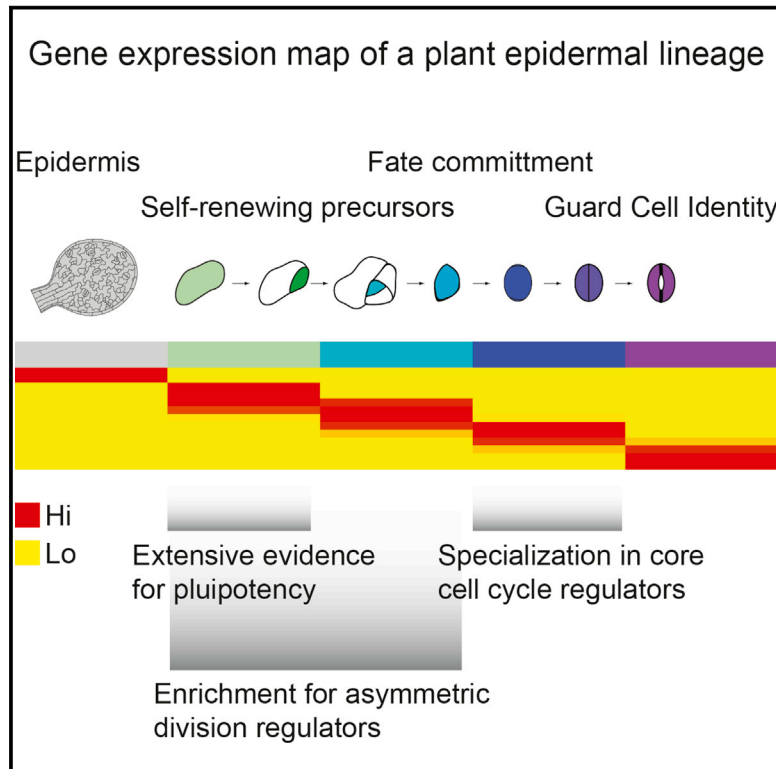# Developmental Cell

Transcriptome Dynamics of the Stomatal Lineage: Birth, Amplification, and Termination of a Self-Renewing Population

Graphical Abstract

Authors

Jessika Adrian, Jessica Chang, ..., Kenneth D. Birnbaum, Dominique C. Bergmann

Correspondence

dbergmann@stanford.edu

In Brief

Stomata facilitate plant gas exchange with the atmosphere. Adrian et al. profile the developing stomatal lineage, revealing increasing canalization of gene expression as cells become committed to specific fates and linking cell types previously thought to be independent. The data serve as a resource for further investigation of lineage specification in plants.

Highlights

- A gene expression atlas has been created for the *Arabidopsis* stomatal lineage

- Individual cell-type profiles complement profiles from shoot meristems and roots

- Profiling of meristemoids reveals unexpected pluripotency in the early lineage

- ENODLs and CYCD7 are regulators of cell division in stem cell contexts

Accession Numbers

GSE58857

CrossMark

CellPress

Transcriptome Dynamics of the Stomatal Lineage: Birth, Amplification, and Termination of a Self-Renewing Population

Jessika Adrian,[1] Jessica Chang,[1,4] Catherine E. Ballenger,[3] Bastiaan O.R. Bargmann,[2,5] Julien Alassimone,[1] Kelli A. Davies,[1] On Sun Lau,[1] Juliana L. Matos,[1] Charles Hachez,[1,6] Amy Lanctot,[1,7] Anne Vatén,[1] Kenneth D. Birnbaum,[2] and Dominique C. Bergmann[1,3,*]

[1]Department of Biology, Stanford University, Stanford, CA 94305, USA
[2]Biology Department, Center for Genomics and Systems Biology, New York University, New York, NY 10003, USA
[3]Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA
[4]Present address: Department of Genetics, Stanford Medical School, Stanford, CA 94305, USA
[5]Present address: Cibus US LLC, San Diego, CA 92121, USA
[6]Present address: Institut des Sciences de la Vie, Université Catholique de Louvain, Louvain-la-Neuve 1348, Belgium
[7]Present address: Department of Biology, University of Washington, Seattle, WA 98195, USA
*Correspondence: dbergmann@stanford.edu
http://dx.doi.org/10.1016/j.devcel.2015.01.025

SUMMARY

Developmental transitions can be described in terms of morphology and the roles of individual genes, but also in terms of global transcriptional and epigenetic changes. Temporal dissections of transcriptome changes, however, are rare for intact, developing tissues. We used RNA sequencing and microarray platforms to quantify gene expression from labeled cells isolated by fluorescence-activated cell sorting to generate cell-type-specific transcriptomes during development of an adult stem-cell lineage in the *Arabidopsis* leaf. We show that regulatory modules in this early lineage link cell types that had previously been considered to be under separate control and provide evidence for recruitment of individual members of gene families for different developmental decisions. Because stomata are physiologically important and because stomatal lineage cells exhibit exemplary division, cell fate, and cell signaling behaviors, this dataset serves as a valuable resource for further investigations of fundamental developmental processes.
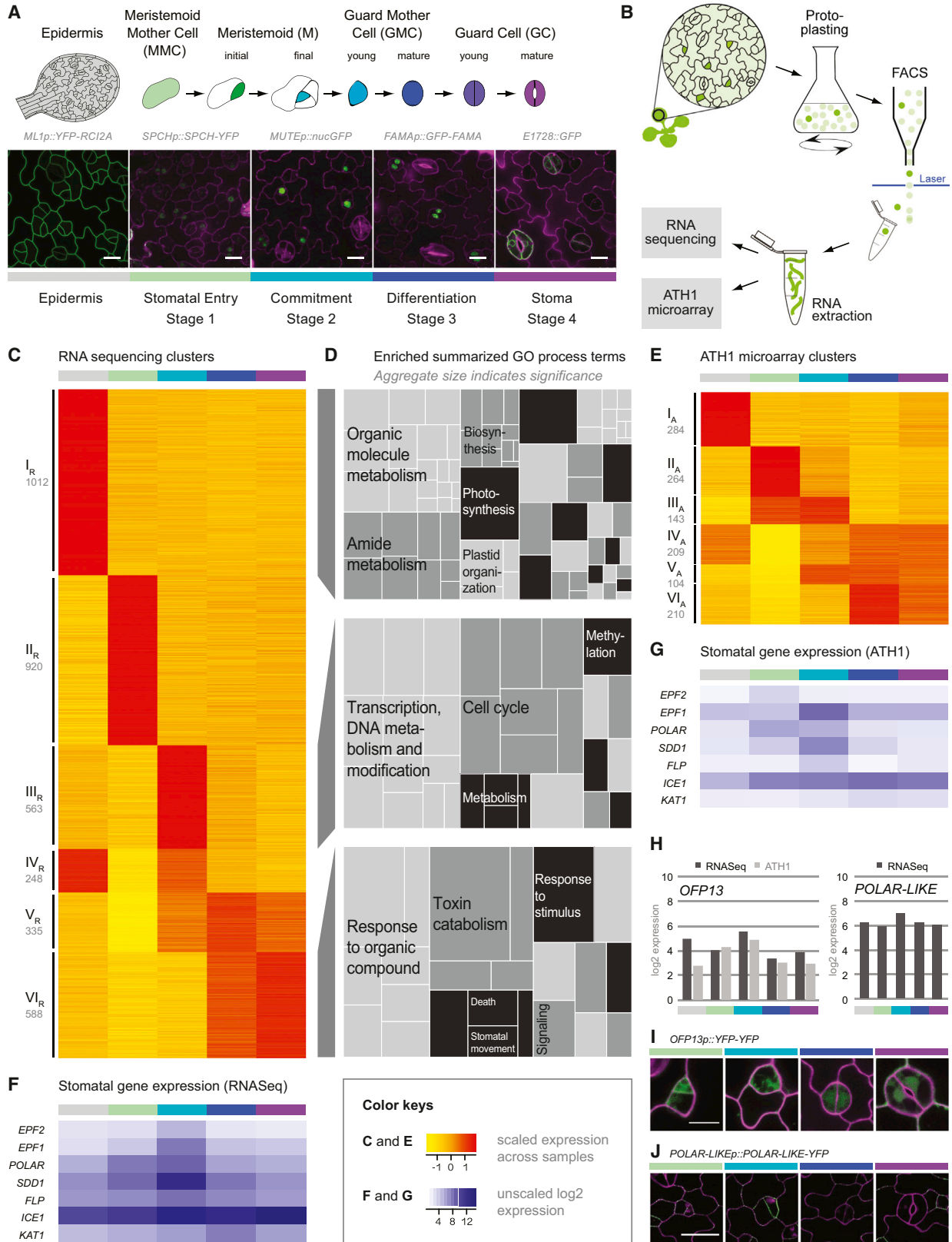
INTRODUCTION

Multicellular organisms are comprised of diverse cell types that exhibit unique transcriptional profiles appropriate to their identity and function. The development of these cell types from a common precursor requires a profound set of changes in gene expression. Recent studies following the programming and reprogramming of embryonic stem cells or induced pluripotent cells have revealed a complex, yet fairly ordered set of changes (Xie et al., 2013; Young, 2011). Similar dynamic transcriptional profiles in intact developing organisms, however, have been more

challenging to obtain. Profiles of individual cell types from intact plants have revolutionized the way cell fates and responses can be understood, but these profiles largely feature terminally differentiated cell types (e.g., Birnbaum et al., 2003; Deal and Henikoff, 2010; Yang et al., 2008). Computational approaches have been used to infer the developmental states of specific cells (Brady et al., 2007), but we lack profiles isolated directly from true intermediate cell types along a developmental trajectory.

The production and pattern of stomata in the *Arabidopsis* epidermis have received considerable recent attention as a model for cell fate determination, cell-cell communication, and cell polarity and provide a clear and accessible model for adult stem cell lineages (Pillitteri and Torii, 2012). The stomatal lineage can be parsed into discrete intermediate steps, and cells representing those intermediate steps can be identified by gene expression markers, making this an ideal system from which to generate transcriptional profiles tracing the intermediate identities and fate transitions during development. The stomatal lineage begins with asynchronous and indeterminate early divisions and lacks a strict prepattern, allowing for flexible development. Flexibility is key because the stomatal lineage generates the majority of cells in the leaf epidermis and has the potential to modify both numbers and cell types in response to environmental cues (Hetherington and Woodward, 2003).

Beyond its utility as a developmental model, the lineage produces, as its ultimate products, stomatal guard cells (GCs), that act as valves facilitating plant/atmosphere gas exchange. Because they are essential for plant physiology and are present on all large land plants, stomata have been the subject of studies ranging from probes of single molecules to global scale ecophysiology. As a consequence of the wide-scale interest in stomatal properties, mature GC transcriptomes, proteomes, and metabolomes have been generated and stomatal activities modeled (Misra et al., 2014; Yang et al., 2008; Zhao et al., 2008). Because of increasing interest and progress elucidating the integration of environmental cues (such as light and carbon dioxide) with endogenous circuits to control stomatal production and activity (e.g., Casson and Hetherington, 2014; Engineer

A Epidermis | Meristemoid Mother Cell (MMC) | Meristemoid (M) initial · final | Guard Mother Cell (GMC) young · mature | Guard Cell (GC) young · mature

ML1p::YFP-RCI2A · SPCHp::SPCH-YFP · MUTEp::nucGFP · FAMAp::GFP-FAMA · E1728::GFP

Epidermis | Stomatal Entry Stage 1 | Commitment Stage 2 | Differentiation Stage 3 | Stoma Stage 4

B Proto-plasting · FACS · Laser · RNA sequencing · ATH1 microarray · RNA extraction

C RNA sequencing clusters

I$_R$ 1012 · II$_R$ 920 · III$_R$ 563 · IV$_R$ 248 · V$_R$ 335 · VI$_R$ 588

D Enriched summarized GO process terms
Aggregate size indicates significance

Organic molecule metabolism · Biosynthesis · Photo-synthesis · Amide metabolism · Plastid organization

Transcription, DNA metabolism and modification · Cell cycle · Methylation · Metabolism

Response to organic compound · Toxin catabolism · Response to stimulus · Death · Stomatal movement · Signaling

E ATH1 microarray clusters

I$_A$ 284 · II$_A$ 264 · III$_A$ 143 · IV$_A$ 209 · V$_A$ 104 · VI$_A$ 210

G Stomatal gene expression (ATH1)

EPF2 · EPF1 · POLAR · SDD1 · FLP · ICE1 · KAT1

H
■ RNASeq ■ ATH1
OFP13
log2 expression

■ RNASeq
POLAR-LIKE
log2 expression

I OFP13p::YFP-YFP

J POLAR-LIKEp::POLAR-LIKE-YFP

F Stomatal gene expression (RNASeq)

EPF2 · EPF1 · POLAR · SDD1 · FLP · ICE1 · KAT1

Color keys

C and **E** scaled expression across samples
−1 0 1

F and **G** unscaled log2 expression
4 8 12

(legend on next page)

et al., 2014), transcriptional profiles of developing GCs or their precursors would be invaluable community resources.

Profiles of mutant seedlings enriched in precursor and mature stomatal lineage types have been useful to identify new stomatal regulators (Bergmann et al., 2004; Pillitteri et al., 2011), but these experiments profile heterogeneous (and mutant) tissues in plants that are physiologically impaired by lack of stomata. To generate a comprehensive view of WT development, we turned to fluorescence-activated cell sorting (FACS) of stomatal lineage cells derived from intact, developing plants, and generated cell-type-specific RNA expression profiles. Using both RNA sequencing (RNA-Seq) to obtain the most complete inventory of gene expression possible and ATH1 microarray profiling to enable comparisons between the transcriptomes of the stomatal lineage and other individual cell types, we resolved gene expression profiles during critical developmental events. We found that expression profiles of early stomatal lineage stages are distinct and more variable than those from committed or differentiating cells. Some of this behavior may be attributable to pluripotency of early lineage cells, as we uncovered evidence for shared expression and function of stomatal and trichome regulators. We have also validated expression of genes identified as differentially expressed (DE) in this developmental series and show mutant phenotypes related to the stages in which they are expressed. Because stomata are physiologically important and because stomatal lineage cells exhibit exemplary division, cell fate, and cell signaling behaviors, this dataset serves as a valuable resource for further investigations of fundamental processes in plants and in developing systems.

RESULTS

Identification and Isolation of Specific Stomatal Lineage Stages

Capturing cell-type-specific transcriptome changes during the development of dispersed self-renewing populations in leaves is a technical challenge because the stomatal lineage cell types are rare and transient (Figure 1A). The lineage is initiated when pluripotent meristemoid mother cells (MMCs) divide asym-

metrically, creating meristemoids as their smaller daughters. Meristemoids typically continue dividing asymmetrically two to three times, retaining meristemoid identity in the smaller daughter, before differentiating into oval-shaped guard mother cells (GMCs). Becoming a GMC marks a commitment to make GCs, which proceeds via symmetric division of GMC and subsequent coordinated morphological and gene expression changes in the daughters to form the functional stomatal unit. The larger daughters of MMC or meristemoid divisions may differentiate into pavement cells or, through secondary asymmetric divisions, create new meristemoids and MMCs (Pillitteri and Torii, 2012).

To isolate homogeneous cell populations corresponding to cells in these discrete intermediate stages along the stomatal development trajectory, we needed to identify tools capable of isolating cells within a short developmental window. We found that this developmental constraint precluded the use of INTACT or TRAP methods (Figures S1A and S1B), most likely because their efficacy is linked to the highly stable proteins used to isolate RNA (Deal and Henikoff, 2010; Mustroph et al., 2009). FACS, alternatively, could be used with fluorescent markers fused to proteins with degradation signals such that they were only present in discrete stages. Cells representing specific developmental stages isolated by FACS after a short protoplasting step included stomatal entry (stage 1, *SPCHp::SPCH-YFP*, *SSY*), commitment (stage 2, *MUTEp::nucGFP*, *MG*), and differentiation (stage 3, *FAMAp::GFP-FAMA*, *FGF*) stages, as well as mature stomata (stage 4, enhancer trap *E1728::GFP*, *E1728G*) and a marker of the entire aerial epidermis (*ML1p::YFP-RCI2A*, *ML1Y*). The cell-type-specific expression patterns of the five marker lines used have been extensively characterized previously (Davies and Bergmann, 2014; Gardner et al., 2009; MacAlister et al., 2007; Matos et al., 2014; Ohashi-Ito and Bergmann, 2006; Pillitteri et al., 2007; Robinson et al., 2011; Roeder et al., 2010). *SPCHp::SPCH-YFP* is expressed most brightly just prior to and after asymmetric divisions of MMCs, thus marking a potentially mixed population of precursors; a time course of this expression is provided in Figures S1D and S1E. To minimize transcriptional differences due to plant age or circadian rhythms, all cell types were sorted from same-aged aerial rosettes and commenced at the same time of day. The cell-type specificity

---

Figure 1. Transcriptional Profiling of Stomatal Lineage Cells Isolated by FACS

(A) Cartoon of stages in stomatal development with confocal images of markers used for FACS. Specific reporters used to mark cell stages are *ML1p::YFP-RCI2A*, epidermal cells (including stomatal lineage cells), gray; *SPCHp::SPCH-YFP*, stomatal entry, green; *MUTEp::nucGFP*, commitment, light blue; *FAMAp::GFP-FAMA*, differentiation, violet; and *E1728::GFP*, maturation, purple. Confocal images show cell-type-specific expression of fluorescent markers (green) in second true leaves of 14-day-old seedlings. Cell outlines are in magenta; scale bars represent 10 μm.

(B) Scheme of cell isolation protocol. Aerial seedling tissues expressing markers were protoplasted and FACS for RNA extraction. Expression profiles of sorted cells were generated using RNA-Seq and ATH1 microarrays (ATH1).

(C and E) Clustering of DE genes identified six dominant expression patterns (clusters I–VI; indices $_R$ and $_A$ for RNA-Seq and ATH1, respectively). Heat maps show expression of genes assigned to a cluster (clustering coefficient cutoff 0.6). Mean and median expression values are scaled per gene across samples; low expression is in yellow, and high expression is in red. The number of genes/cluster is indicated below the cluster name.

(D) Enriched GO process terms for clusters I_R, II_R, and VI_R (from C) summarized using REVIGO (Supek et al., 2011). Related GO terms are displayed in similar colors; aggregate size indicates significance of overrepresentation of a group of GO terms.

(F and G) Expression profiles of known stomatal genes generated from sorted cells profiles by RNA-Seq (F) and ATH1 microarray (G) are highly correlated with published *in planta* data. Heat maps show unscaled mean and median log2-transformed expression values; low expression is in white, and high expression is in blue.

(H–J) Validation of transcriptional map by reporter analysis of two genes not previously assigned to the stomatal lineage: *OVATE FAMILY PROTEIN 13* matches peak expression in stage 2 in both RNA-Seq and ATH1 profiles and *POLAR-LIKE (*not on ATH1) matches RNA-Seq expression. In (H), the y axis represents log2-transformed expression values; in (I) and (J), YFP signal is depicted in green and cell walls in magenta; scale bars represent 10 μm.

of reporter expression pattern at this common time point was confirmed via confocal microscopy (Figure 1A).

We coupled the isolation of RNA from cell types at specific developmental stages with two independent means of assessing gene expression: microarrays and next generation sequencing. The ATH1 arrays have been extensively used by the *Arabidopsis* community for single-cell type studies and thus are useful for cross-tissue comparisons. RNA-Seq provides increased coverage and sensitivity, and thus, we employed this strategy to gather the most comprehensive analysis of transcriptional activity possible. We present the RNA-Seq data first, followed by the ATH1 data to facilitate the narrative transitions from analysis within a lineage to analysis between lineages.

Libraries suitable for RNA-Seq were generated using RNA extracted from FACS-isolated protoplasts (20,000 cells per replicate except for the stomatal entry marker line at 4,000–5,000 cells per replicate; Figure S2); 22–41 million reads (50 bp) per sample replicate were generated and aligned to 33,602 genes of the *Arabidopsis* TAIR10.18 genome assembly via Bowtie2 and normalized using DESeq2 (Figure S3A; Table S1) (Anders and Huber, 2010; Langmead and Salzberg, 2012). RNA integrity and measures of library quality were equivalently high among samples, but we noticed that the stage 1 (*SSY*) replicates were more divergent than replicates from the later stages. We therefore sequenced an additional *SSY* replicate and resequenced libraries from the two original *SSY* samples in a common lane (Figure S3C). To validate all cell-type profiles, we surveyed previously known regulators of early and late stomatal development and found that their expression matched previously published expression patterns and functions (Figures 1F and S4).

DE Genes and Dominant Expression Patterns within the Stomatal Lineage

Transcript abundance at each intermediate developmental stage is useful, but a more powerful use of the resource is to characterize cohorts of genes and biological processes that define cell states and state transitions in development. As genes defining or changing cell identity may show dynamic expression during lineage progression, we first identified the subset of genes whose expression changed during development and then identified dominant expression patterns within this filtered subset by fuzzy k-means clustering; 11,956 genes were defined as being DE in at least one pairwise comparison using DESeq2 (Figure S3A; Table S1). Because the three stage 1 (*SSY*) samples were transcriptionally distinct (Figure S3C), we considered how best to represent the diverse nature of stomatal entry cells. We ran parallel analyses using either three samples or the *SSY* replicate that clusters closest to the other stomatal cell types in a principal component analysis (Figure S3D). We found results were largely similar using a single *SSY* replicate or all three (Figures S3F and S3G), so to capture the features of the stomatal entry population most broadly, all subsequent analyses (and numbers mentioned below) were carried out using the three *SSY* replicates. To look at overall expression trends among this cohort of potential cell fate regulators, we used an unbiased fuzzy k-means clustering approach. With a stringent cluster membership probability of 0.6, 3,666 genes could be placed in six dominant expression patterns. Several of these patterns corresponded to single-cell stages, such as stomatal entry or

committed cells (Figure 1C, clusters I_R–III_R), whereas others bridged adjacent later stages in the progression of the stomatal lineage (Figure 1C, cluster V_R–VI_R).

We hypothesized that genes assigned to one of these clusters are important for the identity or function of a specific cell type or for transitions between developmental stages. To identify processes associated with those genes and to distinguish one stage from another on a global scale, we looked for enriched gene ontology (GO) terms and promoter sequence motifs within the clustered genes (Figure 1D; Table S3). Not only did the GO terms associated with each cluster correspond to known activities of the cell types they encompass, but the GO terms and dominant expression patterns also reflect a developmental continuum. Transcript accumulation revealed transitions from an undifferentiated cell type (stage 1, cluster II_R) to cells that are still proliferative but establishing an identity (stage 2, cluster III_R) to cells that are differentiating into mature stomata (stages 3 and 4, cluster VI_R). Genes dominantly expressed in the epidermis (cluster I_R) participated in processes such as biosynthesis and metabolism, while genes enriched in differentiating and mature stomata (cluster VI_R) mediate response to signals and stomatal movement, consistent with known activities of these cell types (Kalve et al., 2014; Kim et al., 2010). The biological process enrichment of genes expressed in committed stomatal lineage cells (cluster III_R) is interesting because cell cycle, DNA and chromatin modification and methylation terms point out that these still dividing cells might be poised to switch from a pluripotent to committed state. Equally intriguing is the lack of enriched terms at the stomatal entry stage (cluster II_R and V_R), possibly due to the uncommitted or pluripotent state of cells in this population.

One surprise was the enrichment of the photosynthesis term in clusters I_R (pan-epidermal). This was true using GO term photosynthesis (GO:0015979) or just core light harvesting and carbon fixation genes (Figure S5A). Classical studies suggested that within the epidermis only mature GCs have chloroplasts and the associated photosynthetic gene expression. By analyzing reporters of light harvesting complex (*LHCB1.1*) and carbon fixation (*RBCS2B*) genes (Kim et al., 2003; Sawchuk et al., 2008), however, we confirmed their broader epidermal expression (Figure S5B). Thus, the increased resolution RNA-Seq data provided about transcript abundance have the potential to reveal previously overlooked phenomena.

Comparison of Gene Expression Trends between the Stomatal Lineage and Other Tissues

A major question in developmental biology is to what extent regulatory programs are shared among lineages that must solve similar patterning and fate specification issues, but that ultimately produce different cell types. The stomatal lineage dataset is a clearly linked developmental series, but there are also some available cell profiles derived from young and mature populations of the same cell type (Birnbaum et al., 2003; Brady et al., 2007; He et al., 2012; Yadav et al., 2009, 2014). Most of these profiles were acquired using the ATH1 microarray platform. To be able to compare the stomatal lineage to other individual *Arabidopsis* cell types, we also performed stage-specific genome-wide expression analyses with Affymetrix ATH1 microarrays (Figures 1A and 1B). Intensity values for 20,996 genes were RMA normalized, and informative and non-informative (I/NI) calls

were made to enhance the identification of DE genes (Figure S3B; Table S3). I/NI filtering excludes both noisy genes and housekeeping genes that tend to be expressed across all conditions (Hochreiter et al., 2006; Talloen et al., 2007); 3,114 informative genes were subjected to fuzzy k-means clustering to obtain dominant expression patterns comprised of 1,214 genes at a cluster coefficient cutoff of 0.6. Cluster patterns from the ATH1 dataset were very similar to those derived from RNA-Seq (note clusters I, II, and V in Figures 1C and 1E). Patterns of gene expression obtained by this method were also validated by comparison to known stomatal regulators (Figures 1G and S4) and by the creation of reporters to previously uncharacterized genes (Figures 1H and 1I).

The concurrence in dominant expression patterns between RNA-Seq and ATH1 suggests that both techniques identify major developmental trends, and individual stomatal lineage genes behave similarly in the two experiments (Figures 1F, 1G, and S4). The composition of genes in each cluster, however, differs between the two datasets. This may not be surprising given differences in sample preparation and the differences in how transcript abundance is measured in RNA-Seq and microarray platforms (counting discrete reads versus intensity scores derived from hybridization), which consequentially require different computational analysis pipelines. Moreover, RNA-Seq captures genes that are not present on the ATH1 array (e.g., POLAR-LIKE, Figure 1H, and CYCLIND7;1, Figure 3E). A similarly low correlation was also found when comparing RNA-Seq and ATH1-based transcriptomes derived from female gametophytes (Schmid et al., 2012). We found, in general, that RNA-Seq data captured the patterns of genes expressed at very low levels (e.g., MUTE; Figure S4D) better than the ATH1 array, but that DE patterns appeared more distinct in the ATH1 array (e.g., FLP, Figures 1F, 1G, and S4D).

As with the RNA-Seq data, ATH1-generated late-stage development profiles (clusters V_A–VI_A; Figure 1E) were generally similar to each other and distinct from early stages and from the epidermis (clusters I_A–II_A; Figure 1E). We tested this trend explicitly and quantitatively by using Spearman's rank correlation coefficients (r_s) as a measure of expression profile correlation in pairwise comparisons. Among stomatal lineage cell types, gene expression is highly correlated ($r_s = 0.87$–0.97) with stomatal entry cells (stage 1) being the most distinct cell type and the differentiated and mature stages (stages 3 and 4) the most highly correlated (Figure 2A; Table S4). We then compared the stomatal profiles to profiles of single-cell type populations derived from shoots, roots, and callus (Figure 2B; Table S4) (Brady et al., 2007; He et al., 2012; Yadav et al., 2009, 2014). All expression arrays used in this analysis were reanalyzed in a common computational pipeline to avoid analysis-based biases (see Supplemental Experimental Procedures). These comparisons resulted, not unexpectedly, in lower correlation values ($r_s = 0.54$–0.69), but the overall trend was that the stomatal lineage most closely resembled meristematic and young populations (Figure 2B).

Because comparing overall expression correlations between tissues gives only a very broad view of how similar or dissimilar two cell types are, we used additional methods to define relationships among cell types. Reasoning that regulatory networks comprised of paralogous genes might regulate the development of different tissues, we developed a ranking approach to identify process similarities embedded within unique gene behaviors. The ranking approach is independent of absolute gene expression values, circumventing the problem that different ATH1 datasets showed significantly different hybridization values even when analyzed in a common pipeline. Our ranking approach established "high-priority" genes for a given cell-type dataset by ranking its gene expression values from high to low and then comparing this cell type to high-low rankings derived from other, similar, cell types. Most housekeeping genes are expressed at comparable levels in each cell, but the 5% tails of the distribution represent genes that are "higher priority" in one or the other cell type (Figure 2C). To find genes that might be the cell-type specific solution to a general problem, we looked for common processes (shared enriched GO terms) within these 5% extremes. We choose stage 1 as an example of a transient and uncommitted cell type and stage 3 as an example of a terminally differentiating population. Each stage was compared to the ten nonstomatal lineage cell types with which it exhibited the highest r_s (from experiments reported in Brady et al., 2007; He et al., 2012; Pillitteri et al., 2011; Yadav et al., 2009). Figures 2D and 2E show a selection of genes prioritized in entry and differentiating cells of stomatal lineage development (full gene lists in Table S4).

Among early and late stomatal priority genes, we found several key stomatal regulators that confirm the efficacy of this approach (Figures 2D and 2E). Interestingly, many of the genes not previously associated with stomatal development were members of gene families whose paralogs act in other developmental processes. For example, among stage 1 priority genes, the putative signaling peptide gene ROOT GROWTH FACTOR 9 (RGF9) belongs to a family whose members are required for maintenance of the root stem cell niche (Matsuzaki et al., 2010). RGF9, despite the name, is expressed in leaves but not in roots (Fernandez et al., 2013). Priority gene BEL1-LIKE HOMEODOMAIN 2 (BLH2) is a paralogue of BLR, a regulator of meristem identity and architecture (Kumar et al., 2007). Among stage 3 priority genes, PHOTOTROPIN 2 (PHOT2) and ATP-BINDING CASSETTE B14 (ABCB14) regulate stomatal aperture responses (Briggs and Christie, 2002; Kinoshita et al., 2001); interestingly, our analysis also identified ABCB14's homolog ABCB2, whose function has yet to be ascertained. The ranking approach identified genes whose expression was prioritized in a cross-tissue comparison but tended to show modest transcriptional differences within the stomatal lineage (Tables S1 and S3), thereby providing complementary information to that derived from the fuzzy k-means clustering.

Characterization of Stomatal Division Regulators Identified by Expression Pattern

We initiated these studies, in part, to identify regulators of stomatal development not accessible by classical genetic screens because of redundancy or pleiotropy. To test the utility of our datasets for the first of these issues, we analyzed the EARLY NODULIN-LIKE PROTEIN (ENODL) family that encodes 22 glycosylphosphatidylinisitol (GPI)-anchored proteins whose function has not been ascertained (Mashiguchi et al., 2009). The ENODLs display intriguing patterns in our datasets; expression of ENODL15 and its two most closely related family members, ENODL13 and ENODL14, peaks in stage 1 (Figure 3A). We

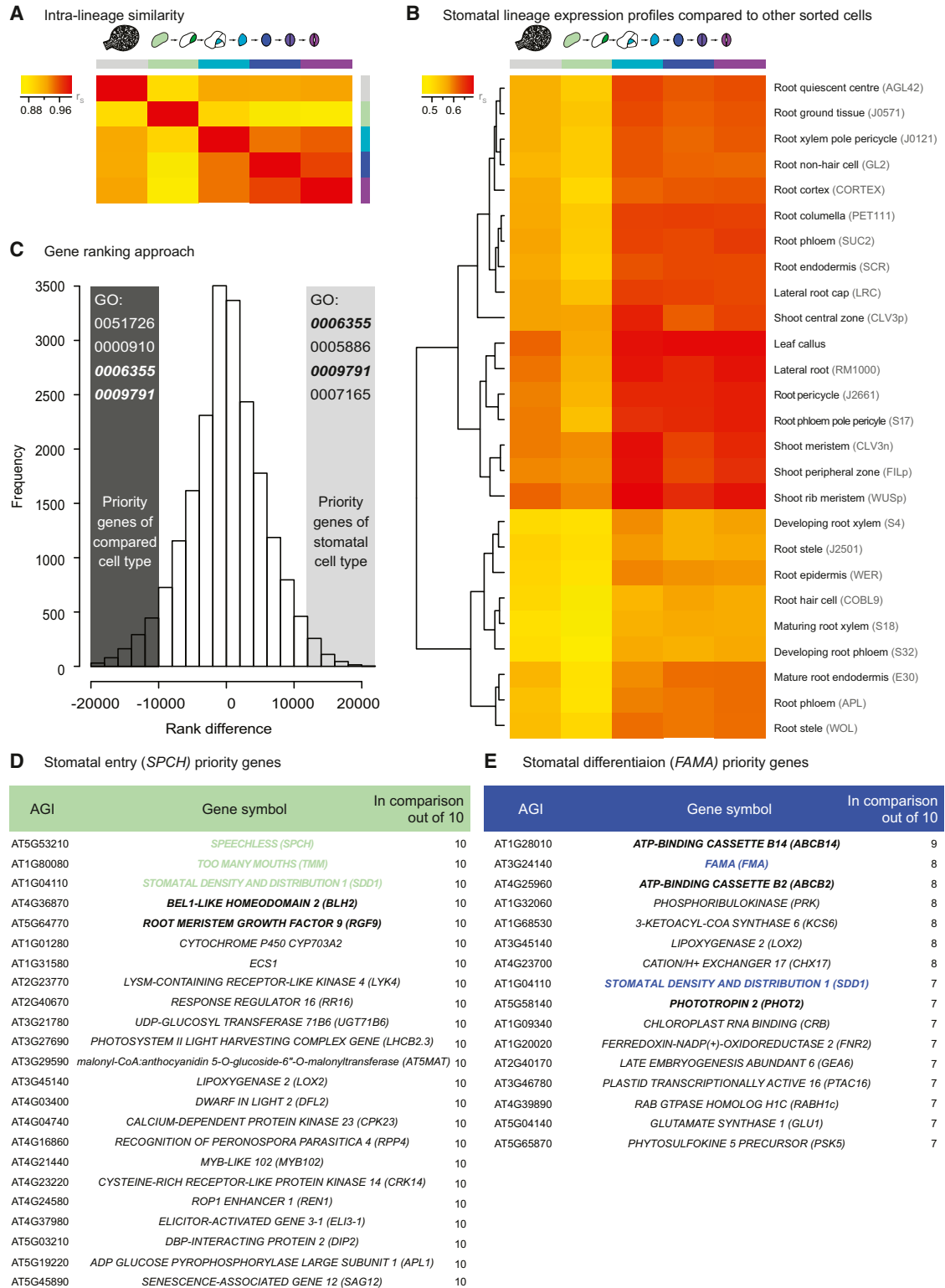**A** Intra-lineage similarity

B Stomatal lineage expression profiles compared to other sorted cells

Root quiescent centre (AGL42)
Root ground tissue (J0571)
Root xylem pole pericycle (J0121)
Root non-hair cell (GL2)
Root cortex (CORTEX)
Root columella (PET111)
Root phloem (SUC2)
Root endodermis (SCR)
Lateral root cap (LRC)
Shoot central zone (CLV3p)
Leaf callus
Lateral root (RM1000)
Root pericycle (J2661)
Root phloem pole pericycle (S17)
Shoot meristem (CLV3n)
Shoot peripheral zone (FILp)
Shoot rib meristem (WUSp)
Developing root xylem (S4)
Root stele (J2501)
Root epidermis (WER)
Root hair cell (COBL9)
Maturing root xylem (S18)
Developing root phloem (S32)
Mature root endodermis (E30)
Root phloem (APL)
Root stele (WOL)

C Gene ranking approach

D Stomatal entry (*SPCH*) priority genes

AGI	Gene symbol	In comparison out of 10
AT5G53210	*SPEECHLESS (SPCH)*	10
AT1G80080	*TOO MANY MOUTHS (TMM)*	10
AT1G04110	*STOMATAL DENSITY AND DISTRIBUTION 1 (SDD1)*	10
AT4G36870	*BEL1-LIKE HOMEODOMAIN 2 (BLH2)*	10
AT5G64770	*ROOT MERISTEM GROWTH FACTOR 9 (RGF9)*	10
AT1G01280	*CYTOCHROME P450 CYP703A2*	10
AT1G31580	*ECS1*	10
AT2G23770	*LYSM-CONTAINING RECEPTOR-LIKE KINASE 4 (LYK4)*	10
AT2G40670	*RESPONSE REGULATOR 16 (RR16)*	10
AT3G21780	*UDP-GLUCOSYL TRANSFERASE 71B6 (UGT71B6)*	10
AT3G27690	*PHOTOSYSTEM II LIGHT HARVESTING COMPLEX GENE (LHCB2.3)*	10
AT3G29590	*malonyl-CoA:anthocyanidin 5-O-glucoside-6"-O-malonyltransferase (AT5MAT)*	10
AT3G45140	*LIPOXYGENASE 2 (LOX2)*	10
AT4G03400	*DWARF IN LIGHT 2 (DFL2)*	10
AT4G04740	*CALCIUM-DEPENDENT PROTEIN KINASE 23 (CPK23)*	10
AT4G16860	*RECOGNITION OF PERONOSPORA PARASITICA 4 (RPP4)*	10
AT4G21440	*MYB-LIKE 102 (MYB102)*	10
AT4G23220	*CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 14 (CRK14)*	10
AT4G24580	*ROP1 ENHANCER 1 (REN1)*	10
AT4G37980	*ELICITOR-ACTIVATED GENE 3-1 (ELI3-1)*	10
AT5G03210	*DBP-INTERACTING PROTEIN 2 (DIP2)*	10
AT5G19220	*ADP GLUCOSE PYROPHOSPHORYLASE LARGE SUBUNIT 1 (APL1)*	10
AT5G45890	*SENESCENCE-ASSOCIATED GENE 12 (SAG12)*	10

E Stomatal differentiaion (*FAMA*) priority genes

AGI	Gene symbol	In comparison out of 10
AT1G28010	*ATP-BINDING CASSETTE B14 (ABCB14)*	9
AT3G24140	*FAMA (FMA)*	8
AT4G25960	*ATP-BINDING CASSETTE B2 (ABCB2)*	8
AT1G32060	*PHOSPHORIBULOKINASE (PRK)*	8
AT1G68530	*3-KETOACYL-COA SYNTHASE 6 (KCS6)*	8
AT3G45140	*LIPOXYGENASE 2 (LOX2)*	8
AT4G23700	*CATION/H+ EXCHANGER 17 (CHX17)*	8
AT1G04110	*STOMATAL DENSITY AND DISTRIBUTION 1 (SDD1)*	7
AT5G58140	*PHOTOTROPIN 2 (PHOT2)*	7
AT1G09340	*CHLOROPLAST RNA BINDING (CRB)*	7
AT1G20020	*FERREDOXIN-NADP(+)-OXIDOREDUCTASE 2 (FNR2)*	7
AT2G40170	*LATE EMBRYOGENESIS ABUNDANT 6 (GEA6)*	7
AT3G46780	*PLASTID TRANSCRIPTIONALLY ACTIVE 16 (PTAC16)*	7
AT4G39890	*RAB GTPASE HOMOLOG H1C (RABH1c)*	7
AT5G04140	*GLUTAMATE SYNTHASE 1 (GLU1)*	7
AT5G65870	*PHYTOSULFOKINE 5 PRECURSOR (PSK5)*	7

Figure 2. Comparing Expression Profiles across Different Cell Types in *Arabidopsis*

(A and B) Heat maps of Spearman's rank correlation coefficients in pairwise comparisons; low correlation is in yellow, and high correlation is in red. Within the stomatal lineage, gene expression of stage 1 cells is least correlated to other stomatal lineage cell types (A). Lower correlations are seen comparing FACs-isolated root, shoot, and leaf callus cells to the stomatal lineage (particularly in stage 1) (B).

confirmed this expression with reporters for *ENODL14* and *ENDOL15* (Figures 3B and 3C). Interestingly, a translational ENODL15 reporter associates with the newly formed cell walls (Figure 3C, inset of meristemoid). Triple mutant plants of genotype *enodl13-1;enodl14-1;enodl15-1* exhibit significant defects in stomatal patterning, a typical consequence of defects in division regulation (Figure 3D).

Association of ENODLs with cell division prompted us to consider cell division regulators more generally. Stages 1–3 in the stomatal lineage are comprised of actively dividing cells (Pillitteri and Torii, 2012). Whether cells at these stages have characteristic cycling behaviors is not known, but such behaviors could be deduced from DE of regulators associated with S phase or G2/M phase. Analysis of core cell-cycle regulators did not reveal enrichment of particular cell-cycle phases, but individual members from among specific groups, such as the cell-cycle inhibitory KIP-RELATED PROTEIN or SIAMESE-RELATED families, and the division-promoting CYCLIN (CYC) D family, exhibited DE (Table S1).

CYCDs are critical for G1/S phase transitions in most organisms (Kalve et al., 2014), and CYCD3 isoforms promote divisions in most *Arabidopsis* tissues (Menges et al., 2006). The stage-restricted expression patterns of *CYCD4;1*, *CYCD6;1*, and *CYCD7;1* (Figure 3E), therefore, are particularly interesting. *CYCD4;1* expression peaks in stage 1 and was previously linked to control of meristemoid divisions through overexpression and loss-of-function studies (Kono et al., 2007). *CYCD6;1* and *CYCD7;1* peak at stage 2, but exhibit different overall patterns with *CYCD7;1* continuing to be enriched in later stages. Consistent with this, a CYCD7;1 reporter accumulates in GMCs and during the GMC to GC transition (Figure 3F). An additional copy of *CYCD7;1* under its native promoter does not affect early lineage proliferation, but instead promotes divisions in GCs (Figure 3G). *CYCD6;1*, which is associated with asymmetric stem-cell divisions in the root (Sozzani et al., 2010), has a profile in the stomatal lineage that parallels the asymmetric amplifying division stage, although the overall expression level is low (Figure 3E), and no stomatal lineage defects have been reported for *cycd6;1*.

Another division control point, the activation of the anaphase promoting complex/cyclosome (APC/C), is regulated by the CDC20 and CCS52 families. The genes encoding these regulators showed stomatal lineage specificity by high expression of a single member, CCS52B, broadly in the stomatal lineage (Figures 3H and 3I). Interestingly, while CCS52A1 and CCS52A1 function in endoreduplication (Liu et al., 2012; Vanstraelen et al., 2009), CCS52B has not been linked to this process. The enriched expression of CCS52B in the stomatal lineage, which does not undergo endoreduplication, suggests that the activity of this paralogue is fundamentally different from the rest of the family.

Evidence for Additional Pluripotency among SPCH-Expressing Leaf Epidermal Cells

The transcriptional map of developmental transitions in the stomatal lineage also provides a backdrop for understanding the gene regulatory landscape of specific "master regulator" basic helix-loop-helix (bHLH) transcription factors. SPEECHLESS (SPCH), like its mammalian homolog MyoD, initiates a cell lineage, and in this role, might potentially reset cells from one state to another. In the RNA-Seq profiles derived from cells in which SPCH protein is active (stage 1), no dominant biological processes (as defined by enriched GO terms) were identified (Table S2). Moreover, these cells stand apart from the later morphologically distinct stomatal lineage cell types, as well as other committed cell types from other organs (Figures 2A and 2B). SPCH is associated with thousands of binding sites in the genome and hundreds of genes are differentially regulated upon its induction (Lau et al., 2014). When the stomatal lineage profiles were compared with the targets of SPCH, there was a significant enrichment of SPCH targets among genes expressed in stages 1 and 2 (Lau et al., 2014). Thus, SPCH has a dominant role in the regulatory hierarchy in these early stages. It is attractive to speculate that some of the expression level variation observed in this early phase reflects the large-scale (SPCH-guided) reprogramming of protodermal cells when they enter the stomatal lineage, and part of this programming may be permissive for the cells to later assume multiple fates.

When considering previously characterized stage 1 enriched SPCH target genes, we were surprised at the extent of overlap between regulators of trichome patterning and the early stomatal lineage. For many years, these two cell types were described as being under independent control in leaves (Figure 4A) (reviewed in Kalve et al., 2014), and profiles derived from whole seedlings overproducing stomata or precursors showed no significantly different expression of trichome-related genes relative to WT (Bergmann et al., 2004; Pillitteri et al., 2011). With our stage-specific profiles, however, we found enrichment of trichome specification genes, such as those encoding the bHLH transcription factors MYC1 and TRANSPARENT TESTA 8 and R3MYB-type transcription factors ENHANCER OF TRIPTYCON AND CAPRICE (ETC) 2 and ETC3, in the early cell stages of stomatal development (Figure 4B). We subsequently confirmed stomatal lineage expression of *MYC1* and *ETC3* reporters (Figures 4C and 4D).

Why might trichome regulators be expressed in the stomatal lineage and be direct targets of SPCH (Figures 4B and 4E)? Based on the prevailing ideas in the literature, we first considered antagonism between the cell types (i.e., stomata are produced at the cost of trichomes and vice versa). If this were true, then negative regulators of trichome development should be preferentially represented in the stomatal lineage, SPCH could promote stomatal identity via upregulation of the trichome repressors, and mutations that reduce trichome numbers should result in overproduction of stomata. *ETC2/3* do indeed repress trichome fate (Wester et al., 2009), and *ETC3* is upregulated in response to SPCH induction (Figure 4F). However, *MYC1* promotes trichome formation (Zhao et al., 2012) and is also upregulated by SPCH (Figure 4F). Loss of *MYC1* does result in a significantly higher stomatal index (Figure 4G), but there is no change in

(C–E) A ranking approach compares stomatal stages 1 and 3 with the ten most highly correlated non-stomatal cell-type specific datasets from (B). Genes were ranked corresponding to their expression within a dataset and the difference in ranking calculated (C). High-priority genes fall in the top and bottom 5% of the graph; from these, common enriched GO terms were used to enrich for genes that contribute to a similar process. Selection of genes prioritized in multiple comparisons in entry (D) and differentiation (E) samples (full gene lists in Table S4).
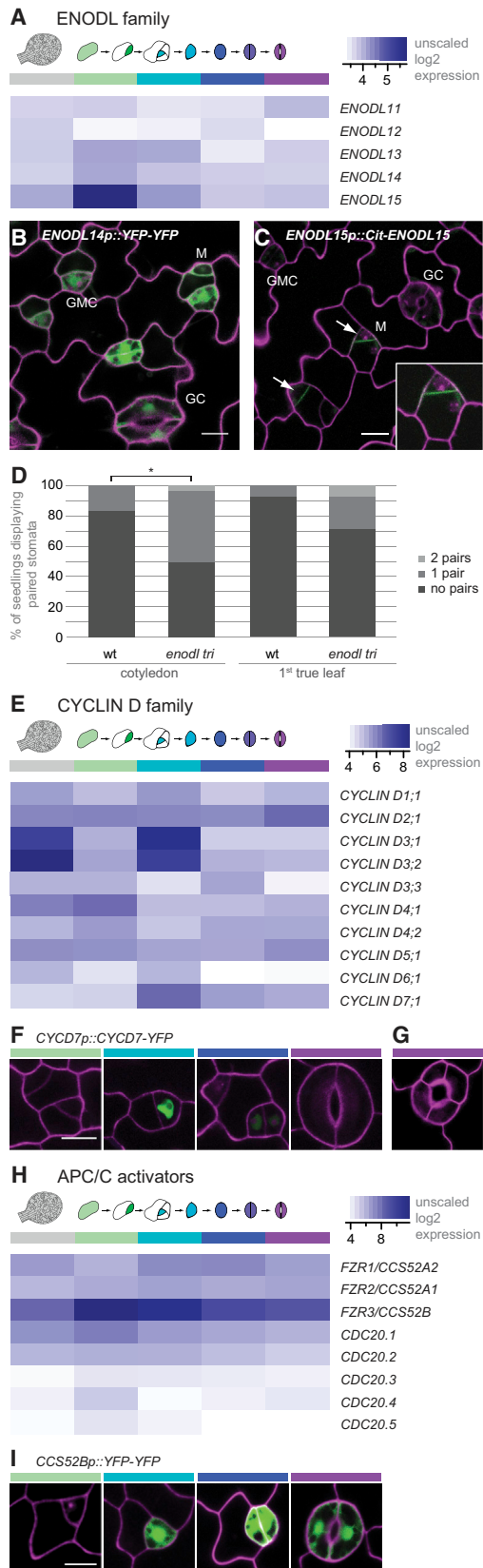
A ENODL family

ENODL11
ENODL12
ENODL13
ENODL14
ENODL15

B *ENODL14p::YFP-YFP*

C *ENODL15p::Cit-ENODL15*

D

E CYCLIN D family

CYCLIN D1;1
CYCLIN D2;1
CYCLIN D3;1
CYCLIN D3;2
CYCLIN D3;3
CYCLIN D4;1
CYCLIN D4;2
CYCLIN D5;1
CYCLIN D6;1
CYCLIN D7;1

F *CYCD7p::CYCD7-YFP*

G

H APC/C activators

FZR1/CCS52A2
FZR2/CCS52A1
FZR3/CCS52B
CDC20.1
CDC20.2
CDC20.3
CDC20.4
CDC20.5

I *CCS52Bp::YFP-YFP*

Figure 3. Genes DE in Stages of the Stomatal Lineage Map Have Roles in Stomatal Development

(A) Expression patterns of *ENODLs* suggest roles for *ENODL15*, *ENODL14*, and *ENODL13* during stomatal lineage development.

(B and C) Confocal images of *ENODL14* transcriptional (B) and *ENODL15* translational (C) reporters confirm cell-type-specific expression in 4-day-old cotyledons. White arrow points to ENODL15 accumulation at division planes. The inset shows meristemoid at higher magnification.

(D) *enodl13-1;enodl14-1;enodl15-1* mutants exhibit a higher frequency of mispatterned stomata in cotyledons and true leaves. y axis shows percentage of seedlings displaying stomatal pairs in a given leaf area. Mann Whitney test, *$p < 0.05$.

(E) Stage-specific expression enrichment for some *CYCD* family members.

(F) GMC-specific expression of *CYCD7p::CYCD7;1-YFP*; bars are color coded as in Figure 1A.

(G) Expression of *CYCD7p::CYCD7;1-YFP* promotes extra GC divisions.

(H) Expression pattern of APC/C activator genes emphasizes uniquely high expression of *CCS52B* in the stomatal lineage.

(I) Confocal images of *CCS52B* reporter expression in stomatal lineage cells. All heat maps show unscaled mean and median log2-transformed expression values; low expression is depicted in white and high expression in blue. In confocal images, reporter signal is in green; cell outlines are in magenta. Scale bars represent 10 μm.

stomatal production in *etc1;etc2;etc3* mutants (Figure 4H) and overexpression of the stomatal repressor *TOO MANY MOUTHS* reduces trichome numbers (Yan et al. 2014). Together, these data do not support an antagonism model. Instead, they suggest the stem-cell like precursor stage that generates pavement cells and stomata may actually contribute to all epidermal cell types (Figure 4A, red arrow). Genes associated with trichome maturation were not enriched in the stomatal lineage (Table S1), consistent with a bifurcation in cell fate and gene expression occurring when cells progress past this pluripotent early stage.

DISCUSSION

The sequencing and microarray-based profiles of cells transiting through the stomatal lineage are resources that can be explored and exploited in numerous ways. We identified regulators of stomatal development among the genes showing restricted expression patterns: here with the *ENODL* family and *CYCD7;1* and in a recent complementary study with genes encoding Brassinosteroid signaling pathway components and an asymmetric division-regulating kinesin (Lau et al., 2014). For the stomatal signaling community, the extended stage and transcript coverage provided by our RNA-Seq profiles of developing and mature GCs is likely to be of considerable use. The ability to compare related and still-proliferating cell types allowed us to find core cell-cycle genes that were nonetheless tailored to participation in different specific divisions (e.g., *CYCD4;1* and *CYD7;1*). The fact that the early-stage divisions are asymmetric and genes implicated in the regulation of asymmetric divisions (*BASL, POLAR, ARK3*) are enriched precisely in that stage (Tables S1 and S3) suggests the stomatal lineage expression map will be a good resource for uncovering more regulators of asymmetric and oriented divisions.

To use these data to guide future investigations, it is important to have confidence that the techniques used here—RNA-Seq and microarray—faithfully recapitulate in vivo expression data. Many different organisms and biological questions previously analyzed with microarrays are now shifting to sequencing based
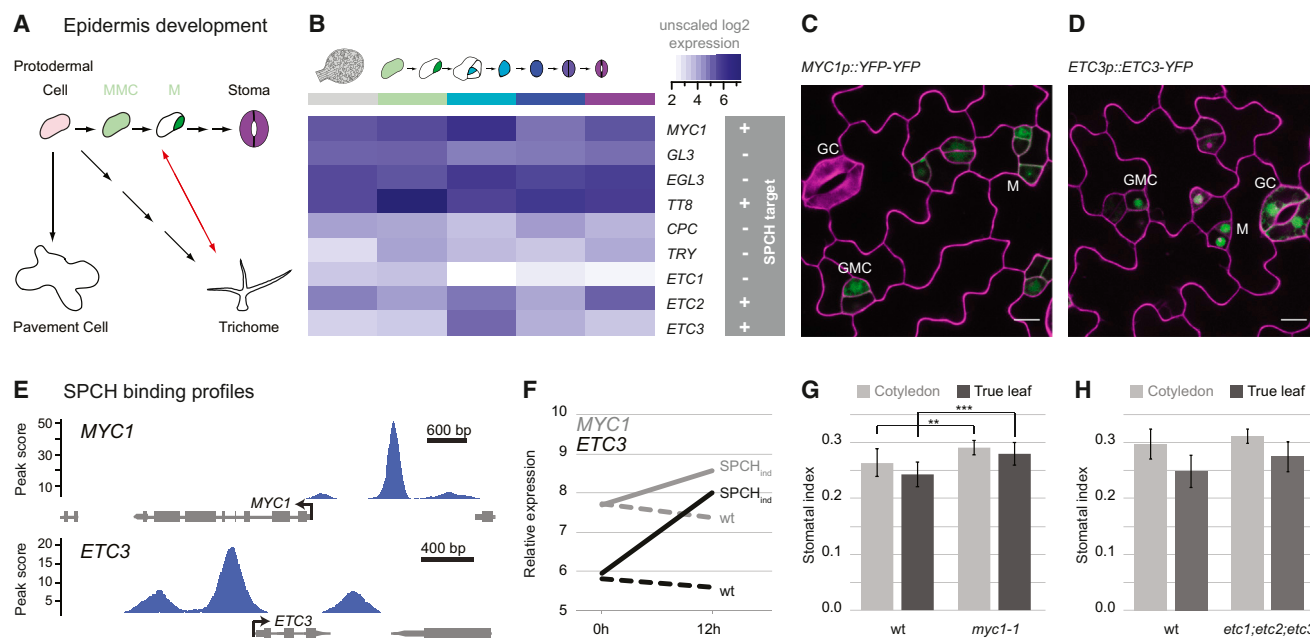
Figure 4. Cross-Talk between Developmental Pathways in the Leaf Epidermis

(A) Classical (black arrows) and updated (red arrow) view of cell lineages in the leaf. Our data suggest a pluripotent stage from which both trichomes and stomata are derived.

(B) Trichome-specifying transcription factors *MYC1*, *TT8*, *ETC2*, and *ETC3* show high transcript abundance in early stomatal cells. Heat map shows unscaled mean and median log2-transformed expression values; low expression is depicted in white and high expression in blue.

(C and D) Expression of *MYC1* and *ETC3* reporters (green) at stage 1 and 2 of stomatal development in 4-day-old cotyledons. Scale bars represent 10 μm.

(E) *MYC1* and *ETC3* are direct targets of the stomatal lineage key regulator SPCH. The y axis represents the computed enrichment score of SPCH binding, and arrows indicate transcriptional start sites and orientation of genes; data are derived from Lau et al., (2014).

(F) Induction of SPCH leads to upregulation of *MYC1* and *ETC3* transcripts. y axis represents relative expression; x axis represents times after induction in hours. Data are derived from (Lau et al., 2014).

(G) *myc1* plants show a higher stomatal index than WT, suggesting interactions between regulation of trichomes and stomata. Stomatal index is shown as mean ± SD of 7–17 seedlings. Mann Whitney test, **p < 0.01, ***p < 0.001.

(H) *etc1;etc2;etc3* plants show a non-significant stomatal index increase relative to WT.

expression profiling; in each of these cases, there is discussion about which platform provides the most accurate results (Wang et al., 2009, 2014; Zhao et al., 2014), but as yet, there are no sophisticated methods for cross-platform analyses. There are merits to both approaches, as we have explored in this paper, and we have been able to validate specific expression of reporters derived from both; however, it is clear that we are reading two different types of signals and consequently sampling different transcript populations in the analysis pipeline. RNA-Seq has a greater dynamic range and can detect very low expression levels, and as a consequence of including these rare transcripts, can exhibit more sample to sample variance. Given the wealth of data in both platforms, it will be important to develop better computational tools; parallel datasets such as those provided here are essential for those efforts.

We also pondered the meaning of the expression variance among replicates in the earliest stomatal lineage samples. These populations are extremely interesting from a developmental perspective, but because they are transient and rare, they are difficult to access. We ruled out purely technical explanations for variation among stage 1 samples (RNA integrity and measures of library quality were equivalent to later stages) and lack of enrichment for GO term categories for stress or environmental

response suggests that stage 1 cells are not more sensitive to perturbation by sample preparation. Stage 1 includes both MMCs and meristemoids, but this alone is not sufficient to explain the variation because the replicates of *ML1Y* samples (composed of all epidermal cell types) are highly correlated. One explanation for the variance in the stage 1 samples that encompasses technical and biological issues comes from the fact that FACS is a quantitative detection of florescence signals, thresholded for each experiment. *SSY* expression peaks just prior to and after the asymmetric cell divisions of MMCs and is also brightest in the youngest leaves (Figures S1C–S1E). Small random fluctuations in the brightness of the reporter between batches of plants could lead to some replicates only capturing the youngest, brightest cells (Figure S1C, right), and others containing a broader representation of Meristemoids and MMC stages (Figure S1C, left). Importantly, despite the potential for gene expression differences between these substages, all of these cells are still part of the stomatal precursor population, and all of our data suggest that the source of variation at these early stages is intrinsic to the biology of these uncommitted cell types. In the future, single-cell sequencing, better markers, or more sophisticated computational approaches may help to dissect the source and meaning of the expression variation.

Moving beyond study of the leaf epidermis, the stomatal lineage serves as an important counterpoint to apical meristems of the shoot and root. Because there is low correspondence between gene expression studies derived from single cell types and those from whole seedlings enriched for those same cell types (e.g., between this study and Pillitteri et al., 2011), the single cell-type profiles will enable more sophisticated analysis of gene expression patterns across organs. For example, quantitative methods to assign expression specificity values to individual genes (Birnbaum and Kussell, 2011) can be trained with these additional cell types. As new RNA-Seq based datasets emerge from other tissues, the stomatal lineage data can be immediately incorporated into those comparisons. To facilitate use of these data by the developmental and systems biology communities, these data have been provided as extensively annotated tables (Tables S1 and S3) and have been deposited in GEO as GSE58857. A user-friendly graphical representation of expression levels along the developmental progression has been made compatible with the online eFP browser (Winter et al., 2007).

Monitoring of transcriptional profiles during reprogramming of mammalian cells into induced pluripotent stem cells (iPSCs) revealed initial stochastic gene expression followed by more predictable and hierarchical patterns as cells acquired well-defined fates (Buganim et al., 2012). The overall pattern is similar among stomatal lineage cell types: stage 1 samples have the most variation among replicates (Figures S3C and S3E), followed by decreasing amounts of variation in stages 2–4. In stage 2 there is a strong enrichment of GO terms related to DNA methylation and chromatin modification, followed by enrichment of modules for differentiation of specific cell types in stages 3 and 4. Such a pattern would be consistent with a pluripotent early state expressing transcripts for many possible outcomes followed by a permanent setting of a more limited program for a single cell type. Although specific genes that specify plant and mammalian stem cells are likely to differ, the underlying regulatory logic may be similar. In light of this, the unique attributes of plant development—new organs and new stem cell lineages initiated postembryonically and continuously—may lend themselves to elucidation of developmental regulatory mechanisms more difficult to address in the hidden stem-cell niches of animals.

EXPERIMENTAL PROCEDURES

Plant Reporter Lines and Mutants

Previously described reporters used for cell isolation were *ML1p::YFP-RCI2A* (Roeder et al., 2010), *SPCHp::SPCH-YFP* (Davies and Bergmann, 2014), *MUTEp::nucGFP* (MacAlister et al., 2007), *FAMAp::GFP-FAMA* (Ohashi-Ito and Bergmann, 2006), and *E1728::GFP* (Gardner et al., 2009). New reporters were created by amplifying appropriate genome sequences (PCR amplified from Col) into vectors compatible with the binary *R4pGWB* destination vector system (Nakagawa et al., 2008; Tanaka et al., 2011).

FACS and RNA Extraction

Protoplast isolation and FACS from reporter line seedlings were performed as described in (Bargmann and Birnbaum, 2010) for FACS on a FACSAria II (BD Biosciences) fitted with a 100 μm nozzle. Non-GFP/YFP protoplasts from WT seedlings were used to define gate boundaries (Figures S2A and S2B). Protoplast signals gated for RNA sequencing analysis are shown in Figure S2A. Positive events were sorted directly into 350 ml RNA extraction buffer (RNeasy

Micro Kit, QIAGEN) and total RNA of extracted with the RNeasy Micro Kit (QIAGEN) including on column DNase treatment.

RNA Sequencing and Data Analysis

cDNA libraries for high-throughput sequencing were prepared from ~10 ng of total RNA from each cell sample. cDNA was generated using the PrepX SPIA RNA-Seq Kit, and libraries were generated using the ILM DNA Library Kit (Wafergen) according to the manufacturer's protocol. Fifty-bp single end reads were generated from a HiSeq2000 sequencer (Illumina) in high-output mode. The TruSeq SR Cluster Kit v3–cBot–HS (GD-401-3001) was used for cluster generation and TruSeq SBS Kit v3–HS (50 cycles) (FC-401-3002) for sequencing. All sequencing and data analysis (to fastq files) were done using Illumina's standard protocols and bcl2fastq software. RNA-Seq data files are in GEO (GSE58856). Reads were aligned to the TAIR10.18 genome assembly via Bowtie2 (Langmead and Salzberg, 2012) and counts normalized via DESeq2 (Anders and Huber, 2010), both using default settings. Analyses are based on the mean expression values of two replicates (except for SSY in which 3 or 1 were used). DE was calculated for all possible pairwise comparisons via DESeq2 (Anders and Huber, 2010) with FDR < 0.05. DE genes were clustered via FANNY (Maechler et al., 2012) with k = 6 and a probability cutoff of 0.6.

Microarray Hybridization and Data Analysis

RNA extracted for ATH1-microarray based expression profiling (three replicates per cell type) was reverse transcribed and amplified using the Ovation Pico WTA System (NuGen). cDNA was labeled with the Encore Biotin Module (NuGen) before hybridization to the ATH-121501 microarray (Affymetrix) and processed using standard procedures on a GeneChip Fluidics Station 450 and a GeneChip Scanner. Microarray data files are in GEO (GSE58855). Gene expression was normalized using the R package, *affy* (Gautier et al., 2004) with following parameters: RMA background correction, quantile normalization across all arrays, no perfect match (PM) probe correction, and the median polish method for summarization (Irizarry et al., 2003). Computation of log2 scale expression values from probe sets was carried out by the median polish method. Informative and noninformative (I/NI) calls were made via *FARMS* (Hochreiter et al., 2006; Talloen et al., 2007). Median expression values (from three replicates) of informative genes were clustered via FANNY (Maechler et al., 2012) with k = 6 and a probability cutoff of 0.6.

Cell-type specific comparison datasets (ATH1-based) from (Brady et al., 2007; He et al., 2012; Yadav et al., 2014; 2009) were reanalyzed alongside stomatal linage data to make pairwise Spearman correlations. The ten cell types with the highest Spearman correlations to either stage 1 or stage 3 were subsequently used in the ranking analysis.

ACCESSION NUMBERS

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.devcel.2015.01.025.

AUTHOR CONTRIBUTIONS

J. Adrian, J.C., B.O.R.B., and D.C.B. helped with conception and design. J. Adrian contributed to cell sorting, ATH1 data analysis, phenotypic analysis of transgenic lines, and project coordination. J.C. provided ATH1 ranking analysis, RNA-Seq data analysis, and revised the article. C.E.B. helped with cloning, screening, and phenotypic analysis of transgenic lines. B.O.R.B. set up the sorting experiment. J. Alassimone performed confocal analysis of reporter lines. K.A.D., O.S.L., A.L., J.L.M., C.H., and A.V. performed candidate gene characterization. K.D.B. contributed consultation on experimental design and initial sorting with B.O.R.B. D.C.B. interpreted data. D.C.B. and J. Adrian drafted and revised the article.

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Bargmann, B.O.R., and Birnbaum, K.D. (2010). Fluorescence activated cell sorting of plant protoplasts. J. Vis. Exp. *36*, 1673.

Bergmann, D.C., Lukowitz, W., and Somerville, C.R. (2004). Stomatal development and pattern controlled by a MAPKK kinase. Science *304*, 1494–1497.

Birnbaum, K.D., and Kussell, E. (2011). Measuring cell identity in noisy biological systems. Nucleic Acids Res. *39*, 9093–9107.

Birnbaum, K., Shasha, D.E., Wang, J.Y.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., and Benfey, P.N. (2003). A gene expression map of the Arabidopsis root. Science *302*, 1956–1960.

Brady, S.M., Orlando, D.A., Lee, J.Y., Wang, J.Y.Y., Koch, J., Dinneny, J.R., Mace, D., Ohler, U., and Benfey, P.N. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. Science *318*, 801–806.

Briggs, W.R., and Christie, J.M. (2002). Phototropins 1 and 2: versatile plant blue-light receptors. Trends Plant Sci. *7*, 204–210.

Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell *150*, 1209–1222.

Casson, S.A., and Hetherington, A.M. (2014). phytochrome B Is required for light-mediated systemic control of stomatal development. Curr. Biol. *24*, 1216–1221.

Davies, K.A., and Bergmann, D.C. (2014). Functional specialization of stomatal bHLHs through modification of DNA-binding and phosphoregulation potential. Proc. Natl. Acad. Sci. USA *111*, 15585–15590.

Deal, R.B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. Dev. Cell *18*, 1030–1040.

Engineer, C.B., Ghassemian, M., Anderson, J.C., Peck, S.C., Hu, H., and Schroeder, J.I. (2014). Carbonic anhydrases, EPF2 and a novel protease mediate CO2 control of stomatal development. Nature *513*, 246–250.

Fernandez, A., Drozdzecki, A., Hoogewijs, K., Nguyen, A., Beeckman, T., Madder, A., and Hilson, P. (2013). Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation. Plant Physiol. *161*, 954–970.

Gardner, M.J., Baker, A.J., Assie, J.M., Poethig, R.S., Haseloff, J.P., and Webb, A.A.R. (2009). GAL4 GFP enhancer trap lines for analysis of stomatal guard cell development and gene expression. J. Exp. Bot. *60*, 213–226.

Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics *20*, 307–315.

He, C., Chen, X., Huang, H., and Xu, L. (2012). Reprogramming of H3K27me3 is critical for acquisition of pluripotency from cultured Arabidopsis tissues. PLoS Genet. *8*, e1002911.

Hetherington, A.M., and Woodward, F.I. (2003). The role of stomata in sensing and driving environmental change. Nature *424*, 901–908.

Hochreiter, S., Clevert, D.A., and Obermayer, K. (2006). A new summarization method for Affymetrix probe level data. Bioinformatics *22*, 943–949.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249–264.

Kalve, S., De Vos, D., and Beemster, G.T.S. (2014). Leaf development: a cellular perspective. Front. Plant Sci. *5*, 362.

Kim, J.Y., Yuan, Z., and Jackson, D. (2003). Developmental regulation and significance of KNOX protein trafficking in Arabidopsis. Development *130*, 4351–4362.

Kim, T.-H., Böhmer, M., Hu, H., Nishimura, N., and Schroeder, J.I. (2010). Guard cell signal transduction network: advances in understanding abscisic acid, CO2, and Ca2+ signaling. Annu. Rev. Plant Biol. *61*, 561–591.

Kinoshita, T., Doi, M., Suetsugu, N., Kagawa, T., Wada, M., and Shimazaki, K. (2001). Phot1 and phot2 mediate blue light regulation of stomatal opening. Nature *414*, 656–660.

Kono, A., Umeda-Hara, C., Adachi, S., Nagata, N., Konomi, M., Nakagawa, T., Uchimiya, H., and Umeda, M. (2007). The Arabidopsis D-type cyclin CYCD4 controls cell division in the stomatal lineage of the hypocotyl epidermis. Plant Cell *19*, 1265–1277.

Kumar, R., Kushalappa, K., Godt, D., Pidkowich, M.S., Pastorelli, S., Hepworth, S.R., and Haughn, G.W. (2007). The Arabidopsis BEL1-LIKE HOMEODOMAIN proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. Plant Cell *19*, 2719–2735.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lau, O.S., Davies, K.A., Chang, J., Adrian, J., Rowe, M.H., Ballenger, C.E., and Bergmann, D.C. (2014). Direct roles of SPEECHLESS in the specification of stomatal self-renewing cells. Science *345*, 1605–1609.

Liu, Y., Ye, W., Li, B., Zhou, X., Cui, Y., Running, M.P., and Liu, K. (2012). CCS52A2/FZR1, a cell cycle regulator, is an essential factor for shoot apical meristem maintenance in Arabidopsis thaliana. BMC Plant Biol. *12*, 135.

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

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2012). Cluster: Cluster Analysis Basics and Extensions. R package version 1.2.

Mashiguchi, K., Asami, T., and Suzuki, Y. (2009). Genome-wide identification, structure and expression studies, and mutant collection of 22 early nodulin-like protein genes in Arabidopsis. Biosci. Biotechnol. Biochem. *73*, 2452–2459.

Matos, J.L., Lau, O.S., Hachez, C., Cruz-Ramírez, A., Scheres, B., and Bergmann, D.C. (2014). Irreversible fate commitment in the Arabidopsis stomatal lineage requires a FAMA and RETINOBLASTOMA-RELATED module. eLife. Published online October 10, 2014. http://dx.doi.org/10.7554/eLife.03271.

Matsuzaki, Y., Ogawa-Ohnishi, M., Mori, A., and Matsubayashi, Y. (2010). Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. Science *329*, 1065–1067.

Menges, M., Samland, A.K., Planchais, S., and Murray, J.A.H. (2006). The D-type cyclin CYCD3;1 is limiting for the G1-to-S-phase transition in Arabidopsis. Plant Cell *18*, 893–906.

Misra, B.B., Assmann, S.M., and Chen, S. (2014). Plant single-cell and single-cell-type metabolomics. Trends Plant Sci. *19*, 637–646.

Mustroph, A., Zanetti, M.E., Jang, C.J.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J. (2009). Profiling translatomes of discrete

cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. Proc. Natl. Acad. Sci. USA *106*, 18843–18848.

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.

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

Pillitteri, L.J., and Torii, K.U. (2012). Mechanisms of stomatal development. Annu. Rev. Plant Biol. *63*, 591–614.

Pillitteri, L.J., Sloan, D.B., Bogenschutz, N.L., and Torii, K.U. (2007). Termination of asymmetric cell division and differentiation of stomata. Nature *445*, 501–505.

Pillitteri, L.J., Peterson, K.M., Horst, R.J., and Torii, K.U. (2011). Molecular profiling of stomatal meristemoids reveals new component of asymmetric cell division and commonalities among stem cell populations in Arabidopsis. Plant Cell *23*, 3260–3275.

Robinson, S., Barbier de Reuille, P., Chan, J., Bergmann, D., Prusinkiewicz, P., and Coen, E. (2011). Generation of spatial patterns through cell polarity switching. Science *333*, 1436–1440.

Roeder, A.H.K., Chickarmane, V., Cunha, A., Obara, B., Manjunath, B.S., and Meyerowitz, E.M. (2010). Variability in the control of cell division underlies sepal epidermal patterning in Arabidopsis thaliana. PLoS Biol. *8*, e1000367.

Sawchuk, M.G., Donner, T.J., 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.

Schmid, M.W., Schmidt, A., Klostermeier, U.C., Barann, M., Rosenstiel, P., and Grossniklaus, U. (2012). A powerful method for transcriptional profiling of specific cell types in eukaryotes: laser-assisted microdissection and RNA sequencing. PLoS ONE *7*, e29685.

Sozzani, R., Cui, H., Moreno-Risueno, M.A., Busch, W., Van Norman, J.M., Vernoux, T., Brady, S.M., Dewitte, W., Murray, J.A.H., and Benfey, P.N. (2010). Spatiotemporal regulation of cell-cycle genes by SHORTROOT links patterning and growth. Nature *466*, 128–132.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE *6*, e21800.

Talloen, W., Clevert, D.-A., Hochreiter, S., Amaratunga, D., Bijnens, L., Kass, S., and Göhlmann, H.W.H. (2007). I/NI-calls for the exclusion of non-informative genes: a highly effective filtering tool for microarray data. Bioinformatics *23*, 2897–2902.

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.

Vanstraelen, M., Baloban, M., Da Ines, O., Cultrone, A., Lammens, T., Boudolf, V., Brown, S.C., De Veylder, L., Mergaert, P., and Kondorosi, E. (2009). APC/C-CCS52A complexes control meristem maintenance in the Arabidopsis root. Proc. Natl. Acad. Sci. USA *106*, 11806–11811.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*, 57–63.

Wang, C., Gong, B., Bushel, P.R., Thierry-Mieg, J., Thierry-Mieg, D., Xu, J., Fang, H., Hong, H., Shen, J., Su, Z., et al. (2014). The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. Nat. Biotechnol. *32*, 926–932.

Wester, K., Digiuni, S., Geier, F., Timmer, J., Fleck, C., and Hülskamp, M. (2009). Functional diversity of R3 single-repeat genes in trichome development. Development *136*, 1487–1496.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS ONE *2*, e718.

Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D., et al. (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell *153*, 1134–1148.

Yadav, R.K., Girke, T., Pasala, S., Xie, M., and Reddy, G.V. (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. Proc. Natl. Acad. Sci. USA *106*, 4941–4946.

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

Yan, L., Cheng, X., Jia, R., Qin, Q., Guan, L., Du, H., and Hou, S. (2014). New phenotypic characteristics of three tmm alleles in Arabidopsis thaliana. Plant Cell Rep. *33*, 719–731.

Yang, Y., Costa, A., Leonhardt, N., Siegel, R.S., and Schroeder, J.I. (2008). Isolation of a strong Arabidopsis guard cell promoter and its potential as a research tool. Plant Methods *4*, 6.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940–954.

Zhao, Z., Zhang, W., Stanley, B.A., and Assmann, S.M. (2008). Functional proteomics of Arabidopsis thaliana guard cells uncovers new stomatal signaling pathways. Plant Cell *20*, 3210–3226.

Zhao, H., Wang, X., Zhu, D., Cui, S., Li, X., Cao, Y., and Ma, L. (2012). A single amino acid substitution in IIIf subfamily of basic helix-loop-helix transcription factor AtMYC1 leads to trichome and root hair patterning defects by abolishing its interaction with partner proteins in Arabidopsis. J. Biol. Chem. *287*, 14109–14121.

Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K., and Liu, X. (2014). Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS ONE *9*, e78644.