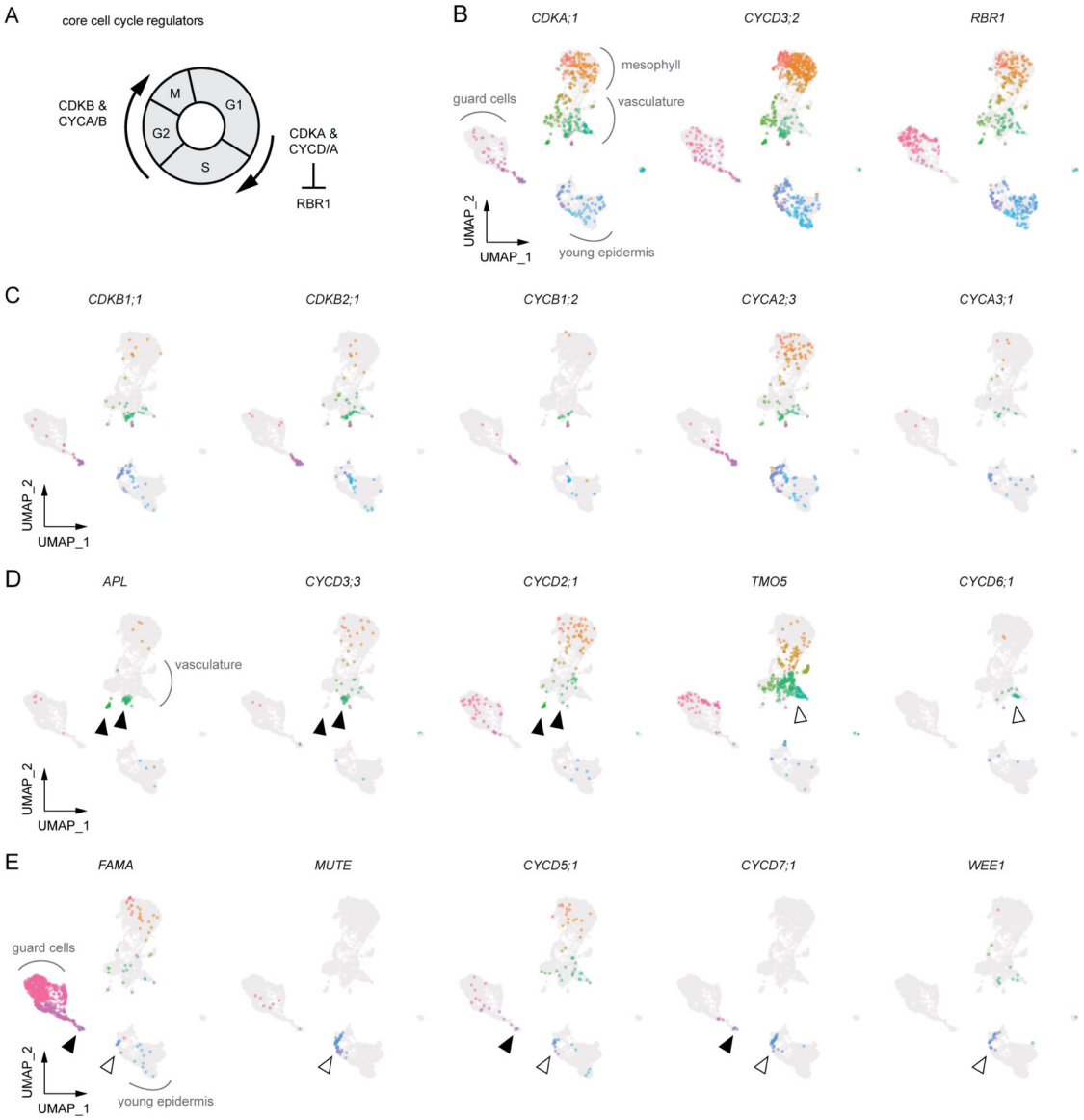


Figure 2. Cell cycle regulators exhibit distinct, yet overlapping, expression profiles.

(A) Schematic of the cell cycle. (B-E) UMAPs depict gene expression profiles of cell cycle regulators within the developing leaf. Colors represent graph-based clustering with resolution 1. (B) Core regulators of G1/S phase. (C) *CDKB/CYCB* markers of G2/M phase and *CYCAs*. While *CYCA2;3* expression is relatively broad, *CYCA3;1* is likely expressed only in G1/S phase. (D) Co-expression of *CYCD* regulators with the phloem transcriptional regulator *APL* (black arrowheads) and the procambium/xylem transcriptional regulator *TMO5* (white arrowhead). (E) Co-expression of *CYCDs* and *WEE1* with the epidermal transcription factors *FAMA* (black arrowhead) and *MUTE* (white arrowhead). See also Figure S1 and Table S3.

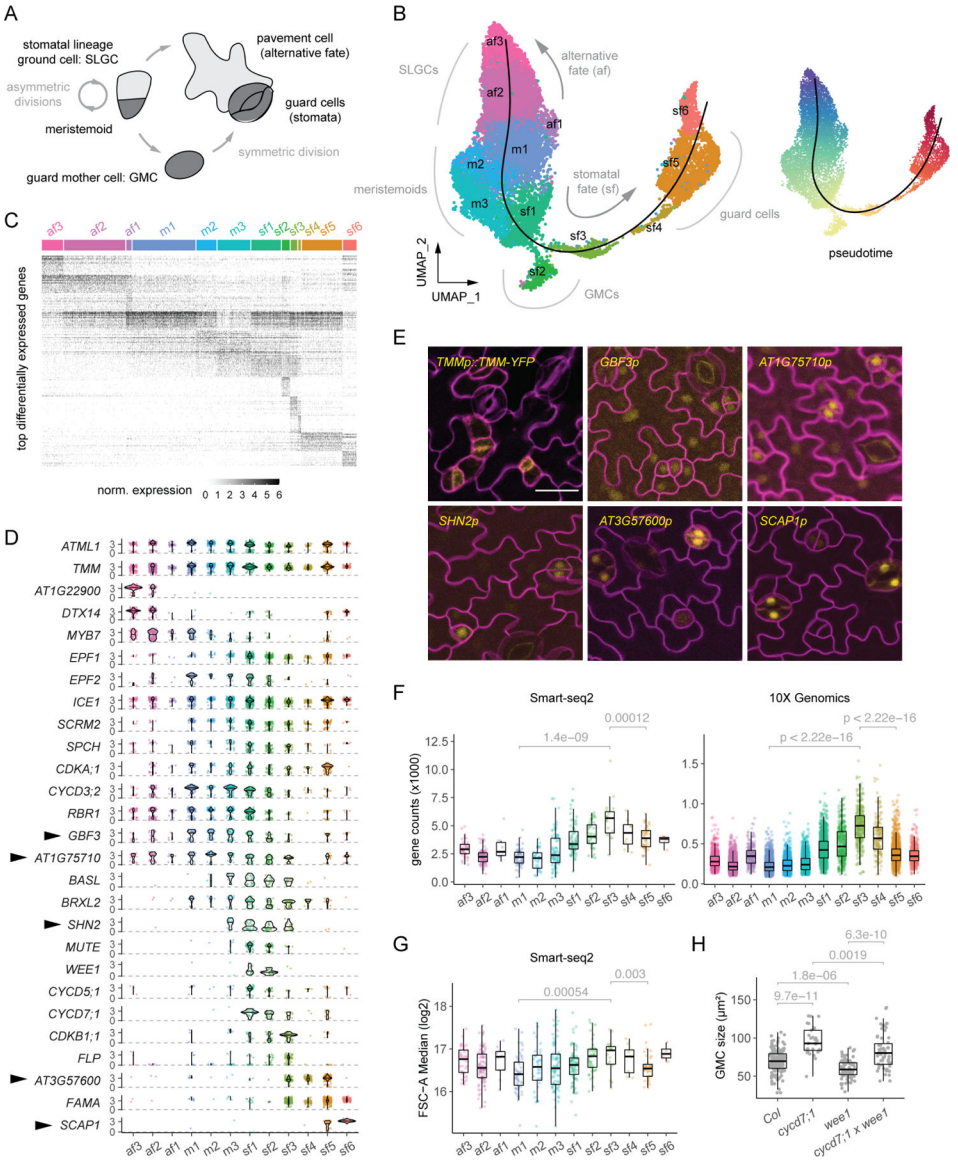


Figure 3. Stomatal lineage flexibility defined as a continuum of cell states.

(A) Model of stomatal lineage. Early stem cells – meristemoids and stomatal lineage ground cells (SLGCs) – either divide asymmetrically (left) or commit to differentiation (right). Differentiation ultimately yields guard cells and pavement cells. (B) UMAPs of integrated scRNA-seq (10X Genomics v3, 12,933 cells; Smart-seq2, 478 cells) derived from *TMMp::TMM-YFP*-expressing shoot tissue. Colors depict graph-based clustering with resolution 0.8 (left; alternative fate: af, meristemoid: m, stomatal fate: sf) and the relative pseudotime scale (right, color transitions correspond to the pseudotime axis line). The pseudotime axis (black curved line) is overlaid on the UMAPs. (C) Heatmap of the top 20 differentially expressed genes (width of colored label corresponds to the number of cells per cluster) (see Table S5A–B) and (D) violin plots depict a range of gene expression profiles within pseudotime, aligned with cluster resolution 0.8 (colors). Arrowheads indicate genes for which reporters were generated. The pseudotime trajectory is predicted to initiate with

SPCH expression, and it represents transitions from meristemoids to SLGCs in one direction (towards left, af clusters) and transitions from meristemoids to GMCs and guard cells (towards right, sf clusters). (E) Confocal images of promoter-driven *YFP-NLS* reporters corroborate predicted expression profiles. The abaxial epidermis was imaged from first true leaves of 10 dpg seedlings. All images were taken at the same magnification (scale bar: 20 μm). (F) Boxplots depict transcriptional diversity within pseudotime, based on total gene counts from respective datasets. (G) Boxplots depict relative cell size (log2 values of the median forward scatter area, FSC-A) within pseudotime. FSC was measured and indexed by FACS, and it corresponds to the Smart-seq2 dataset. Colors in (F, G) represent cluster resolution 0.8, and clusters are aligned with pseudotime, as in (C, D). (H) Quantification of GMC area in Col wild type (n = 154), *cycd7;1* (n = 29), *wee1* (n = 72), and *cycd7;1* x *wee1* (n = 63) cotyledons at 4 dpg. P-values in (F-H) are from pairwise comparisons using the Wilcoxon rank sum test. See also Figure S4 and Table S5.

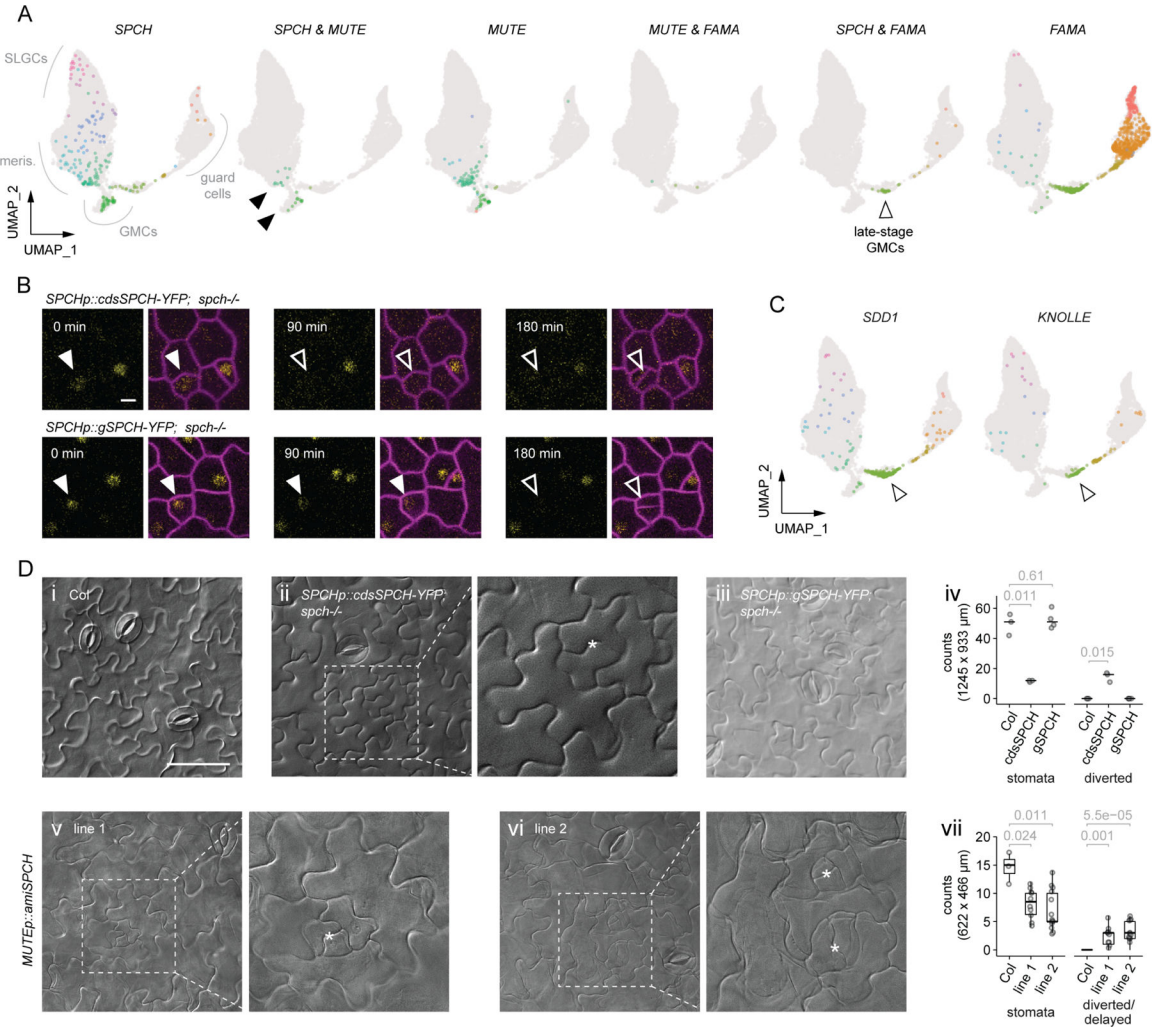


Figure 4. *SPCH* is expressed and required beyond the early stages of the lineage.

(A) Depiction of extended *SPCH* gene expression. UMAPs indicate cells and clusters (resolution 0.8) that express both *SPCH* and *MUTE* (black arrowheads, 21 cells or 8% of total cells that express *SPCH*) are distinct from the cluster that expresses *SPCH* and *FAMA* (white arrowhead, 27 cells or 10% of total cells). (B) Time-lapse confocal images of *SPCHp::cdsSPCH-YFP; spch*−/− and *SPCHp::gSPCH-YFP; spch*−/− from the abaxial epidermis of 4 dpg seedlings. Arrowheads follow the GMC division that yields young stomatal guard cells (at 180min). Filled white arrowhead: reporter expression detected, outlined arrowhead: no expression detected. All images were taken at the same magnification (scale bar: 5 µm). (C) Examples of putative state-specific targets of SPCH, and their respective UMAP expression profiles (colors: cluster resolution 0.8). 33 and 16 cells respectively co-express *SDD1* or *KNOLLE* with *SPCH*, representing 13% and 6% of cells that express *SPCH*. (D) Diverted cell fates (asterisks: cells with diverted/delayed identities) in representative differential interference contrast DIC images of abaxial cotyledon epidermis from Col wildtype control (i), *SPCHp::cdsSPCH-YFP spch*−/− (ii), *SPCHp::gSPCH-YFP spch*−/− (iii), and two independent lines of *MUTE::amiSPCH* (v, vi). Images were all taken at 15 dpg with the same magnification (scale bars: 100 µm), except for

insets demarked by white dashed boxes. Diverted/delayed fate events are quantified in (iv, vii). Early lobing of an SLGC is depicted in (vi, top asterisk). Plots in (iv, vii) include grey points that each represent counts from an individual seedling (at least n=3 seedlings were measured per genotype). median: black line. p-values: two-sample t-test for respective comparisons except for diverted events in (iv), wherein the p-value represents one-sample t-test. See also Figure S4 and Table S5.

Table 1.
Dataset attributes of scRNA-seq approaches.

Included are library preparation approaches, tissues and reporters sampled, total number of cells and mean values of reads that passed mapping preprocessing, and postprocessing median values of UMIs and genes. Technical replicates from independent pools of tissue are indicated as such (e.g. two pools of tissue were used to generate the Smart-seq2 *ATML1p::YFP-RCI2A* dataset), as are reporter replicates from datasets generated using shared genotypes *ATML1p::YFP-RCI2A* or *TMMp::TMM-YFP*, which are comprised of three and five replicates, respectively. See also Figure S2 and STAR Methods.

Method	Tissue	Reporter	Total cells	Mean # reads	Mean # UMIs	Median # genes	Technical replicates	Reporter replicates
Smart-seq2	Aerial tissue	*ATML1p::YFP-RCI2A*	42	731,385	-	2,858	pool 1	*ATML1p* rep 1
			101	778,841	-	3,305	pool 2	*ATML1p* rep 2
		TMMp::TMM-YFP	142	997,485	-	3,319	pool 1	*TMMp* rep 1
			336	1,191,501	-	2,620	pool 2	*TMMp* rep 2
10X Genomics v2	True leaves	*ATML1p::YFP-RCI2A*	5,021	67,873	5,026	1,870	pool 1	*ATML1p* rep 3
10X Genomics v3		*TMMp::TMM-YFP*	4,310	25,101	371	282	pool 1	*TMMp* rep 3
			4,099	25,673	333	257	pool 2	*TMMp* rep 4
			4,524	23,307	354	272	pool 3	*TMMp* rep 5

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Escherichia coli TOP10	N/A	N/A
Agrobacterium tumefaciens GV3101	N/A	N/A
Chemicals, Peptides, and Recombinant Proteins		
Propidium iodide	Invitrogen / Molecular Probes	Cat#P3566
Cellulase R-10	Yakult	N/A
Macerozyme R-10	Yakult	N/A
Mannitol	Sigma	M1902
Critical Commercial Assays		
pENTR 5' TOPO TA cloning kit	Invitrogen	Cat#K591-20
Gateway LR clonase II enzyme mix	Invitrogen	Cat#11791100
Chromium Single Cell 3' v2 / v3 Reagent Kits	10X Genomics	N/A
High Sensitivity NGS Fragment Analyzer Kit	Advanced Analytical	DNF-474
Deposited Data		
raw and analyzed scRNA-seq datasets	this manuscript	GSE167135
Experimental Models: Organisms/Strains		
Arabidopsis: ATML1p::YFP-RCI2A	Roeder et al., 2010	N/A
Arabidopsis: TMMp::TMM-YFP	Nadeau and Sack, 2002	N/A
Arabidopsis: SPCHp::cdsSPCH-YFP	Vatén et al., 2018	N/A
Arabidopsis: SPCHp::gSPCH-YFP	Vatén et al., 2018	N/A
Arabidopsis: AT5G54630p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: BNQ2p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: ANAC034p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: AT1G11850p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: AT1G75710p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: GBF3p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: SHN2p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: AT3G57600p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: SCAP1p::YFP-NLS ATML1p::mCherry-RCI2A	this manuscript	N/A
Arabidopsis: cycd7;1-1	Weimer et al., 2018	FLAG_369E02
Arabidopsis: wee1-1	De Schutter et al., 2007	GABI_270E05
Arabidopsis: cycd7;1-1 wee1-1	this manuscript	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Arabidopsis: MUTEp::amiSPCH	this manuscript	N/A
Oligonucleotides		
see Table S1 for the list of primers used to generate transgenics and genotype lines	N/A	N/A
Recombinant DNA		
Plasmid: R4pGWB601	Nakagawa et al., 2008; Tanaka et al., 2011	RIKEN BRC: pdi00133
Plasmid: YFP-NLS pENTR	Lee et al., 2019	N/A
Plasmid: pRS300	Schwab et al., 2006	Addgene #22846
Software and Algorithms		
Cell Ranger v2.1.1 / v3.1.0	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome
samtools	Li et al., 2009	http://www.htslib.org/
velocyto	La Manno et al., 2018	http://velocyto.org/
cutadapt	Marcel, 2011	https://github.com/marcelm/cutadapt
STAR	Dobin et al., 2012	https://github.com/alexdobin/STAR
HTSeq	Anders et al., 2015	https://github.com/htseq/htseq
Seurat v3	Butler et al., 2018	https://satijalab.org/seurat/index.html
clustree	Zappia and Oshlack, 2018	https://cran.r-project.org/web/packages/clustree/index.html
clusterProfiler	Yu et al., 2012	https://github.com/YuLab-SMU/clusterProfiler
scVelo	Bergen et al., 2020	https://github.com/theislab/scvelo
Slingshot	Street et al., 2018	https://github.com/kstreet13/slingshot
R/RStudio	R Core Team, 2019; RStudio Team, 2019	https://www.r-project.org/
ggplot2	Wickham, 2016	https://ggplot2.tidyverse.org/
biomaRt	Durinck et al., 2009, 2005	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
org.At.tair.db	Carlson et al., 2019	https://bioconductor.org/packages/release/data/annotation/html/org.At.tair.db.html
Jupyter/IPython Notebook	Perez and Granger, 2007; Van Rossum and Drake, 2009	https://ipython.org/notebook.html
Web MicroRNA Designer platform	Schwab et al., 2006	http://wmd3.weigelworld.org
Fiji	Schindelin et al., 2012	https://imagej.net/Fiji