Stem cells and fate control in plant stomatal development

Kylee M. Peterson

A dissertation
submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington
2013

Reading Committee:
Keiko Torii, Chair
Christine Queitsch
Verónica di Stilio

Program Authorized to Offer Degree:
Biology

University of Washington

Abstract

Stem cells and fate control in plant stomatal development

Kylee M. Peterson

Chair of the Supervisory Committee:
Professor Keiko U. Torii
Department of Biology

The plant epidermis is a critical interface between the atmosphere and internal plant tissues, which allows plants to succeed on land by restricting their exposure to the environment. Stomata, closable pores on the plant surface bounded by specialized guard cells, are an integral part of epidermal function. By controlling water loss and carbon dioxide uptake, stomata regulate global carbon and water cycles. Stomata undergo a complex course of development involving cell-cell signaling and sequential action of master regulatory transcription factors, making stomata a suitable system for studying widely applicable developmental processes necessary for tissue and organ development. This dissertation presents an examination of molecular characteristics, gene function, and protein dynamics in stomatal lineage cells, including preliminary results of manipulating cell fate during stomatal development.

We performed molecular profiling of the stem-cell-like stomatal precursor, the meristemoid, by enriching cell types through synthetic mutations. This uncovered new genes involved in stomatal development as well as molecular commonalities with the main plant stem-cell niches at the shoot and root apices. A novel gene, *POLAR*, was found to localize asymmetrically in dividing stomatal-lineage cells. Also uncovered through meristemoid profiling was a

transcription factor, HOMEODOMAIN GLABROUS 2 (HDG2), which is highly expressed in stomatal lineage cells and sufficient to convert internal leaf cells to stomata when ectopically expressed. Loss of function in *HDG2* hinders stomatal development after initiation and causes aberrant stomata; further loss of its close relative *AtML1* magnifies the effect. To further investigate the role of transcription factors in stomatal development, we used time lapse microscopy to observe protein dynamics of stomatal regulators in germinating cotyledons. Cell-cell signaling in the cotyledon was perturbed using laser ablation of stomatal-lineage cells, and preliminary results indicate that cell fate was thus affected. Cotyledon time lapse revealed an unexpected developmental sequence indicating possible epidermal prepatterning, so we employed an embryo time lapse technique to discover that both regulatory genes and signaling components were active in the embryo, indicating that stomatal development begins during embryonic development. This work demonstrates that stomatal development exemplifies crucial developmental processes and provides novel insight into how cell fate is dynamically specified by tissue-level regulation, cell-cell signaling, and cell-autonomous molecular mechanisms in plants.

Supplementary Files

Movie 2-1. Time-lapse imaging of POLAR localization in germinating cotyledon..

Movie 2-2. Time-lapse imaging of POLAR localization in germinating cotyledon of *basl*.

Movie 3-1. Complete optical sectioning of a wild-type 10-day-old cotyledon.

Movie 3-2. Complete optical sectioning of a 10-day-old cotyledon under *HDG2* induced ectopic overexpression.

Movie 3-3. Complete optical sectioning of a 10-day-old cotyledon under *AtML1* induced ectopic overexpression.

Movie 4-1. Time-lapse imaging of *SPCHpro::SPCH-GFP* localization in germinating cotyledon..

Movie 4-2. Time-lapse imaging of *MUTEpro::MUTE-GFP* localization in germinating cotyledon.

Movie 4-3. Time-lapse imaging of *SCRMpro::GFP-SCRM* localization in germinating cotyledon.

Movie 4-4. Laser cell ablation of *SCRMpro::GFP-SCRM* putative GMC in germinating cotyledon.

Movie 5-1. *SPCHpro::SPCH-GFP* expression in asymmetrically dividing embryonic cotyledon cells.

Movie 5-2. *MUTEpro::MUTE-GFP* expression in non-dividing embryonic cotyledon cells.

Supplementary Data Set 2-1. Genes showing high expression in *scrm-D mute*.

Supplementary Data Set 2-2. Genes showing high expression in *scrm-D*.

Supplementary Data Set 2-3. Genes showing high expression in *spch*.

Supplementary Data Set 2-4. Genes significantly deregulated in the meristemoid-enriched background, *scrm-D mute*.

Supplementary Data Set 2-5. Genes significantly deregulated in the stomata-enriched background, *scrm-D*.

Supplementary Data Set 2-6. Genes significantly deregulated in the pavement-cell-enriched background, *spch*.

Table of Contents

Chapter 1.

Introduction.

This chapter has been revised and updated from:

Kylee M. Peterson*, Amanda L. Rychel*, and Keiko U. Torii (2010) Out of the Mouths of Plants: The Molecular Basis of the Evolution and Diversity of Stomatal Development. *Plant Cell* 22: 296–306.

* These authors contributed equally to this work.

Introduction

The evolution of stomata (Greek for mouths; singular, stoma) was a crucial adaptation occurring some 400 million years ago that permitted plants to thrive on land, which they began to colonize some 450 million years ago. To survive in the dry atmosphere, plants must protect against desiccation yet allow for the gas exchange necessary for photosynthesis and respiration. A watertight epidermis with closable valves, the stomata, was the solution. Stomata are composed of paired, symmetric guard cells that operate by changing turgor pressure, a system that forms an effective fail-safe against dehydration, since low turgor closes stomata. Transpiration, or water loss through stomata, promotes upward and outward water movement in plants and cools the plant surface via thermal dissipation. The distribution of stomata is far from random; they are evenly distributed, an adaptation allowing gases to reach interior tissues effectively, and they adhere closely to the one-cell spacing rule (Sachs, 1991), which means that stomata are always found with at least one non-stomatal epidermal cell separating them. This one-cell spacing rule is important for proper opening and closing of the stomatal aperture, which requires efficient exchange of water and ions

with neighboring nonstomatal cells. Stomatal development must be tightly controlled to accomplish these goals.

The fundamental physiological importance of stomata is underscored by the observation that their evolution predates that of flowers, leaves, roots, and even vascular systems (Freeman, 2008) (Figure 1-1). The fossil record indicates that the earliest nonvascular and vascular plants, such as *Cooksonia*, *Rhynia*, and *Aglaophyton*, possessed a simple architecture of bifurcated stems with apical spore capsules (Edwards et al., 1998), very different from modern land plants. Yet these ancient plants possessed stomata that were strikingly similar to those of extant land plants, such as *Arabidopsis thaliana* (Edwards et al., 1998; Taylor et al., 2005). In Figure 1-1, mature stomata and their neighboring epidermal cells over a large phylogenetic distance are presented to demonstrate the evolutionarily ancient derivation of the stoma.

If the evolution of stomata was indeed a single vital event in land plant evolution, what are the underlying genes and molecular mechanisms that gave rise to stomata? What are the molecular bases that explain the diversity of stomatal patterns seen today? Recent discoveries of key genes controlling stomatal development in the model plant *Arabidopsis* now provide molecular tools with which to tackle these questions. In this review, we first introduce *Arabidopsis* genes controlling stomatal development. We further describe the evolutionary conservation and uniqueness of stomatal development as well as the functions of orthologous genes in other plant species as an attempt to promote interest and discussion among the diverse disciplines of botany, paleobotany, and evolutionary biology.

Stomatal development in *Arabidopsis*

Cell-state transitions

Stomata in *Arabidopsis* develop through a process of asymmetric cell division (Figure 1-2) that produces the anisocytic stomatal complexes commonly seen in the Brassicaceae, in which each stoma has one smaller and two larger neighboring cells arising from asymmetric divisions within the stomatal lineage (Esau, 1977). First, in the protoderm, a population of cells initiates the stomatal lineage and divides asymmetrically to form a stomatal precursor cell called a meristemoid. The meristemoid reiterates asymmetric divisions several times, usually from one to three, creating sister cells called stomatal lineage ground cells (SLGCs) of decreasing size while renewing the meristemoid state (Figure 1-2). The meristemoid then differentiates into a guard mother cell (GMC). The GMC divides once symmetrically to form a pair of cells that develop the thickened walls and mature chloroplasts characteristic of stomatal guard cells (GCs). SLGCs eventually become pavement cells but may divide asymmetrically as well, producing a meristemoid away from the original stoma (Figure 1-2). The stomatal complex may be formed exclusively through multiple asymmetric divisions, in which case it is clonal, or if fewer asymmetric divisions occur the stoma may be adjacent to the lineage of another protodermal cell (Serna et al., 2002).

Two groups of basic helix-loop-helix (bHLH) transcription factors regulate the major cell-state transitions through stomatal differentiation (Figure 1-2). The first group is encoded by the paralogs (closely related genes arising from duplication within a genome) *SPEECHLESS* (*SPCH*), *MUTE*, and *FAMA*, which operate sequentially to regulate the cellular identity of each step in the developmental process (Ohashi-Ito and Bergmann, 2006; MacAlister et al., 2007; Pillitteri et al., 2007b). *SPCH* controls the initial asymmetric division with which protodermal cells enter the stomatal lineage; *spch*

mutants are entirely without stomatal lineage cells, having only puzzle-piece-shaped pavement cells interlocking over the entire shoot surface (MacAlister et al., 2007; Pillitteri et al., 2007b). Next to act is *MUTE*, which governs the meristemoid's cessation of asymmetric divisions and its transition to GMC identity. Plants without functional *MUTE* also develop no stomata, but rather than lacking the stomatal lineage entirely they show rosette-like patterns of inwardly spiraling meristemoid cell divisions with aborted meristemoids in the center (MacAlister et al., 2007; Pillitteri et al., 2007b, 2008). Lastly, *FAMA* appears to control the switch from the GMC to GC identity; in *fama* mutants, the stomatal precursors halt at the GMC state, resulting in "caterpillars" of excessively symmetrically divided cells instead of mature GCs from a single symmetric division (the many "false mouths" of the Roman goddess of rumor, Fama) (Ohashi-Ito and Bergmann, 2006).

Regulating the sequential action of the master regulators are the second group of bHLH proteins, encoded by the broadly expressed paralogs *SCRM* and *SCRM2*, which have partially overlapping functions. The originally discovered gain-of-function mutation, *scrm-D*, causes all epidermal cells to take on stomatal fate, while successive loss of *SCRM* and *SCRM2* creates phenotypes recapitulating *spch*, *mute*, and *fama* mutants (Kanaoka et al., 2008). The bHLH proteins function through dimerization (Massari and Murre, 2000), and SCRM and SCRM2 are capable of interacting with SPCH, MUTE, and FAMA in vivo (Figure 1-2) (Kanaoka et al., 2008). This suggests a possible mode of regulation in which transiently expressed stomatal-lineage-specific bHLH proteins heterodimerize with widely expressed bHLH partner proteins to control developmental progression (Pillitteri and Torii, 2007).

Other factors also act to promote stomatal development. Formation of satellite (secondary) meristemoids is controlled by microRNA *miR824*, which targets

AGAMOUS-LIKE16 (*AGL16*) mRNA for degradation (Kutter et al.,2007). An AGL16 mRNA that no longer matches miR824 due to silent mutations causes additional reentry of SLGC into the stomatal lineage, resulting in more satellite meristemoids, a phenotype opposite to that of the *agl16-1* dosage reduction mutant (Kutter et al., 2007).

Loss of function in a MYB transcription factor, *FOUR LIPS* (*FLP*), causes parallel pairs of stomata (Yang and Sack, 1995). Its paralog, At *MYB88*, is partially redundant with *FLP*: in *flp myb88* double mutants, massive *fama*-like stacks of cells are frequently formed, although some are able to differentiate terminally into GCs (Lai et al., 2005). It is therefore likely that *FLP* and *MYB88* together restrict the symmetric division of GMC, likely through their interactions with Cyclin A2 genes (reviewed in Pillitteri & Torii, 2012). Consistent with their redundant roles in GC differentiation, FLP and MYB88 share an amino acid substitution that is unique among known plant MYBs (Lai et al., 2005). The roles of *FLP* and *MYB88* echo the ongoing pattern in which a gene and its paralog have redundant or related functions, which is consistent with evolution through gene duplication.

Patterning and Signaling

A repressive cell–cell signaling pathway in *Arabidopsis* maintains proper stomatal patterning as defined by the one-cell spacing rule (Figure 1-2A). Genetic interactions among the partially redundant ERECTA (ER) family of leucine-rich-repeat receptor-like kinases (LRR-RLKs) and the LRR receptor-like protein TOO MANY MOUTHS (TMM) affect stomatal spacing and density (Shpak et al., 2005), and the ER-family has been shown to form homomers and also to associate with TMM (Lee et al., 2012). The function of TMM is organ-dependent and complex: in *tmm* mutants, leaves show clusters of stomata, but stems develop no stomata at all despite the entry of some cells

into the stomatal lineage (Bhave et al., 2009). Similarly, the three ER family RLKs exhibit combination-specific interactions with TMM in specific organ types, implying that different combinations of receptor complexes inhibit particular steps of stomatal differentiation.

Two cysteine-rich peptides, EPIDERMAL PATTERNING FACTOR1 (EPF1) (Hara et al., 2007) and EPF2 (Hara et al., 2009; Hunt and Gray, 2009) restrict the later and the earlier steps of stomatal cell fate specifications, respectively (Figure 1-2A). EPF2 signals through ERECTA, while EPF1's primary receptor is ERL1 (Lee et al., 2012). Additionally, EPF2 interacts with TMM directly, while EPF1 does not (Lee et al., 2012). Stomatal initiation and asymmetric division are thus regulated by receptor-ligand communication among epidermal cells.

The *EPF-LIKE (EPFL)* gene family contains an additional nine members (Hara et al., 2009), of which *STOMAGEN*/*EPFL9* promotes stomatal differentiation through *TMM* and the *ERECTA* family (Kondo et al., 2010; Sugano et al., 2010). *STOMAGEN* is not expressed in the epidermis, but rather in the mesophyll, the photosynthetic cell layer within the leaf (Kondo et al., 2010; Sugano et al., 2010); this signaling between cell layers coordinates the development of two tissues toward the organ's overarching goal of photosynthesis. *EPFL6*/*CHALLAH* affects stomatal patterning in the *tmm* mutant background or when overexpressed (Abrash and Bergmann, 2010; Uchida et al., 2012), probably due to its structural similarity to stomatal EPFs, but its endogenous function appears to be as an ERECTA ligand in stem elongation along with EPFL4 (Uchida et al., 2012). These additional players highlight the roles of the EPF family of signaling ligands in various aspects of development and imply an evolutionary relationship between the EPF ligand family and the ERECTA receptor family. Ongoing studies of ligand binding

and receptor associations will clarify the complex actions of TMM/ER family RLKs and how conflicting signals can be transmitted through the same receptors.

The receptor-ligand interactions appear to activate a mitogen-activated protein (MAP) kinase signaling cascade including the MAPKKK YODA (Bergmann et al., 2004), MPKK4/5, and MAPK MPK3/6 (Wang et al., 2007), to suppress stomatal development in neighboring cells. In general, MAP kinase cascades target transcription factors and regulate their activities via phosphorylation. SPCH has been shown to be a MAP kinase target, thus providing a mechanism by which stomatal suppression signaling might affect stomatal cell state identity directly (Lampard et al., 2008).

STOMATAL DENSITY AND DISTRIBUTION1 (SDD1) also acts as a negative regulator of stomatal formation (Berger and Altmann, 2000), but although it has been shown to require TMM for function (von Groll et al., 2002), it seems to be independent of EPF1 (Hara et al., 2007), EPF2 (Hara et al., 2009; Hunt and Gray, 2009), STOMAGEN/EPFL9 (Kondo et al., 2010; Sugano et al., 2010), and EPFL6/CHALLAH (Abrash and Bergmann, 2010). Although SDD1 encodes a membrane-associated subtilisin-type protease (Berger and Altmann, 2000), it appears unlikely that it directly processes any known member of the EPF family of stomatal ligands.

Innate Polarity

Asymmetric cell division within the stomatal lineage is highly polar: The meristemoid divides at angles of 120° from each previous division, forming an orderly inward spiral of SLGCs with a triangular meristemoid inside (Serna et al., 2002; Pillitteri et al., 2007b). When spacing divisions occur, the satellite meristemoids always arise away from any neighboring stoma to maintain the one-cell spacing rule. While cell–cell signaling and bHLH transcription factors determining the state of asymmetrically dividing stomatal precursors are well characterized, factors directing the innate polarity

that allows spacing to emerge from the asymmetric cell division process remained elusive until the discovery of a unique protein, BREAKING OF ASYMMETRY IN THE STOMATAL LINEAGE (BASL) (Dong et al., 2009). The BASL protein fused with green fluorescent protein (GFP) is localized to the nuclei of dividing stomatal lineage cells and also to the cell cortex opposite the site of the most recent asymmetric division, which appears to be BASL's main site of activity (Dong et al., 2009) (Figure 1-3). The ectopic overexpression of *BASL* confers abnormal cell polarization. Interestingly, the loss of asymmetry in a *basl* mutant resulted in adoption of the stomatal fate, as depicted by the accumulation of a *MUTEpro::nuclearGFP* reporter fusion protein in both daughter cells (Dong et al., 2009). This observation implies that the asymmetric divisions in stomatal lineage cells are necessary for specifying the fate of the SLGC, but not the meristemoid, and a rapid cell polarization via a BASL-mediated mechanism is required to do so. This *basl* phenotype is the opposite of *tmm er*, where both daughter cells of a meristemoid adopt the SLGC fate (Shpak et al., 2005). While genetic studies suggest that *BASL* operates in a separate pathway from *TMM* and the *ERECTA* family (Dong et al., 2009), the latter is required for proper positioning of GFP-BASL. Understanding the mechanism that achieves a peripheral localization of BASL may illuminate the intersection of signaling and innate asymmetry.

An excellent review of polarity in stomatal development, including cytological mechanisms, was published recently (Facette & Smith, 2012). Chapter 2 of this work also introduces an intriguing new player in stomatal cell polarity.

Stomatal Patterning and Differentiation in Grasses

In grasses, stomatal complexes develop from nonclonal cells through a polarized division process (Figure 1-4). The cells are organized in cell files throughout the plant,

which are visible on the macro scale as the typical parallel venation of the monocot (Figure 1-1F). Spacing is maintained by the specification of nonadjacent cell files to develop stomata, which begin with an asymmetric division toward the leaf apex for all cells in the file. One of the cells produced by the asymmetric division differentiates directly into a GMC, which before dividing again signals the neighboring cells in adjacent cell files to divide asymmetrically toward the GMC. These divisions produce subsidiary cells (SCs) important for stomatal action (Figure 1-4). Subsequently, the GMC makes a single symmetric division and differentiates two GCs in a distinctive dumbbell shape (Gallagher and Smith, 2000). GCs with this shape require less water to open than those with a kidney shape (Raschke, 1979), so they may be an adaptation to drought conditions.

The recent availability of genome sequences and reverse-genetics tools for model monocot plants, rice (*Oryza sativa*) and maize (*Zea mays*), allows for the examination of the conservation of genes important in stomatal patterning and differentiation. Grasses contain orthologs (genes in different species that evolved from a common ancestor) of the master regulatory genes of Arabidopsis stomatal differentiation, *SPCH*, *MUTE*, and *FAMA* (Liu et al., 2009). Although transcripts of the two *Os SPCH* genes were detected only in coleoptiles, one of them showed a decrease in entry into the stomatal lineage when mutated (Liu et al., 2009). This suggests that in rice, at least one of the *SPCH* genes controls initial cell lineage divisions; it may have additional functions that are redundant with the second *Os SPCH*, in which a mutation has not yet been found. *Os MUTE* and maize *Zm MUTE* were expressed earlier in development than their *Arabidopsis* ortholog, at the time of cell file specification. Ectopic overexpression of *Os MUTE* and *Zm MUTE* in *Arabidopsis* caused conversion of epidermal cells into stomata to varying degrees, a phenotype similar to that of *At MUTE* overexpression (Liu et al.,

2009), and they were able to partially rescue the *Arabidopsis mute* mutant when expressed under the native *Arabidopsis* promoter (Liu et al., 2009). In mutants of *Os FAMA*, GCs fail to differentiate into the dumbbell shape of wild-type GCs, implying that *Os FAMA* controls GC fate, though there are none of the extra cell divisions seen in *Arabidopsis fama* plants (Liu et al., 2009).

This series of studies provides interesting implications. First, the *SPCH-MUTE-FAMA* bHLH genes diverged before the evolutionary split between monocots and dicots. Second, they have maintained their overall functions as regulators of stomatal differentiation across these widely separated taxa. Third, the orthologs have differences in their specific functions that reflect the differences in the specifics of stomatal ontogeny between dicots (*Arabidopsis*) and monocots (grasses). For instance, while both *At fama* and *Os fama* mutants fail to differentiate mature GCs, only the *At fama* mutant forms characteristic tumors of supernumerary GMC-like cells (Ohashi-Ito and Bergmann, 2006; Liu et al., 2009). These differences could be reconciled if we assume that, in *Arabidopsis*, a transition from a proliferative state to terminal differentiation of stomata requires a strong brake to halt cell division machinery. There is no actively dividing meristemoid state during stomatal differentiation in grasses; consequently, *Os FAMA* may not have acquired (or retained) the specific function to restrict cell division. The role of *Os MUTE* earlier in development, relative to *At MUTE*, may also suggest that these bHLH proteins acquired specific functions reflecting the uniqueness of stomatal development in grasses.

Formation of stomatal complexes in grasses (Figure 1-4) involves asymmetric cell divisions in subsidiary mother cells (SMCs), members of GMC-neighboring cell files that give rise to SCs, which have a highly specialized function as ion reserves for GC movement. Such an asymmetric cell division outside the stomatal lineage does not exist

in *Arabidopsis*. Therefore, it is of great interest to unravel the molecular identities of the genes controlling subsidiary cell division and polarity in grasses. Loss-of-function mutations in *PANGLOSS1* (*PAN1*) lead to aberrant asymmetric cell division and SC patterning (Gallagher and Smith, 2000). *PAN1* was cloned and shown to encode an atypical group III LRR-RLK without kinase activity (Cartwright et al., 2009). It follows that PAN1 might associate with other RLKs with functional kinase domain to transmit signals. In *Arabidopsis*, an atypical LRR-RLK without kinase activity is known to regulate developmental processes (Chevalier et al., 2005). Additionally, PAN1 associates with and is necessary for polar localization of ROP (Rho of plants) GTPases (Humphries et al., 2011), and ROP2 and 9 were shown to be required for SMC polarization (Humphries et al., 2011).

Interestingly, PAN1 protein is localized specifically at the plasma membrane of the SMC at the sites of contact with GMC. This is consistent with the hypothesis that PAN1 receives GMC signals to regulate the site of cytokinesis. Such polar localization has not been observed for the *Arabidopsis* stomatal LRR receptor TMM (Nadeau and Sack, 2002), and no member of LRR group III in *Arabidopsis* has been found to affect stomatal development. Thus, PAN1's function may highlight the unique regulation of stomatal patterning in grasses.

More recently, *PANGLOSS2 (PAN2)* has been found to have similar function to *PAN1* in SMC polarity. Maize PAN2 is required for proper subcellular localization of PAN1, placing it upstream in the SMC polarization process (Cartwright et al., 2009). Surprisingly, *PAN2* has been identified as an unrelated LRR-RLK that also lacks kinase activity (Zhang et al., 2012). Although it can bind to itself, PAN2 does not associate with PAN1, leaving two hypothesized roles for an interacting RLK with functional kinase domain still open.

Diversity of Stomata and Stomatal Development

In extant land plant species, stomata appear on the sporophyte shoot epidermis of nearly every lineage. For most plants, stomata are evenly distributed in the shoot areas where they appear, and they adhere closely to the one-cell spacing rule (Sachs, 1991), allowing them to open and close effectively. Therefore, the diversity of stomata consists of variations on a theme: stomata are spaced at least one cell apart, but the specifics differ. Classic anatomical studies suggest that differences in stomatal patterning reflect the types, positions, and numbers of cell divisions that give rise to the stomatal complex (Esau, 1977; Payne, 1979). In this section, we describe stomatal development in other plant taxa and speculate regarding their possible alterations of regulatory mechanisms known in *Arabidopsis*.

In mosses and lycophytes (Figures 1-1A and B), stomata develop through a simple process of a single asymmetric division followed by differentiation of a GMC, then GCs (Payne, 1979). We could speculate that the ancient function of SPCH-MUTE-FAMA family proteins may be to initiate the stomatal cell lineage and differentiate into GMC directly. In ferns, MMC in a cell file undergo one, or sometimes two, asymmetric divisions toward the leaf apex, then differentiate into GMC and GC (Apostolakos et al., 1997) (Figure 1-5A). This implies cell polarity mechanisms different from those of *Arabidopsis*.

In basal angiosperms, such as *Victoria* (Figure 1-1D), stomata appear to arise without amplifying divisions. (For terminology and evolution of stomatal complexes among basal angiosperms, see Carpenter [2005].) By contrast, the magnoliid *Houttuynia* shows a clear and extended inward spiral of cells around each stoma, which we hypothesize arise through an extended process of asymmetric division resulting from delayed *MUTE* activity (Figure 1-5B).

Gardenia and *Begonia* (Figures 1-1G and H) show some of the diversity found within eudicots, a category that also includes *Arabidopsis* (Figure 1-2). In *Gardenia*, it appears that most, if not all, protodermal cells enter the stomatal lineage, implying that the unknown mechanism by which MMCs arise operates at higher density than in *Arabidopsis*. In *Begonia*, small groups of stomata are separated by pavement cells. Figure 1-5C shows a hypothetical developmental process giving rise to such a stomatal complex: Instead of dividing asymmetrically, the MMC divides symmetrically one or more times and all daughter cells retain MMC identity. Asymmetric division then proceeds as in *Arabidopsis*, allowing the one-cell spacing rule to be maintained, and the mature *Begonia* epidermis has groups of up to four stomata (Figure 1-5C).

Phylogenetic perspective on stomatal development

The recent availability of diverse land plant genomes, such as the moss *Physcomitrella patens* and model crop monocot plants rice and maize, allows for examination of the origin and conservation of genes important in stomatal patterning and spacing. Since stomata evolved only once in the ancestor of the mosses (Raven, 2002), it is expected that plants that arose from different basal land plant lineages, such as *P. patens*, would also possess some if not all of the genes found to be essential for stomatal patterning and differentiation in *Arabidopsis*. Stomata in moss form through a single asymmetric cell division to produce a GMC, which divides partially or completely to produce a single GC or a pair of GCs surrounding a pore (Payne, 1979). Congruent with this relatively simplified developmental sequence to mature GC formation, the moss *P. patens* does not have genes encoding *SPCH* or *MUTE*. *P. patens*, however, has two *FAMA*-like genes (Figures 1-6A and 1-S1), which likely function in moss as they do in *Arabidopsis*, rice, and maize: to control the final differentiation step to

GCs. Based on our *SPCH-MUTE-FAMA* gene tree and taxonomic sample, it is not clear whether there was only a single stomatal master regulatory bHLH gene in the ancestor of land plants or if *SPCH* and *MUTE* have been subsequently lost in *P. patens*. It is fascinating, however, to speculate that a single stomatal bHLH protein in early land plants may have had a function to create stomatal-lineage cells and immediately trigger GC differentiation. Later, as this gene duplicated, its roles may have subfunctionalized and even acquired new functions to mediate amplifying and spacing asymmetric cell divisions. The evolution of the leaf blade (Figure 1-1) in vascular plants perhaps benefited from additional mechanisms to space stomata correctly during laminar leaf expansion, such as amplifying divisions of meristemoids and recruiting cell–cell signaling components to impinge directly on the stomatal differentiation process. In this regard, it is interesting that *At SPCH* possesses a specific MAP kinase target domain not found in *At FAMA* (Lampard et al., 2008).

Land plants also share common signaling components for the spacing of stomata. TMM, a member of the LRR-RLP family, functions as part of a signaling complex that regulates the spacing of stomata in *Arabidopsis* (Nadeau and Sack, 2002). *TMM* is present in a single copy in poplar, grass, and moss genomes, which may reflect its integral role in enforcing correct stomatal spacing (Figures 1-6B and 1-S2). In *Arabidopsis*, TMM is proposed to associate with ERECTA family receptors, which are Ser/Thr kinases (Shpak et al., 2005). TMM is specific to the stomatal lineage, whereas ERECTA family members play additional roles in various developmental processes in *Arabidopsis* (Shpak et al., 2004; Pillitteri et al., 2007a; Hord et al., 2008; van Zanten et al., 2009, 2010; Uchida et al., 2012). Thus, TMM may be necessary for recognition of ligands or receptor complex activation specific to the stomatal lineage in all land plants.

YODA is a MAPKKK thought to act downstream of TMM and is present in multiple copies in rice and poplar and in single copy in *Arabidopsis* and *P. patens* (Figure 1-6C). Although MAPKKKs comprise a large gene family, *YODA* can be identified by its unique N-terminal Ser-rich regulatory region (Figure 1-S3). Closely related MKKKs in *Arabidopsis* appear to be lacking this region (Figure 1-S3), and if it is deleted in YODA, the protein becomes constitutively active, resulting in a stomata-less phenotype (Lukowitz et al., 2004). Given the high level of sequence conservation in the region of this protein in *P. patens* relative to *Arabidopsis*, we predict that the N-terminal region regulates its YODA activity as well. Like *ERECTA* family RLKs, *YODA* regulates multiple developmental processes in *Arabidopsis* in addition to stomatal patterning, such as embryo polarity/patterning, inflorescence elongation, and flower development (Lukowitz et al., 2004). Therefore, it would be interesting to see whether *YODA* might also be an essential component in the regulation of stomatal spacing in moss.

Although close anatomical descriptions of stomatal development exist for many taxa, little is known of the molecular and evolutionary basis of stomatal formation across land plants. Valuable insights into ancestral and conserved mechanisms of stomatal development will no doubt be gained from further studies on basal land plants. In the moss sporophyte, stomatal development is relatively more similar to that of grasses than that of eudicots, in that one asymmetric division gives rise to a GMC directly, which then divides symmetrically once to form paired GCs (Ziegler, 1987). However, unlike grasses, there are no subsidiary cells formed from outside the stomatal lineage. To reach a deeper understanding of novel mechanisms responsible for generating stomatal diversity (Figure 1-1), comparative studies of closely related, yet morphologically divergent, stomatal phenotypes are needed.

Future Perspectives

With recent advances in understanding the molecular processes controlling stomatal development in the model eudicot *Arabidopsis*, alongside the completion of several genome projects in monocot crop plants and in basal land plants, we are beginning to find that many of the molecular mechanisms that control stomatal development in *Arabidopsis* are part of an ancient machinery. As shown by *PAN1* in maize and *BASL* in *Arabidopsis*, some of the unique aspects of stomatal development relate to the type and number of asymmetric cell divisions associated with the formation of the stomatal complex. Further identification of orthologs of stomatal regulator genes in phylogenetically and evolutionarily important species and analysis of their developmental functions may unravel the molecular basis of the evolution of stomatal development.

Several exciting questions remain: For instance, how is the stomatal lineage first established? What initiates *SPCH* expression, and what are the cellular dynamics that set up the entry into asymmetric cell divisions of the stomatal lineage? When did amplifying cell divisions come to play a key role in stomatal development, and is that innovation correlated with specific features of *MUTE*?

A second line of questions revolves around the intersection of environmental signaling and stomatal development. Some recent work is beginning to explain how light signals integrate with the stomatal development program (Boccalandro et al., 2009; Casson et al., 2009; Kang et al., 2009), while other work is examining how CO_2 regulates stomatal density (Gray et al., 2000). In addition, while a direct role between temperature and stomatal patterning has not been determined, *SCRM*, which is also known as *ICE1*, is a known transcriptional activator of cold response genes (Chinnusamy et al., 2003). During the evolution of land plants, when and how did such network connections of

environmental signaling pathways to stomatal development emerge? Clearly, we now have the molecular-genetic tools in hand to explore the evolution of this landmark developmental innovation.

Introduction to Thesis Research

My major interest in stomatal development is the organization of cells into functional tissues. Multicellularity requires multiple levels of integration and communication among individual cells to create complex structures that fulfill functional and structural objectives for the whole organism. Stomata provide a complex yet accessible system that includes both self-organization within the epidermis and external signals from other tissue layers. In *Arabidopsis*, the stomatal meristemoid has a stem-cell-like character that contributes to the structure of the leaf by allowing laminar expansion and also creates evenly spaced stomata to allow the organ to function correctly. The epidermis is only one of the tissues involved in leaf formation, but it must develop perfectly in coordination with the other cell layers to achieve optimal performance and organismal fitness. We studied the meristemoid cell state using a combination of mutant alleles that greatly enriched different cell states in the epidermis, and the results are presented in Chapter 2.

In Chapter 3, one of the genes discovered through that meristemoid transcriptome analysis is explored in greater detail. *HDG2* encodes a transcription factor highly expressed in meristemoids, which when ectopically expressed in nonepidermal cell layers leads to the development of mature, perfectly formed stomata inside the leaf. This underscores the importance of tissue-level identity to produce correctly formed tissues and organs; the cells diverted to the stomatal pathway would have become functional mesophyll cells rather than ectopic stomata which neither contribute to

mesophyll function nor promote gas exchange due to their position, thus reducing organismal fitness. Tissue organization is critical and requires a hierarchy of genetic regulation; when that organization is disrupted, entire genetic programs may be incorrectly activated to produce misplaced cell types, which underscores the importance of appropriate cell-fate regulation for the formation of useful tissues and organs.

Returning to focus more directly on the endogenous stomatal pathway, I developed a time-lapse protocol to evaluate the activity of stomatal master regulators in the germinating cotyledon epidermis (Peterson & Torii, 2012; see Appendix B). Preliminary results of these investigations are presented in Chapter 4. One promising avenue in this research is my investigation of cell fate using laser ablation of stomatal precursors during development. The time lapse method holds the germinating plant in place and allows observation of protein dynamics, so immediate effects of cell manipulation may be monitored. Statistical significance has not yet been reached, but there are indications that ablating stomatal precursors can affect the fates of neighboring stomatal-lineage cells.

My investigations into germinating cotyledons led me to realize that the sequence of stomatal development there was abbreviated, and I fortunately entered into a collaboration with Professor Tetsuya Higashiyama's ERATO Live-Holonics group at Nagoya University which allowed me to investigate stomatal development during embryogenesis. Preliminary results from this project are presented in Chapter 5. Embryo imaging holds great promise as a system for easily observing de novo stomatal pattern formation at a small, tractable scale, allowing investigation of the effects of signaling and regulatory programs on tissue organization.

Acknowledgements

We thank Laurie Smith (University of California, San Diego) for the images of maize wild-type and *pan1* epidermis (Figure 1-4B) and Lynn Pillitteri for the scanning electron microscopy image of *P. patens* stomata (Figure 1-1A). Our research is supported by grants from the National Science Foundation (IOB-0744892 and MCB-0855659), by the U.S. Department of Energy (DE-FG02-03ER15448), and by PREST, Japan Science and Technology Agency, to K.U.T. K.M.P. is a National Science Foundation Graduate Research Fellow.

Methods

Bioinformatic/Phylogenetic Methods

BLAST searches using *Arabidopsis* amino acid sequences were performed using GRAMENE, *Populus trichocarpa* v1.1, and *Physcomitrella patens* v1.1 genome websites to identify putative orthologs for *SPCH, MUTE, FAMA, TMM*, and *YODA*. All high-scoring matching genes were used in initial phylogenetic analyses to identify true orthologs. Non-orthologous genes were then discarded, and the dataset was realigned and analyzed as described below. Outgroup sequences used for the analyses were identified in previously published comprehensive analyses of *Arabidopsis* bHLH and receptor-like protein gene families (Toledo-Ortiz, 2003; Wang et al. 2008). Identification of the putative N-terminal regulatory region of *Physcomitrella patens YODA* was first performed using BLAST searches, and then additional unannotated sequences were found by identifying open reading frames towards the N-terminal end.

Amino acid sequences for the bHLH and ACT domains from *SPCH, MUTE, FAMA*, and their orthologs were aligned with CLUSTAL X (Jeanmougin et al., 1998). The full-length amino acid sequences from *TMM* and its orthologs were also aligned

with CLUSTAL X. Default CLUSTAL X alignment parameters were used, which include pairwise alignment gap opening penalty of 10 and extension of 0.1, multiple alignment gap opening penalty of 10 and extension of 0.2, and a gonnet series matrix. Neighbor-joining trees were constructed using Phylip 3.68 (Felsenstein, 2005) (Dayhoff PAM matrix) and nodal support was assessed using 100,000 bootstrap replicates. Tree topologies were also computed using Bayesian methods (MrBayes 3.1.2) (Ronquist and Huelsenbeck, 2003) based on a mixed model of amino acid substitution and also allowed for invariant sites and rate heterogeneity using the gamma distribution. For the *TMM* tree, the analysis was run for 435,500 generations in 2 separate analyses, and the first 1000 generations were discarded as burn-in. The *SPCH-MUTE-FAMA* analysis was run for 1,000,000 generations in 2 separate analyses and the first 600 generations were discarded as burn-in. To ensure adequate mixing had occurred in the Bayesian analyses, each of the two separate analyses was analyzed separately to compare clade recovery and posterior probability support values. For both the *TMM* and *SPCH-MUTE-FAMA* analyses, independent run tree topologies were identical, and posterior probability values varied by no more than 2%. Trees shown in Figure 1-6 represent the two separate runs combined. Amino-acid alignments are provided in Figures 1-S1 through 1-S3, and gene IDs are listed in Table 1-S1.

References

Abrash, E.B., and Bergmann, D.C. (2010). Regional specification of stomatal production by the putative ligand CHALLAH. *Development* 137: 447–455.

Apostolakos, P., Panteris, E., and Galatis, B. (1997). Microtubule and actin filament organization during stomatal morphogenesis in the fern *Asplenium nidus*. 1. Guard cell mother cell. *Protoplasma* 198: 93–106.

Berger, D., and Altmann, T. (2000). A subtilisin-like serine protease involved in the regulation of stomatal density and distribution in *Arabidopsis thaliana*. *Genes Dev*. 14: 1119–1131.

Bergmann, D.C., Lukowitz, W., and Somerville, C.R. (2004). Stomatal development and pattern controlled by a MAPKK kinase. *Science* 304: 1494–1497.
Bhave, N.S., Veley, K.M., Nadeau, J.A., Lucas, J.R., Bhave, S.L., and Sack, F.D. (2009). *TOO MANY MOUTHS* promotes cell fate progression in stomatal development of *Arabidopsis* stems. *Planta* 229: 357–367.
Boccalandro, H.E., Rugnone, M.L., Moreno, J.E., Ploschuk, E.L., Serna, L., Yanovsky, M.J., and Casal, J.J. (2009). Phytochrome B enhances photosynthesis at the expense of water-use efficiency in *Arabidopsis*. *Plant Physiol*. 150: 1083–1092.
Carpenter, K.J. (2005). Stomatal architecture and evolution in basal angiosperms. *Am. J. Bot*. 92: 1595–1615.
Cartwright, H.N., Humphries, J.A., and Smith, L.G. (2009). PAN1: A receptor-like protein that promotes polarization of an asymmetric cell division in maize. *Science* 323: 649–651.
Casson, S.A., Franklin, K.A., Gray, J.E., Grierson, C.S., Whitelam, G.C., and Hetherington, A.M. (2009). Phytochrome B and PIF4 regulate stomatal development in response to light quantity. *Curr. Biol*. 19: 229–234.
Chevalier, D., Batoux, M., Fulton, L., Pfister, K., Yadav, R.K., Schellenberg, M., and Schneitz, K. (2005). *STRUBBELIG* defines a receptor kinase-mediated signaling pathway regulating organ development in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* 102: 9074–9079.
Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B.H., Hong, X., Agarwal, M., and Zhu, J.K. (2003). *ICE1*: A regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. *Genes Dev*. 17: 1043–1054.
Dong, J., MacAlister, C.A., and Bergmann, D.C. (2009). *BASL* controls asymmetric cell division in *Arabidopsis*. *Cell* 137: 1320–1330.
Edwards, D., Kerp, H., and Haas, H. (1998). Stomata in early land plants: An anatomical and ecophysiological approach. *J. Exp. Bot*. 49: 255–278.
Esau, K. (1977). Stomata. In *Anatomy of Seed Plants* (New York: John Wiley & Sons), pp. 88–99.
Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package) (Department of Genome Sciences, University of Washington, Seattle: Distributed by the author).
Freeman, S. (2008). Green plants. In *Biological Sciences*, 3rd ed (San Francisco, CA: Pearson Education), pp. 626–663.
Gallagher, K., and Smith, L.G. (2000). Roles for polarity and nuclear determinants in specifying daughter cell fates after an asymmetric cell division in the maize leaf. *Curr. Biol*. 10: 1229–1232.
Gray, J.E., Holroyd, G.H., van der Lee, F.M., Bahrami, A.R., Sijmons, P.C., Woodward, F.I., Schuch, W., and Hetherington, A.M. (2000). The HIC signalling pathway links CO2 perception to stomatal development. *Nature* 408: 713–716.
Hara, K., Kajita, R., Torii, K.U., Bergmann, D.C., 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, K.M., Torii, K.U., and Kakimoto, T. (2009). Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR2, in *Arabidopsis* leaves. *Plant Cell Physiol*. 50: 1019–1031.
Hord, C.L.H., Sun, Y.J., Pillitteri, L.J., Torii, K.U., Wang, H., Zhang, S., and Ma, H. (2008). Regulation of *Arabidopsis* early anther development by the mitogen-

activated protein kinases, MPK3 and MPK6, and the ERECTA and related receptor-like kinases. *Mol. Plant* 1: 645–658.
Humphries, J.A., Vejlupkova, Z., Luo, A., Meeley, R.B., Sylvester, A.W., Fowler, J.E., and Smith, L.G. (2011). ROP GTPases act with the receptor-like protein PAN1 to polarize asymmetric cell division in maize. *Plant Cell* 23(6): 2273-2284.
Hunt, L., and Gray, J.E. (2009). The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Curr. Biol.* 19: 864–869.
Ichimura, K., Shinozaki, K., Tena, G., Sheen, J., Henry, Y., Champion, A., Kreis, M., Zhang, S., Hirt, H., Wilson, C., Heberle-Bors, E., Ellis, B.E., Morris, P.C., Innes, R.W., Ecker, J.R., Scheel, D., Klessig, D.F., Machida, Y., Mundy, J., Ohashi, Y., and Walker, J.C. (2002). Mitogen-activated protein kinase cascades in plants: a new nomenclature. *Trends in Plant Science* 7: 301-308.
Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., and Gibson, T.J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences* 23: 403–405.
Kanaoka, M.M., Pillitteri, L.J., Fujii, H., Yoshida, Y., Bogenschutz, N.L., Takabayashi, J., Zhu, J.K., and Torii, K.U. (2008). *SCREAM/ICE1* and *SCREAM2* specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* 20: 1775–1785.
Kang, C.Y., Lian, H.L., Wang, F.F., Huang, J.R., and Yagn, H.Q. (2009). Cryptochromes, phytochromes, and COP1 regulate light-controlled stomatal development in *Arabidopsis*. *Plant Cell* 21: 2624–2641.
Kondo, T., Kajita, R., Miyazaki, A., Hokoyama, M., Nakamura-Miura, T., Mizuno, S., Masuda, Y., Irie, K., Tanaka, Y., Takada, S., Kakimoto, T., and Sakagami, Y. (2010). Stomatal density is controlled by a mesophyll-derived signaling molecule. *Plant Cell Physiol.* 51: 1–8.
Kutter, C., Schob, H., Stadler, M., Meins, F., Jr., and Si-Ammour, A. (2007). MicroRNA-mediated regulation of stomatal development in *Arabidopsis*. *Plant Cell* 19: 2417–2429.
Lai, L.B., Nadeau, J.A., Lucas, J., Lee, E.K., Nakagawa, T., Zhao, L.,Geisler, M., and Sack, F.D. (2005). The *Arabidopsis* R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. *Plant Cell* 17: 2754–2767.
Lampard, G.R., Macalister, C.A., and Bergmann, D.C. (2008). *Arabidopsis* stomatal initiation is controlled by MAPK-mediated regulation of the bHLH SPEECHLESS. *Science* 322: 1113–1116.
Lee, J.S., Kuroha, T., Hnilova, M., Khatayevich, D., Kanaoka, M.M., McAbee, J.M., Sarikaya, M., Tamerler, C., and Torii, K.U. (2012) Direct interaction of ligand-receptor pairs specifying stomatal patterning. *Genes & Development* 26: 126-136.
Liu, T., Ohashi-Ito, K., and Bergmann, D.C. (2009). Orthologs of *Arabidopsis thaliana* stomatal bHLH genes and regulation of stomatal development in grasses. *Development* 136: 2265–2276.
Lukowitz, W., Roeder, A., Parmenter, D., and Somerville, C. (2004). A MAPKK kinase gene regulates extra-embryonic cell fate in *Arabidopsis*. *Cell* 116: 109–119.
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.
Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* 20: 429–440.

Nadeau, J.A., and Sack, F.D. (2002). Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science* 296: 1697–1700.
Ohashi-Ito, K., and Bergmann, D.C. (2006). *Arabidopsis FAMA* controls the final proliferation/differentiation switch during stomatal development. *Plant Cell* 18: 2493–2505.
Payne, W.W. (1979). Stomatal patterns in embryophytes - Their evolution, ontogeny and interpretation. *Taxon* 28: 117–132.
Peterson, K. M., and Torii, K. U. (2012). Long-term, high-resolution confocal time lapse imaging of *Arabidopsis* cotyledon epidermis during germination. *Journal of Visualized Experiments* (70): e4426.
Pillitteri, L.J., Bemis, S.M., Shpak, E.D., and Torii, K.U. (2007a). Haploinsufficiency after successive loss of signaling reveals a role for *ERECTA*-family genes in ovule development. *Development* 134: 3099–3109.
Pillitteri, L.J., Bogenschutz, N.L., and Torii, K.U. (2008). The bHLH protein MUTE controls differentiation of stomata and the hydathode pore in *Arabidopsis*. *Plant Cell Physiol.* 49: 934–943.
Pillitteri, L.J., Sloan, D.B., Bogenschutz, N.L., and Torii, K.U. (2007b). Termination of asymmetric cell division and differentiation of stomata. *Nature* 445: 501–505.
Pillitteri, L.J., and Torii, K.U. (2007). Breaking the silence: Three bHLH proteins direct cell-fate decisions during stomatal development. *Bioessays* 29: 861–870.
Raschke, K. (1979). Movements of stomata. In *Physiology of Movements, Encyclopedia of Plant Physiology*, Vol. 7, W. Haupt and M.E. Feinleib, eds (Berlin: Springer-Verlag), pp. 383–441.
Raven, J. (2002). Selection pressures on stomatal evolution. *New Phytol.* 153: 371–386.
Ronquist, F., and Huelsenbeck, J.P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572- 1574.
Sachs, T. (1991). Stomata as an example of meristemoid development. In *Pattern Formation in Plant Tissues* (Cambridge, UK: Cambridge University Press), pp. 101–117.
Serna, L., Torres-Contreras, J., and Fenoll, C. (2002). Clonal analysis of stomatal development and patterning in *Arabidopsis* leaves. *Dev. Biol.* 241: 24–33.
Shpak, E.D., Berthiaume, C.T., Hill, E.J., and Torii, K.U. (2004). Synergistic interaction of three ERECTA-family receptor-like kinases controls *Arabidopsis* organ growth and flower development by promoting cell proliferation. *Development* 131: 1491–1501.
Shpak, E.D., McAbee, J.M., Pillitteri, L.J., and Torii, K.U. (2005). Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science* 309: 290–293.
Sugano, S.S., 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.
Taylor, T.N., Kerp, H., and Hass, H. (2005). Life history biology of early land plants: Deciphering the gametophyte phase. *Proc. Natl. Acad. Sci. USA* 102: 5892–5897.
Toledo-Ortiz, G., Huq, E., and Quail, P.H. (2003). The *Arabidopsis* basic/helix-loop-helix transcription factor family. *Plant Cell* 15, 1749-1770.
van Zanten, M., Basten Snoek, L., van Eck-Stouten, E., Proveniers, M.C., Torii, K.U., Voesenek, L.A., Peeters, A.J., and Millenaar, F.F. (2010). Ethylene-induced hyponastic growth in *Arabidopsis thaliana* is controlled by *ERECTA*. *Plant J.* 61: 83–95.

van Zanten, M., Snoek, L.B., Proveniers, M.C., and Peeters, A.J. (2009). The many functions of *ERECTA*. *Trends Plant Sci.* 14: 214–218.
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: 1527–1539.
Wang, G., Ellendorff, U., Kemp, B., Mansfield, J.W., Forsyth, A., Mitchell, K., Bastas, K., Liu, C.-M., Woods-Tor, A., Zipfel, C., de Wit, P.J.G.M., Jones, J.D.G., Tor, M., and Thomma, B.P.H.J. (2008). A genome-wide functional investigation into the roles of receptor-like proteins in *Arabidopsis*. *Plant Physiol.* 147**,** 503-517.
Wang, H., Ngwenyama, N., Liu, Y., Walker, J., and Zhang, S. (2007). Stomatal development and patterning are regulated by environmentally responsive mitogen-activated protein kinases in *Arabidopsis*. *Plant Cell* 19: 63–73.
Yang, M., and Sack, F.D. (1995). The *too many mouths* and *four lips* mutations affect stomatal production in *Arabidopsis*. *Plant Cell* 7: 2227–2239.
Ziegler, H. (1987). The evolution of stomata. In *Stomatal Function*, E. Zeiger, G.D. Farquhar, and I.R. Cowan, eds. (Stanford, CA: Stanford University Press), pp. 29–57.

Figure 1-1.

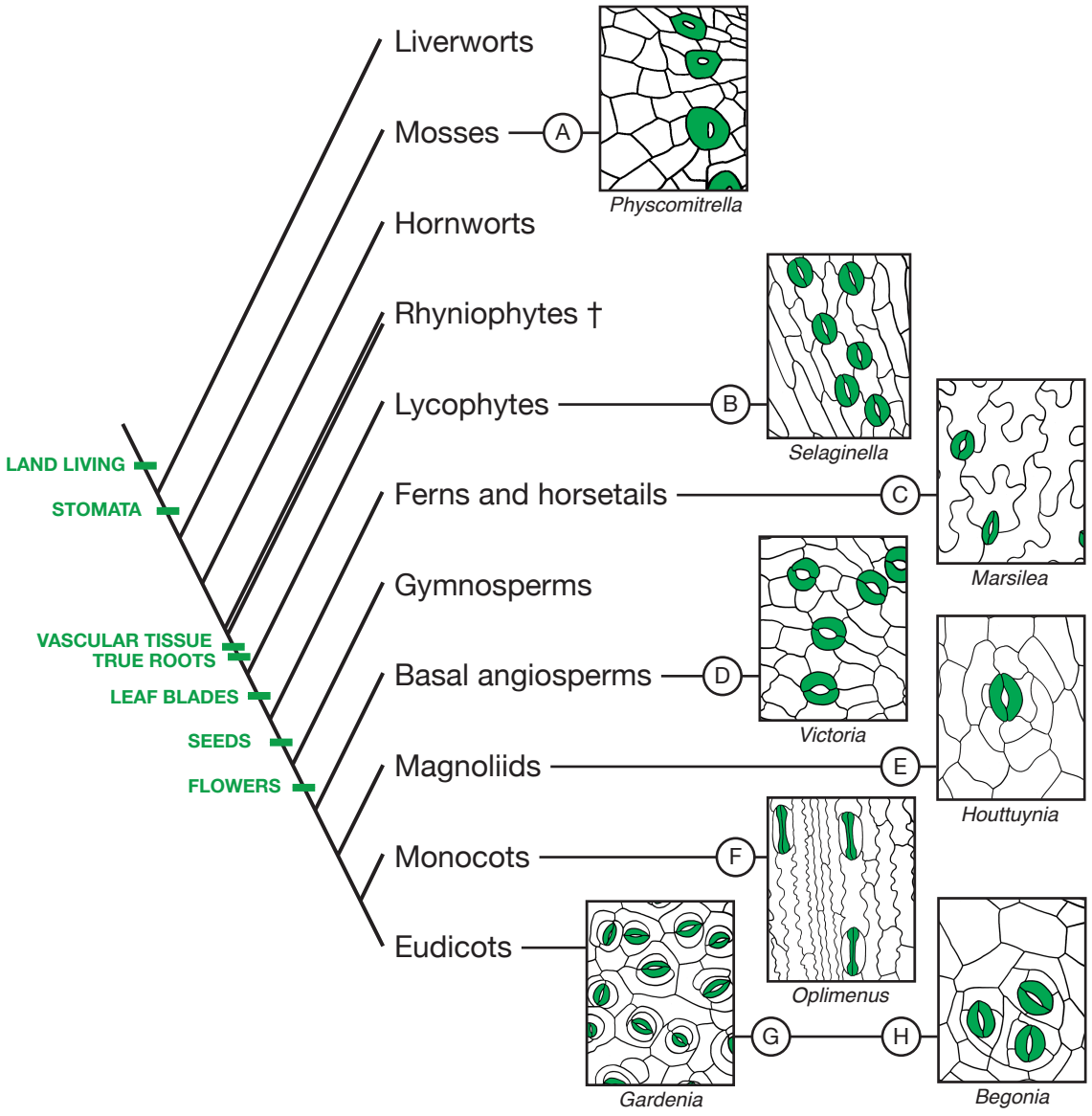


Figure 1-1. Diversity of stomata across land plant taxa.
A phylogenetic tree of extant and extinct (†) land plants includes evolutionary traits supporting success on land. The wide diversity of stomatal complexes among sporophytes in these groups is represented by epidermal tracings of *Physcomitrella patens* (A; nonvascular), *Selaginella kraussiana* (B; lycophyte, vascular), *Marsilea macropoda* (C; fern, vascular), *Victoria amazonica* (D; Nymphaeaceae, basal angiosperm), *Houttuynia cordata* (E; Piperales, magnoliid), *Oplimenus hirtellus* (F; Poales, monocot grass), *Gardenia taitensis* (G; Gentianales, eudicot angiosperm), and *Begonia rex-cultorum* 'Roberta' (H; Cucurbitales, eudicot angiosperm). Stomata are colored green. Note that the stomata of *Physcomitrella* have a single GC, while the GCs of *Oplimenus*, a grass, have a dumbbell shape. Only *Houttuynia*, *Gardenia*, and *Begonia* show evidence of asymmetric amplifying divisions within the stomatal lineage. *Physcomitrella* was traced from a scanning electron microscopy image by L. Pillitteri.

Figure 1-2.

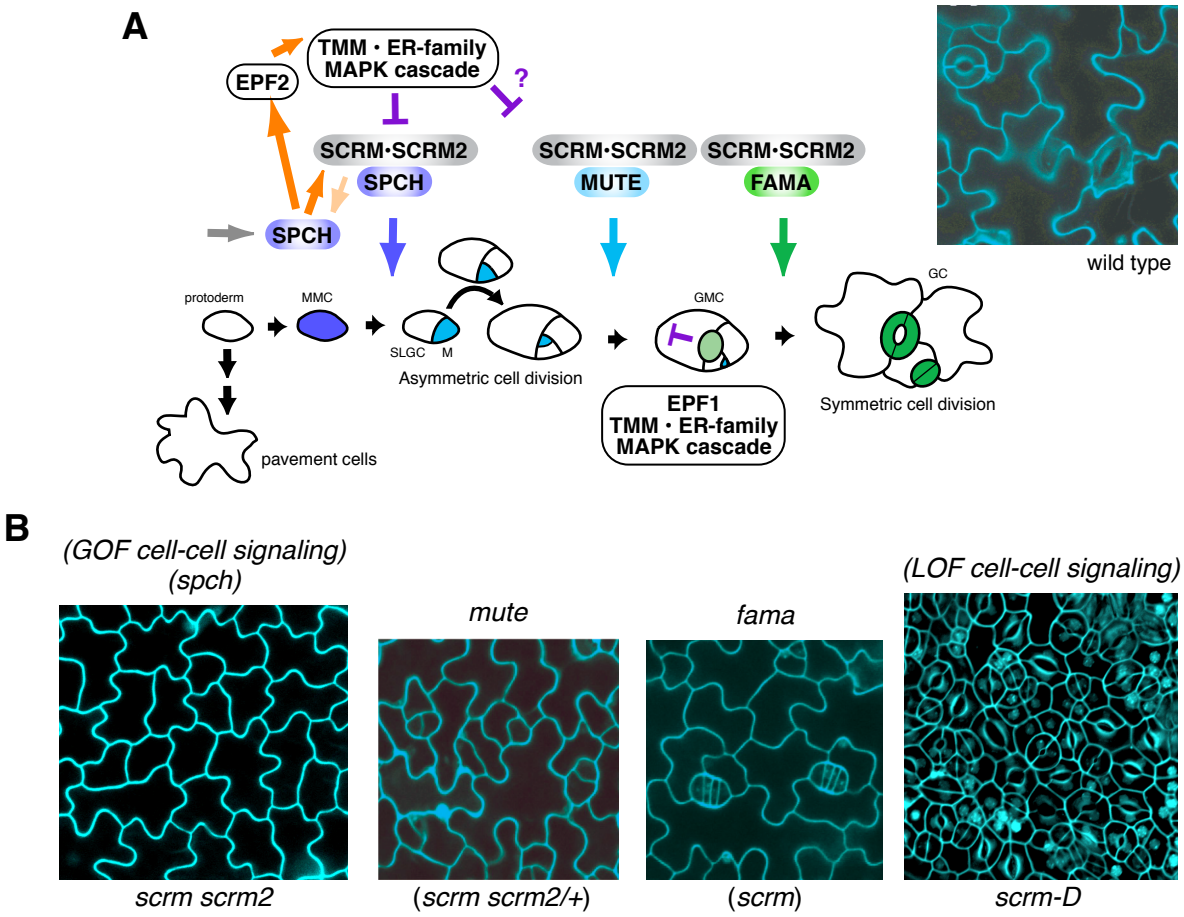


Figure 1-2. Stomatal development in *Arabidopsis*.
(A) Schematic diagram of stomatal development. The cell states of stomatal precursors are driven by three paralogous bHLH transcription factors, which likely dimerize with SCRM and SCRM2 as a mechanism for coordinated action. Initial specification of the stomatal cell lineage, in which a protodermal cell becomes a meristemoid mother cell (MMC), is controlled by *SPCH*. Protodermal cells not entering the stomatal lineage differentiate into pavement cells. The MMC divides asymmetrically to form a meristemoid (M) and SLGC and may reiterate similar divisions several times. *MUTE* controls the cell-state transition from M to GMC, and *FAMA* is required for correct division of the GMC into GCs forming a functional stoma. It is proposed that a MAP kinase signaling cascade following putative ligands EPF1 and EPF2 (EPF1 expressed in GMC, light green, and EPF2 expressed in MMC, blue, and M, cyan) perceived by TMM and the ER family of RLKs acts to suppress stomatal identity in cells adjacent to developing stomata; new meristemoids can differentiate at least one cell away, as shown near GMC. An image of wild-type epidermis is shown at the top right.
(B) Epidermal phenotypes of stomatal differentiation mutants. Shown are the rosette leaf epidermis of (from left) *scrm scrm2*, *mute*, *fama*, and *scrm-D*. *scrm scrm2* produces epidermis solely composed of pavement cells, a phenotype identical to that of spch as well as gain-of-function mutants in stomatal cell–cell signaling genes. *mute* and *fama* produce epidermis with arrested stomatal precursor cells similar to *scrm scrm2*/+ and *scrm*, respectively. *scrm-D* produces epidermis solely composed of stomata, a phenotype similar to loss of function in stomatal signaling genes. Images are reproduced from Kanaoka et al. (2008).

Figure 1-3.

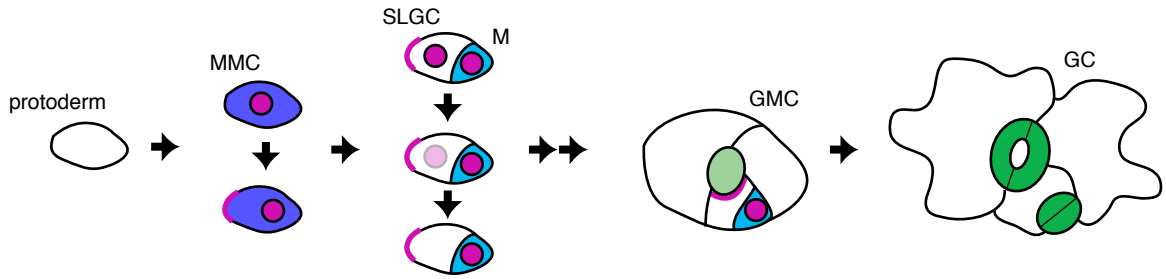


Figure 1-3. Localization of *BASL*.
BASL acts as a molecular signal instructing stomatal lineage cells to divide away from it. BASL protein appears initially in the nuclei of meristemoid mother cells (MMC), which differentiate from protodermal cells. The protein then localizes in a second location at the cell periphery opposite the site of the future asymmetric division. Following that division, BASL remains at the cell periphery but fades away from the nucleus of the larger daughter cell (SLGC), which loses stomatal lineage identity; it remains in the nucleus of the meristemoid (M), which may further asymmetrically divide. BASL is not found in later stomatal lineage cells, such as GMCs or GCs. However, when satellite meristemoids are formed by SLGCs that resume stomatal lineage fate, BASL appears at the SLGC periphery next to the stomatal lineage cell, providing a mechanism for maintenance of the one-cell spacing rule. (Based on data presented in Dong et al. [2009].)

Figure 1-4.

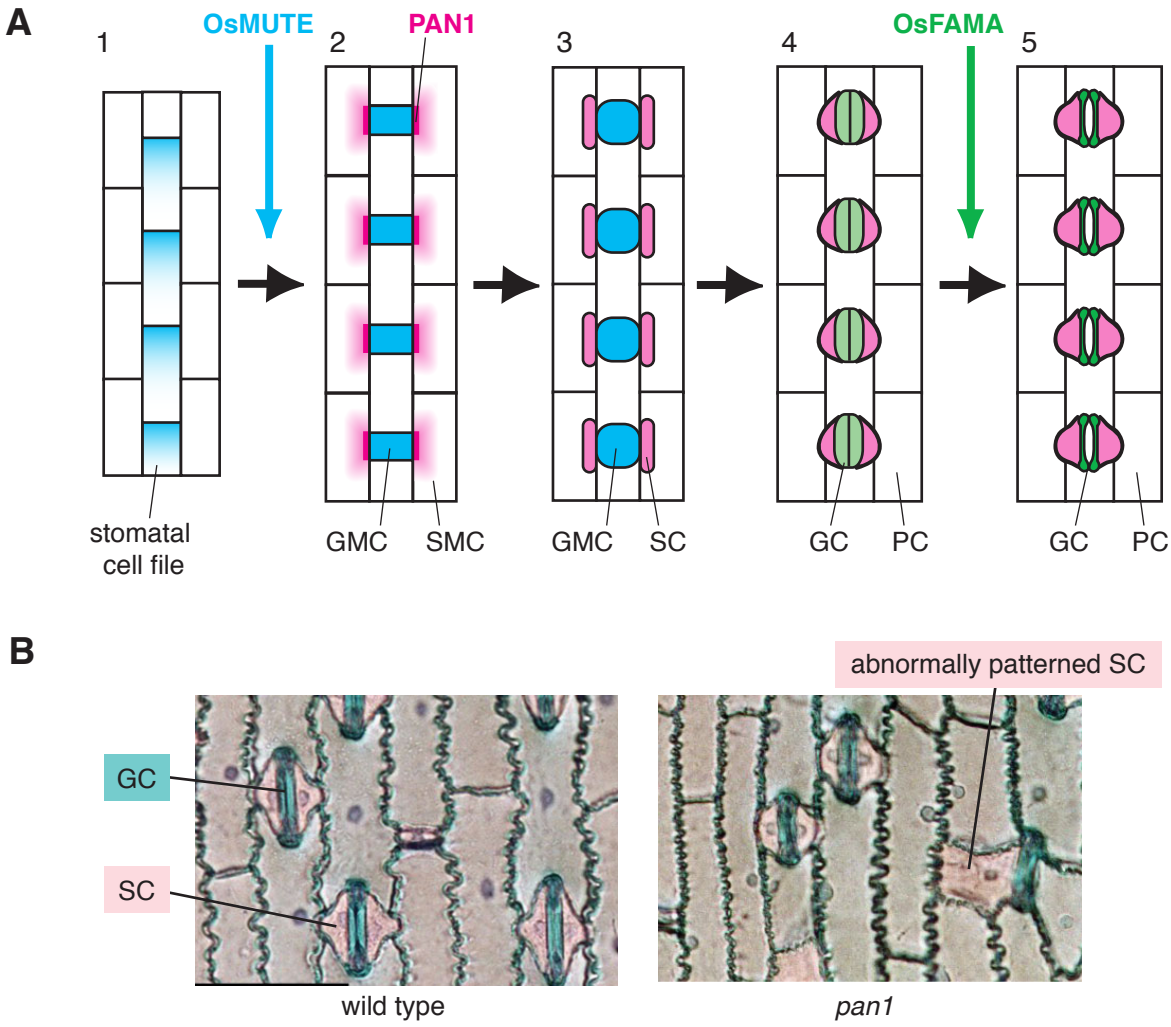


Figure 1-4. Stomatal development in grasses.
(A) Schematic diagram of stomatal development in grasses. (1) During early epidermal development in grasses, stomatal and nonstomatal cell files are specified, and cell division polarity is established in the stomatal cell file. This polarity will ensure that the one-cell spacing rule is maintained. (2) *Os MUTE* controls a single asymmetric division toward the leaf apex in the stomatal cell file, which creates GMCs (blue). Neighboring cell files (SMC; pink gradient) receive a signal via putative receptor PAN1 (magenta), which localizes at the area of GMC contact, and polarize in preparation for division. (3) SMCs divide asymmetrically toward PAN1 to form SCs (pink), which will act as ion reservoirs for the operation of mature stomata. (4) GMCs divide once symmetrically to form GCs (light green). (5) Finally, GC and SC terminally differentiate, forming mature dumbbell-shaped stomata (dark green). *Os FAMA* is required for the differentiation of GCs, though not for their symmetric division.
(B) Leaf epidermal peel from wild-type (left) and *pan1* mutant maize (right). GC and SC are stained in cyan and pink, respectively. Unlike the wild type, the *pan1* mutant occasionally fails to produce proper asymmetric divisions that give rise to SC, resulting in abnormal SC patterning. Images kindly provided by Laurie Smith (University of California, San Diego).

Figure 1-5.

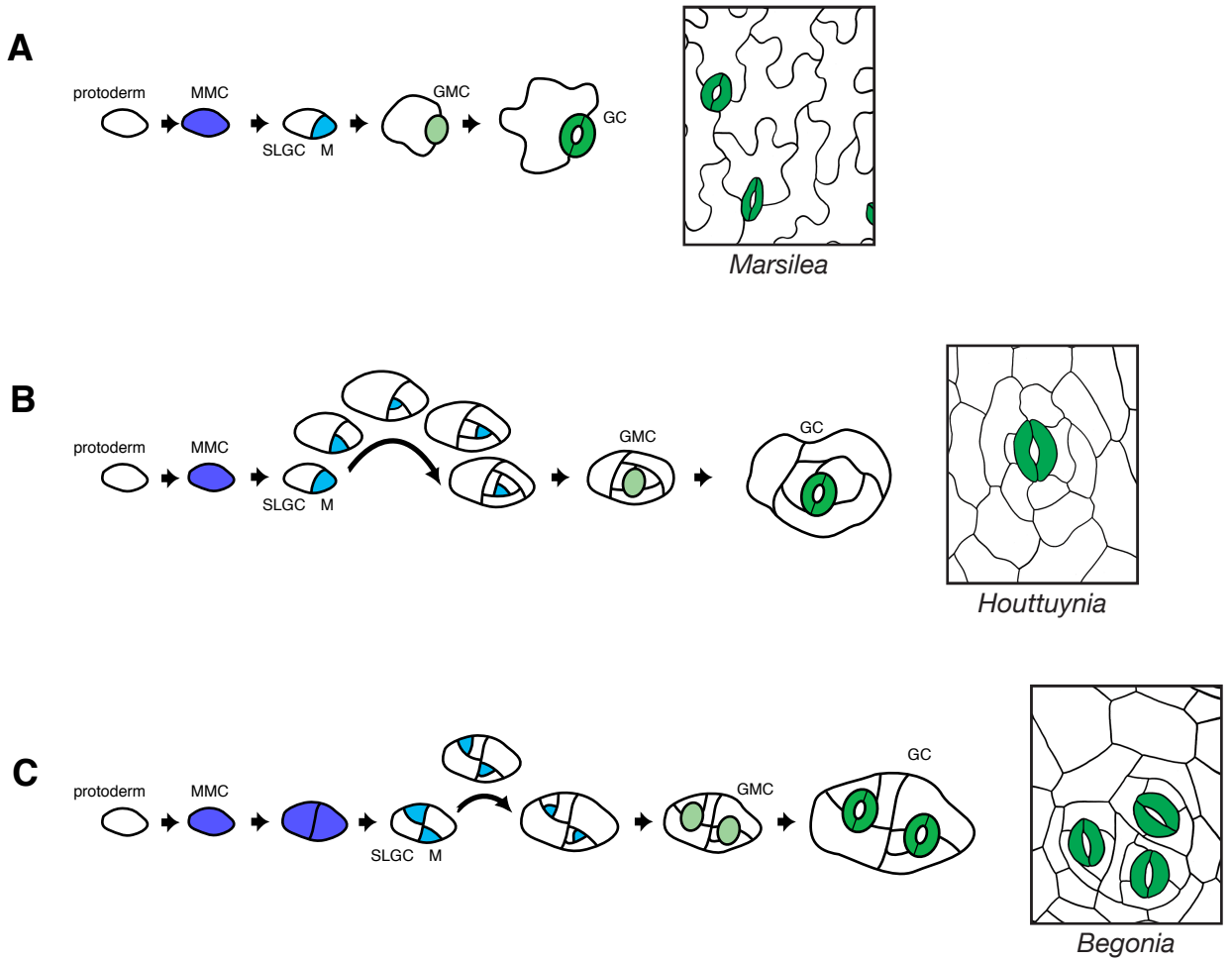


Figure 1-5. Developmental hypotheses for stomatal complex diversity.
(A) Stomata in the fern *Marsilea* appear to develop through a process lacking amplifying divisions of a meristemoid (cf. *Arabidopsis* in Figure 1-2A).
(B) *Houttuynia* (magnoliid) stomata are surrounded by a spiral arrangement of cells that suggests a large number of amplifying divisions.
(C) In *Begonia*, a eudicot, stomata arise in groups that can be explained by early division of a stomatal precursor, such as an MMC, and retention of MMC identity by all daughter cells.

Figure 1-6.

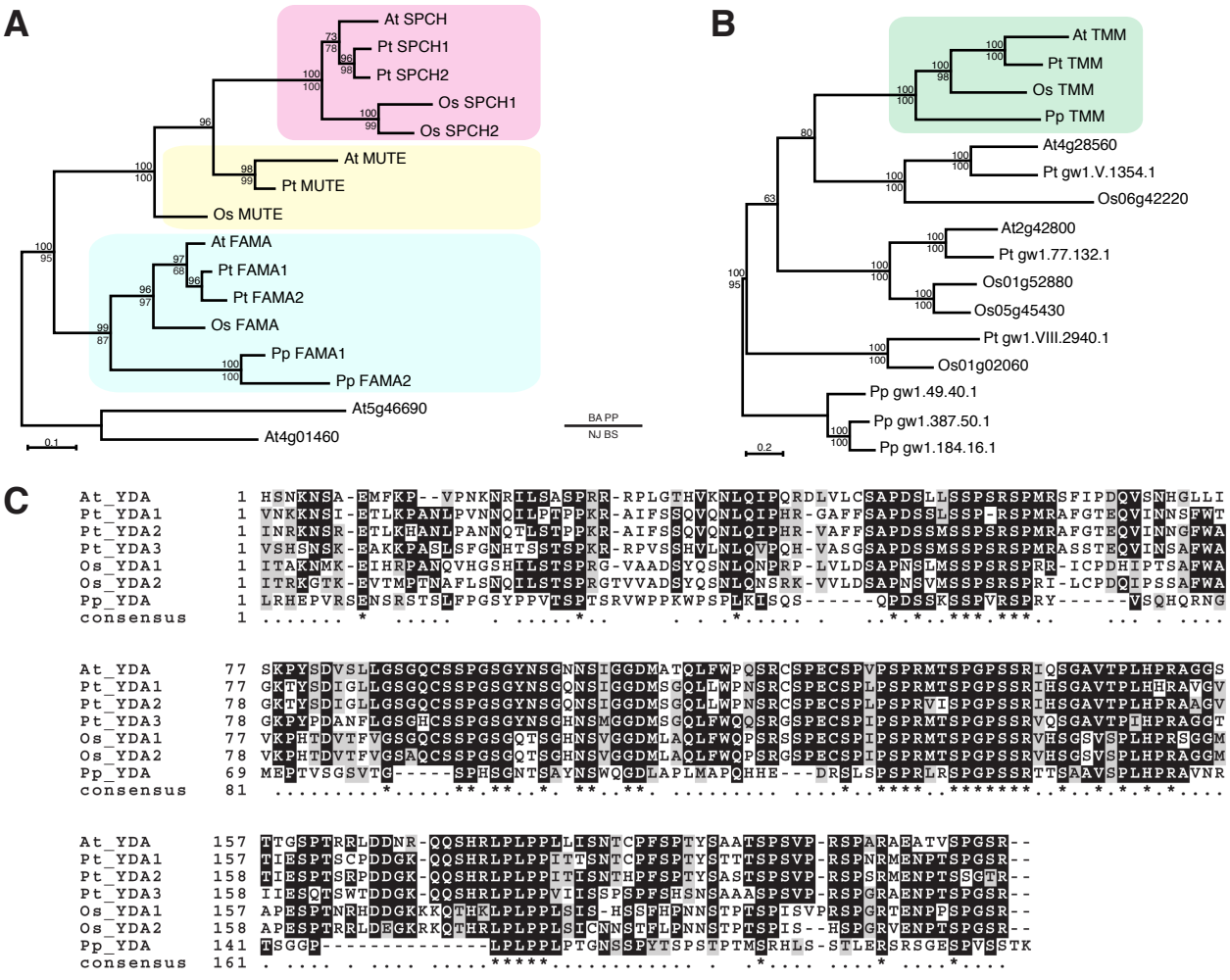


Figure 1-6. Stomatal patterning and spacing genes are conserved among embryophytes.
(A) A gene tree of master regulatory bHLH transcription factors (*SPCH*, *MUTE*, and *FAMA*) in stomatal development. Amino-acid sequences from bHLH and ACT domains were aligned and the tree constructed using Bayesian and neighbor-joining methods. The Bayesian phylogram is shown. The pink, yellow, and cyan rectangles highlight the SPCH, MUTE, and FAMA clades, respectively. Posterior probability values are indicated above the nodes, and bootstrap values over 50 (100,000 replicates) are indicated below the nodes. (B) A gene tree of *TMM*, which encodes an LRR receptor-like protein necessary for proper stomatal spacing in Arabidopsis. The entire amino-acid sequence was aligned and the tree constructed as in (A). The Bayesian phylogram is shown. The green rectangle highlights the single-copy TMM clade. Posterior probability values are indicated above the nodes, and bootstrap values over 50 (100,000 replicates) are indicated below the nodes. (C) Amino-acid alignment of the N-terminal portion of YODA (YDA), a MAPKKK, which is essential to regulate the activity of the protein in Arabidopsis. YDA is also required for appropriate stomatal spacing. At, *Arabidopsis thaliana*; Pt, *Populus trichocarpa*; Os, *Oryza sativa*; Pp, *Physcomitrella patens*. See Methods and Table 1-1 for detailed bioinformatic and phylogenetic methods and gene ID numbers.

Figure 1-S1.

```
At_SPCH       MSHVTVERNRRKQMNEHLTVLRSLMPCFYVKRGDQASIIGGVVEYISELQ
At_MUTE       MSHIAVERNRRRQMNEHLKSLRSLTPCFYIKRGDQASIIGGVIEFIKELQ
At_FAMA       MTHIAVERNRRKQMNEHLRVLRSLMPGSYVQRGDQASIIGGAIEFVRELE
Pt_SPCH2      ISHITVERNRRKQMNEHLSVLRSLMPCFYVKRGDQASIIGGVVDYINELQ
Pt_SPCH1      MSHITVERNRRKQMNEHLSVLRSLMPCFYVKRGDQASIIGGVVDYINELQ
Os_SPCH1      TAHIAVERNRRKQMNENLAVLRSLMPCFYVKRGDQASIIGGVVDYIKELQ
Os_SPCH2      MSHITVERNRRKQMNEHLAVLRSLMPCFYVKRGDQASIIGGVVDYIKELQ
Pt_MUTE       MSHIAVERNRRRQMNEHLKVLRSLTPCFYIKRGDQASIIGGAIEFIKELH
Os_MUTE       MSHIAVERNRRRQMNDHLKVLRSLTPAFYIKRGDQASIIGGAIDFIKELQ
Pp_FAMA1      MTHIAVERNRRKQMNEHLAALRALMPGSYVQKGDQASIVGGAIEFVKELE
Pp_FAMA2      MTHIAVERNRRKQMNEHLTALRALMPGYFIQKGDQASIIGGAIEFVRELE
Pt_FAMA1      MTHIAVERNRRKQMNEHLRVLRSLMPGSYVQRGDQASIIGGAIEFVRELE
Pt_FAMA2      MTHIAVERNRRKQMNEHLRVLRSFMPGSYVQRGDQASIIGGAIEFVRELE
Os_FAMA       MTHIAVERNRRRQMNEYLRVLRSLMPGSYVQRGDQASIIGGAIEFIRELE
At5g46690     MTHIAVERNRRRQMNQHLSVLRSLMPQPFAHKGDQASIVGGAIDFIKELE
At4g01460     MTHIAVERNRRRQMNEHLNSLRSLMPPSFLQRGDQASIVGGAIDFIKELE


At_SPCH       QVLQSLEAKKQRK-----LADVEVKFSGAN--VLLKTVSHKIPGQVMKII
At_MUTE       QLVQVLESKKRRK-----HANVEAKISGSN--VVLRVVSRRIVGQLVKII
At_FAMA       QLLQCLESQKRRR-----LADVEVKLLGFD--AMIKILSRRRPGQLIKTI
Pt_SPCH2      QVLQSLEAKKQRK-----IADVEVKFSGPN--VLLKTVSPRIPGQAVKIV
Pt_SPCH1      QVLQSLEAKKKRK-----IADVEVKFSGPN--VLLKTVSPQIPGQAVKII
Os_SPCH1      QVLHSLEAKKQRK-----LPDVKVEFAGAN--LVLKTVSQRSPGQAVKII
Os_SPCH2      QVLRSLEAKKNRK-----VPDVRVEFAGPN--LVLKTVSHRAPGQALKII
Pt_MUTE       QVLQALESKKQRK-----IADVEAKISGSN--VILKVISRRIPGQIVRII
Os_MUTE       TLLQSLEAQKKRR-----MADVEARISGAN--VLLRTLSRRAP--PVRII
Pp_FAMA1      HLLHCLQAQKRRR-----MASVEVKMVGSDQ-AMVKIMAPRRSGQLLRTV
Pp_FAMA2      HLLHCLQAQKRQR-----QASVNVKMVRIDQ-ALVKVLAPRRSGQLLRTV
Pt_FAMA1      QLLQCLESQKRRR-----LADVEVKLLGFD--AMIKILSRRRPGQLIKAI
Pt_FAMA2      QLLQCLESQKRRR-----LADVEVKLVGFD--AMIKILSRRRPGQLSKTI
Os_FAMA       QLIQCLESQKRRR-----VADIEVRVVGVD--AMIKILSRRRPGQLIKTV
AT5g46690     HKLLSLEAQKHHN-----ETHANIRILSRRRGFRWSTLATTKPPQLSKLV
AT4g01460     QLLQSLEAEKRKD-----TTEVEATVIQNH--VSLKVRCKRGKRQILKAI


At_SPCH       AALED-LALEILQVNINTVD-ETMLNSFTIKIGI--ECQLSAEELAQQIQ
At_MUTE       SVLEK-LSFQVLHLNISSME-ETVLYFFVVKIGL--ECHLSLEELTLEVQ
At_FAMA       AALED-LHLSILHTNITTME-QTVLYSFNVKITS--ETRFTAEDIASSIQ
Pt_SPCH2      SALEG-LALEILHVSISTVDHETMLNSFTIKIGI--ECQLSAEDLAQQIQ
Pt_SPCH1      SALED-LALEILHVSISIVDHETMLNSFTIKIGI--ECQLSAEELAQQIQ
Os_SPCH1      AALEG-RSLEILHAKISTVD-DTAVNSFTVKIGI--ECELSAEELVQVIQ
Os_SPCH2      AALES-LSLEILHVSICTVD-DATVLSFTIKIGI--ECELSAEELVQEIQ
Pt_MUTE       SVLEN-LSFEILHLNISSME-DTVLYSFVIKIGL--ECQVSVEELAVEVQ
Os_MUTE       ALLES-LHLEVLHLNITTMD-DTVLYSFVLKIGL--DCHLSVDDLAMEVH
Pp_FAMA1      VALES-LALTVMHTNITTVH-HTVLYSFHVQISL--HCRLNVDEVAAALH
Pp_FAMA2      MALEG-LALTVLHTNITTVH-HTVLFSFHVHMGL--LCRMSVKEIATVLH
Pt_FAMA1      AALED-LQLNILHTNITTID-QTVLYSFNVKV------------------
Pt_FAMA2      AALED-LQLNIHDTNITTID-QTVLYSFNVKV------------------
Os_FAMA       AALEE-MHMSILHTNITTID-QTVLYSFNVKIAG--DARFSAEDIAGAVH
At5g46690     ASLQS-LSLSILHLSVTTLD-NYAIYSISAKVEES-CQLSSVDDIAGAVH
At4g01460     VSIEE-LKLAILHLTISSSF-DFVIYSFNLKMEDG-CKLGSADEIATAVH
```

```
At_SPCH       QTF
At_MUTE       KSF
At_FAMA       QIF
Pt_SPCH2      QTF
Pt_SPCH1      LTF
Os_SPCH1      QTF
Os_SPCH2      QTF
Pt_MUTE       QSF
Os_MUTE       QSF
Pp_FAMA1      QTF
Pp_FAMA2      GTF
Pt_FAMA1      ---
Pt_FAMA2      ---
Os_FAMA       QIL
At5g46690     HML
At4g01460     QIF
```

Figure 1-S1. Amino-acid alignment of the bHLH domain and the C-terminal ACT-like domain of *SPCH, MUTE,* and *FAMA* orthologs used for the phylogenetic analysis presented in Figure 1-6A. At5g46690 and At4g01460 were used as an outgroup.

Figure 1-S2.

```
At_TMM                -----MARYEFFRQIFIVLSIVSPLVR-SFTVITSDSTAPSALIDGPQTG
Pt_TMM                ------------------MSLLLPLAK-PFTVIVSDSSNQSALVDGPQSG
Os_TMM                MVASALSATAAATVVVAVVAVLVAVCRGEFTVVVPDSSSSAALVNAPQTG
Pp_TMM                --------------------------------------------------
At4g28560             ----------------------MASFTNLSLLLLFLFTATLITSQQSDD
At2g42800             ------------------------MTMKRALPSPSSLLFFFLLITPLFL
Pt_gw1.V.1354.1       ----------------------SSFPIFGLLFLTTLMSSLVISHQ---
Pt_gw1.77.132.1       ------------------------------------------------ILSV
Pt_gw1.VIII.2940.1    -------------------------------LCIMSLG---VWCYGEED
Os06g42220            ----------------------MRMILGRLLVVAVLAAAAAAAES---
Os01g52880            ------------------------------------MAASVAVFVCLLSV
Os05g45430            -----------------------------------MPVASVVLLLHLLSM
Os01g02060            --------------MRQPQSRKLLQLQALSLLLFIIALHSRLHGCSGQGE
Pp_gw1.49.40.1        --------------------------------------------------
Pp_gw1.184.16.1       --------------------------------------------------
Pp_gw1.387.50.1       --------------------------------------------------


At_TMM                FTMTNDGARTEPDEQDAVYDIMRATG----NDWA--AAIPDVCR--GRWH
Pt_TMM                FTMNKNGARTDAREQEAVYDIMRATG----NDWA--TDIPDVCR--GRWH
Os_TMM                LS---DRARTDPAEQRAVQEVMAATG----NGWA--SGIADVCR--GRWH
Pp_TMM                -------ARTDPVELQAVYAMMAATG----NGWA--ASIPDVCVKSGRWH
At4g28560             DENASPQLALDPSEQEAVYRVLDSVN--SAISWR--TIFPDDICASPP-D
At2g42800             CQENRVSASMPPSESETLFKIMESMS--SDQQWR--QSHPNPCAPGSSWP
Pt_gw1.V.1354.1       ------QPLLDSAEQDSLFQVLYSIN--SAIPWR--TLFPDDLCLSAP-H
Pt_gw1.77.132.1       CSSD--AKNMLPSEVETIFKIMESIS--SDEKWR--ISYPNPCNPGSTWL
Pt_gw1.VIII.2940.1    YND------AVPMEEGEKTALYSAIQGFVGNWWNGSDLYPDPCG-WTPVQ
Os06g42220            ------EPRLPAAEQEGVYAVLEAVN--PGFPWR--ASFPDDLCLAGP-H
Os01g52880            AAAAAS---MDPAEREALFLVMEAVS--SDRDWR--SESPDPC--GAPWP
Os05g45430            AAAATAALAMDPAERETLLLVMEAVS--SDREWR--SVGPDPC--GSPWP
Os01g02060            AADGSASTAAAPMEEKEKRALYAAIEGFVGKGWNGSALYPDPCG-WSPIQ
Pp_gw1.49.40.1        -----------PVDLQVLHEMMYSLN--FEEDWN--TYFPDPCV-SGP-Q
Pp_gw1.184.16.1       ------------------------------------YPDPCV-SGP-Q
Pp_gw1.387.50.1       ---MVASQATNSVDLQVLHEMMYSIN--FEYNWD--SYYPDPCI-SGP-Q


At_TMM                GIECMPDQ-----------------DNVYHVVSLSFGALSDD-TAFPTCD
Pt_TMM                GIECMPDK-----------------DNVYHVVSLSFGTLSDD-TAFPTCD
Os_TMM                GIECVPDR-----------------GEVYHVVSLSFGALSDD-TAFPACD
Pp_TMM                GIECVQD------------------GDYYHVVELSFGLVTDQATAFPSC-
At4g28560             GVVCDNLYASQ-----NG-VATS-----VHVTEFHLGYLSDYTQNPPCSS
At2g42800             GIECK-----------------TGPDHLSHVSRLDFG-----SAPNPSCK
Pt_gw1.V.1354.1       GIVCEYFTEEQPPLTPNGSVSTQPPLETAHISELSFGFVSDYTSNPPCSP
Pt_gw1.77.132.1       GIECK-----------------LGQDNHLHVSRLDFG-----THPNPTCK
Pt_gw1.VIII.2940.1    GVSC-----------------DLFDGLWYVTALSIG---PILDNYLDCA
Os06g42220            GVSCDDDDGGN----------------ASHVVGISLGYVSDFSANPSCAA
Os01g52880            GLECK--PAAG-----------DAAAALLRVTRLDFG-----VEPNPSCK
Os05g45430            GLECKPVPAAGN---------VSSAAARLHVTRLDFG-----VAPNPTCK
Os01g02060            GVSC-----------------DLFNGLWYPTVMSIG---PVLDNSLRCS
Pp_gw1.49.40.1        GILCEP----------------DPVTGVLFVTQLQFGY-ISPIDNIIPCS
Pp_gw1.184.16.1       GIVCEV----------------DPVTNLYFVTQMQFGF-ISPIANLIPCS
Pp_gw1.387.50.1       GIVCEA----------------DPVTNLYYVTQMQFGY-ISPIANLIPCS
```

```
At_TMM                ---PQRSYVSESLTRLKHLKALFFYRCLGRAPQRIP----AFLGRLGSSL
Pt_TMM                ---PARSYISESIMKLPHLKTLFFYRCFSDNPQPIP----SFLGQLSPTL
Os_TMM                ---AARATLSPAVLALPHLRSLFFYRCFTANPQPVP----AFLGRLGPAF
Pp_TMM                ---AQHSFISPAVANLTHLRKLGYFACCMENPQRIP----SEIGRLGDSL
At4g28560             ----NATLDPLLFTAFKHLRKLFFYKCFTDARASLPLT---VPEDFGSVL
At2g42800             ---SSASFPSSIFT-LPFLQSVFFFNCFTHFPTTIMFPIK---LIPNSSL
Pt_gw1.V.1354.1       ----NSTINPLVFTSFKFLRKLFFYKCFTEMPVSVPDV---SSSSFGANL
Pt_gw1.77.132.1       ---NTAAFPCEIFD-LPHLQSVFFSQCFTHTKTTLSASPNRGGAFFDSSL
Pt_gw1.VIII.2940.1    ---PIVEFRPQLFE-LKHLKSLTFFSCFVSPHEHPIIIPSKNWEKLAGNL
Os06g42220            PSAATLLTSGLLAASFPRLRSLFVYGCFVGDDDARPLPP--LPWRLPPTL
Os01g52880            ---DTAAFPPLVFSSLPHLQSLFFVGCFKNPAANTSLVLPPAANLSTSSL
Os05g45430            ---DGAAFPHLAFA-LPHLQSLFLVDCFKNPAATTAFTLPPSANLTSSRL
Os01g02060            ---ADAKFSPQLFD-LKRLKTLSFYSCFPATNPTP--IPATSWDKLAGSL
Pp_gw1.49.40.1        ---FNASIPSSIAK-LTRLDTLAFYSCFVNSTATIP----EEITLLGPTL
Pp_gw1.184.16.1       ---WNATIPTSIAN-LKRLDTLSFYNCFMNSTITIP----KEISLLGPTL
Pp_gw1.387.50.1       ---WNATIPASIAN-LTRLDTLSFYNCFSISTVSIP----EDISSLGPTL

At_TMM                QTLVLREN-GFLGPIPDELGNLTNLKVLDLHKNHLN-GSIPLSFNRFSGL
Pt_TMM                QTLVLREN-GHVGLVPAELGNLTRLKVLDLHKNNLN-GSIPVSLGRLAGL
Os_TMM                RSLVLREN-GHVGAIPPELGNLTALRVLDLHGNNLT-SAIPATVQSLAHL
Pp_TMM                EALHLRNN-GHVGIIPEELAGLIKLRTFDVHGNSLA-GTMPIWLSSLTEL
At4g28560             EELVFIENPSLVGEIGAMIGNFTKLRRLVLTGNGFH-GSIPGQIGDLVSL
At2g42800             QQLSLRSNPSLSGQIPPRISSLKSLQILTLSQNRLT-GDIPPAIFSLKSL
Pt_gw1.V.1354.1       EELVFIENPALVGSLSGIIGNFTNLRRLVLTGNGIY-GNIPDGVGSLVNM
Pt_gw1.77.132.1       QQLSLRSNPALVGPIPPQISFLKSLEILTLSQNRLS-GRIPVEIFSLNSL
Pt_gw1.VIII.2940.1    ETLEFRSNPGLIGKVPSSFGSLIRLQSLVLLENGLT-GELPRNVGNLTKL
Os06g42220            QDLVLVNNPALTGRLAISAASLPLLRRLVVASSGLS-GDLPSTP--FPRL
Os01g52880            QQLSIRANPSLSGVMPPQLATLRSLQVLTISQNGLIRGEIPQGIGELTSL
Os05g45430            QQLSVRSNPSLSGTLPPQLSSIRSLQVLTVSQNALIRGEVPQGIGELKSL
Os01g02060            ETLEFRTNPGLTGPIPASLGRLSSLQSLVFVENNLT-GAVPAELGSLVRL
Pp_gw1.49.40.1        RLLSLRRNPSLTGTIPAGIGKLTGLQRLVLSQNGLQ-GEIPAELSNLQNL
Pp_gw1.184.16.1       RLLSFSGNAGLTGAIPAGFGKLTRLQRLVLSQNRLQ-GSIPEDLSNLQRL
Pp_gw1.387.50.1       RRLSFSGNPGLTGAIPSGLGKLTGLQRLVLSQNGLQ-GRIPEELGNLQCL

At_TMM               RSLDLSGNRLTGSIPGFVLP-----ALSVLDLNQNLLTGPVPPTLT-SCG
Pt_TMM               KSLDLSGNKLTGSIFSLSFP-----VLNVLDLSQNMLTGSIPSSLG-FCH
Os_TMM               QLLDLSYNQLAGEVPPFKFQ-----HLSILDLSHNALQGGVPASLG-QCR
Pp_TMM               EAMDISDNTFGGEVDGRTFDN--LERLTVFDASDNEFVGALPDSIG-RLR
At4g28560            EEITLSRNSLTGGFPANATSR--LKNLKVLDFSHNFINGNAPDSIG-DLT
At2g42800            VHLDLSYNKLTGKIPLQLGN---LNNLVGLDLSYNSLTGTIPPTIS-QLG
Pt_gw1.V.1354.1      EEVTVSRNQLSGGVPF-SLAK--LKKLRVLDLSQNYLDGYVPLSVG-NLS
Pt_gw1.77.132.1      VHLDLSYNMLTGPVPIQLGN---LNNLQGLDLSYNSLTGPIPGTIG-RLG
Pt_gw1.VIII.2940.1   KRLVLAANWFYGRIPDNFGG---LNELLILDLSRNLLSGSLPLTLG-GLN
Os06g42220           EQLVLSGSRFAGRIPSALVQG--LANVKILDLSSNLLAGGIPRAIG-GLT
Os01g52880           VHLDLSYNSLTGPVPSEISE---LKSLVGLDLSYNSLSGAIPSRIG-ELR
Os05g45430           VHLDLSYNSLTGTIPSRIGE---LRSLVGLDLSYNSFSGSIPGQLG-DLA
Os01g02060           RRLVLSGNGLSGQIPASLGNGHFAEQLLIMDVSNNSLTGSLPSSLG-GLK
Pp_gw1.49.40.1       IQLDLSHNNLSGSIPATLST---MDSLVNLDLRYNQLDGEFPAGLGQGFG
Pp_gw1.184.16.1      IQLDFSHNNFSGSVPATFGA---MSSIVNMDLRYNHLEGTLPASIIQGMP
Pp_gw1.387.50.1      IQLDFSHNNLSGSVPETFGA---MNSLVNLDLRYNHLDGTLPPSLTQGLP
```

```
At_TMM                SLIKIDLSRNRVTGPIPESINRLNQLVLLDLSYNRLSG-PFPSSLQGLNS
Pt_TMM                SLIKIDFSRNRLSGSIPESIGSLKELILMDLSYNRLSK-PLPTSIRSLNS
Os_TMM                SLLKFDLSQNRFAGTIPDALGDLSDLILLDLSHNALSG-PIPAALGRLSS
Pp_TMM                TLQKLDLSYNNFTGAIPTTIGDLSRLLSLNLAHNRFSG-PLPETMSNLSN
At4g28560             ELLKLDLSFNEFTGEVPSGVGNLKKLVFLDLSYNRFGNFGVPLFLAEMSS
At2g42800             MLQKLDLSSNSLFGRIPEGVEKLRSLSFMALSNNKLKG-AFPKGISNLQS
Pt_gw1.V.1354.1       RLLKLDLSHNRLSGKIPESLVSLQSLEFLDLSFNSFGNYGVPLFLGEMPR
Pt_gw1.77.132.1       MLQKLDLSSNSFIGTIPYSIEKLTLLTFMALSNNKLRG-SIPKGILKLQS
Pt_gw1.VIII.2940.1    SLLKLDLSNNQLVGSLPTVMCYMKNLTLLDLRNNRFSG-GLTKSLQEMYS
Os06g42220            QLVKLDLSSNTLAGPIPGELGGLASLELLDLSNNRLTG-GVPAALRGMTA
Os01g52880            QLQKLDLSSNNLTGGIPVSIANLSSLTFLALSSNGLSG-HFPPGLSGLRN
Os05g45430            MLQKLDLSSNNLTGGVPATITGLTSLTFLALSNNGLSG-HLPAGLSDLLD
Os01g02060            GLLKMDLSNNLLQGSLPPELAGLGSLTLLDLRNNSFTG-GLPSFLQGMAS
Pp_gw1.49.40.1        HLQRLAASYNKLSGSLPDTFTGLKYLTFLDLSYNHLMG-NLPPSLGNLAN
Pp_gw1.184.16.1       QLQRLAVSHNHLSGSLPDTFTGLSSLTFLDLSHNELTG-PLPPSLGRLRN
Pp_gw1.387.50.1       QLQRLALSHNHLSGSLPDTFTGLNSLTFLDLSHNELTG-LLPPSLGHLTN


At_TMM                LQALMLKGNTKFSTTIPENAFKGLKNLMILVLSNTNIQGSIPKSLTR-LN
Pt_TMM                LQALILKGNPMGSTTLASKEFDGMKSLMVLILSNTNLHGPIPESLGR-LT
Os_TMM                LRSLILGDNRMQFTTVPGDIFAGLRALTTLVLSGMGLEGSLPESIGE-LG
Pp_TMM                LKSLDLQRNCFRVPIPAS--LGKLVKLEGLVLSESEFVGPIPSSFGS-LS
At4g28560             LREVHLSGNK-LGGRIP-AIWKNLEGISGIGFSRMGLEGNIPASMGSSLK
At2g42800             LQYFIMDNNPMFV-ALP-VELGFLPKLQELQLENSGYSGVIPESYTK-LT
Pt_gw1.V.1354.1       LKEVYLSGNL-LGGHIP-EIWEKLGGISGIGFSDMGLVGNIPASMGVHLR
Pt_gw1.77.132.1       LQYFIMDDNPMYI-PLP-AEFGKLVKLQELRLSNSCYSGTIPPSFSL-LV
Pt_gw1.VIII.2940.1    LEEMALSNNPIGG-DLQGLEWHSLQNLVVLDLSNMGLTGEIPESIAE-LK
Os06g42220            IREMYLSGNRRLGGRVPADMFAGLKGISAVGLSDAGLTGTIPASLGESLR
Os01g52880            LQCLIMDNNPMNV-PLP-SELGGLPRLQELRLAGSGYSGQIPAAFGQ-LA
Os05g45430            LQYLIMENNPMGV-PLP-SELGDIARLQELRLANSGYSGSIPETLGR-LA
Os01g02060            LQDLLLSNNPLGG-SLGQLGWERLRGLATLDLSNLGLVGAIPESMAA-LT
Pp_gw1.49.40.1        LQDLFLNSNSLDG-EIP-ESLGSLIPLKRLDLSSCGFVGLIPDSLKG-LQ
Pp_gw1.184.16.1       LEDLFLNSNPLDG-NIP-PSLGSMRSLVRLDLSSCLLSSIIPDSLKN-LE
Pp_gw1.387.50.1       LEDLFLNSNSLVG-NIP-APIGMLKSLVRLDLSSCSFGNKIPDSLKN-LE


At_TMM                SLRVLHLEGNNLTGEIPLEFRDVKHLSELRLNDNSLTGPVPFERDTVWRM
Pt_TMM                NLRVIHLDGNHLNGSIPINFKDLKNLGELRLNDNQLTGPVPFGREMVWKM
Os_TMM                HLRVLRLDNNEFTGVIPASFRRLERASELRVDGNRLVGPIPFGKQMMWRL
Pp_TMM                NIRALFLDGNKLTGTIPPALGDLTRVYELELSSNLLAGPVPFSSSFVSRL
At4g28560             NLCFLALDNNNLDGQIPEEFGFLDSAREINLENNNLTGKAPFSDSFRDRI
At2g42800             NLSSLSLANNRLTGEIPSGFESLPHVFHLNLSRNLLIGVVPFDSSFLRRL
Pt_gw1.V.1354.1       NLCYLGLDNNKLEGTVPEELGFLKCGYEINLENNNLSGKIPVT--FTSKV
Pt_gw1.77.132.1       NLSTLSLQNNRLTGKIPEGFSSLSRIYHLNLSGNLLGGVIPFNASFLKRL
Pt_gw1.VIII.2940.1    RLRFLGLRGNRLTGNLSPKLATLPCVSALYLDGNNLTGELKFSGWFYGKM
Os06g42220            NVTYLGLDGNLLEGEVPPALAKMAGRVRLHGNRAVCISPEFLAGAPRPRI
Os01g52880            SLTTLSLEDNNLTGEIPPVLTRLTRMYHLNLSNNGLGGAVPFDGAFLRRL
Os05g45430            SLTTLSLENNNLTGRIPAGLSRLKRMYHLNLSKNGLDGVVPFDGAFLRRL
Os01g02060            RLRFLALDHNRLTGDVPARLAELPNIGALYLNGNNLTGTLQFSPAFYQRM
Pp_gw1.49.40.1        NLRYLSVSNNHLSGPIPASLASLPVLFTLNLDGNQLSGAVPFPSSFIQKM
Pp_gw1.184.16.1       NLRFLSMSNNKLSGFIPASLASLPKIFTLNLDGNKLTGPVPFPPTFVKKM
Pp_gw1.387.50.1       NLRFLSISNNKLSGPIPASLASLPQIFTLNLDGNQLTGPVPFPASFVKKM
```

```
At_TMM                RRKLRLYNNAGLCVNRDSDLDDAFGSKSGSTVRLCDAETSRPAPSGTVQH
Pt_TMM                RRKLKLYNNTGLCYDANSGFEDGLDSTFDSGIGLCN--TPEPGSARTVQH
Os_TMM                GKKLRVGGNEGLCYDAK---QQGLEG----VVALAGVADCDSVRSRTTQH
Pp_TMM                GRRLRLQNNEGLCYEFK---------------------------------
At4g28560             GKKLKLSGNVNLQVKNS----DPHLAGRALYSSARKVLPLIYFPATLLAL
At2g42800             GKNLDLSGNRGLCLNPEDEFSVVKTG--VDVCGKNVSSGGGLSVHSSKKK
Pt_gw1.V.1354.1       AEKLKLKGNSGLCVDGG----D-----------------------------
Pt_gw1.77.132.1       GRNLDLSGNPGLCLSPSEAYNNVKSGSGVGVCGTG---------------
Pt_gw1.VIII.2940.1    GRRFGAWNNPNLCYPVGLMSTGHAPYG-VKPC------------------
Os06g42220            AGVPSCNATQAAPVTRRPVVMPVPLASAEKPAAAAPPPPMRIGSCVVVAM
Os01g52880            GQNLDLSGNAGLCLDDRMVVRGVGVG--VGACH-AGGGGDGPLAPGGVTG
Os05g45430            GRNLDLSGNPGLCVDGRAVLQ-ADVG--VGVCRRAGDGGDIASVSAATDV
Os01g02060            GRRFASWDNPGLCYSNAAVDAAHAPPG-VTVCKVAGGVGDGRKPEASSSL
Pp_gw1.49.40.1        GRNMLLNGNPGLCYTPQ---------------------------------
Pp_gw1.184.16.1       GRNMRLEGNPGLCF------------------------------------
Pp_gw1.387.50.1       GRNMRLGDNPGLCFNSQLVSVKIPDLG-LNHCADPPTMPANAPRSPLATL
```

```
At_TMM                LSR--EEDGALPDGATDVSSTSKSLGFSYLSAFFLVFPNFIFMLISS-
Pt_TMM                LSA--IDGETMPS-TTNKSSAALHKASVFVRLLQQTIAIVSTLIFIST
Os_TMM                LGGRLRNTGGLPAAAAAAPSVGVGAG-VCVGSWHVFVGVLVSLHLARL
Pp_TMM                ------------------------------------------------
At4g28560             YISITQ--------------------------------------------
At2g42800             SQASRYYRSCF---------FANALFPFALFLGLHQRWVL--------
Pt_gw1.V.1354.1       ------------------------------------------------
Pt_gw1.77.132.1       ------------------------------------------------
Pt_gw1.VIII.2940.1    ------------------------------------------------
Os06g42220            LLLMLS--------------------------------------------
Os01g52880            AAATVRGSVDG---------YPFRLLGHACLVVACLVSLN--------
Os05g45430            LSVGTLFRRDGQLWLAGGRWSALLLIRPVAVALCCSCCLLL-------
Os01g02060            MATSSASNLINGFCFFLWMVATSLL----------------------
Pp_gw1.49.40.1        ------------------------------------------------
Pp_gw1.184.16.1       ------------------------------------------------
Pp_gw1.387.50.1       YPNATIRQTSLSLFQWSFWTSTLALLFLEV------------------
```

Figure 1-S2. Amino-acid alignment of the full-length *At TMM* and its orthologs used for the phylogenetic analysis presented in Figure 1-6B.
At4g28560 and At2g42800 and related genes from *Pt*, *Os*, and *Pp* were used as outgroups.

Figure 1-S3.


At_YODA M-PWWSKS--KDEKKK----------TNKES-IID--AFNRKLGFASEDRSSGRSRK
Os_YODA1 MPPWWGKSFSKDAKK----------TTKEN-LID--TFHRLIS-PNDQKGSTKSKR
Os_YODA2 MPPWWGKSSSKEVKK----------TAKEN-LID--TFHRLLS-PNEQKGRTKSRG
Pt_YODA1 ----------
Pt_YODA2 MRSWWGKSSSKEEKKK----------ANKES-FID--TINRKFKITSKEKSNNRSGG
Pp_YODA M-PWWPISSKDSKKGK----------KPTTSGLFDGGSGHSRSPKPRAKEDNGSSKC
At1g53570 MPTWWGRKSCKNKDD----------NHRGIISTDRDIKSSAVVVDPPLTPTRGG
At5g66850 M-RWLPQISFSSPSSSPSSSLKPVASYSESPDPDRNQDRDRFHRRLFRFNRGRLTRQRKL
At1g09000 ----------
At1g54960 ----------
At3g06030 ----------
At_YODA SRRRRDEIVSERGAISRLPSRSPSPSTRVSRCQSF-AERSPAVPLPRPIVRP---HVTST
Os_YODA1 SCRRGNDSSVEKSCRSTTVSRPTSPSKEVSRCQSFSADRPHAHPLPIPGVRP---PVTRT
Os_YODA2 NRRHSKDPTAEKGCWSTAQSRSASPSKEVSRCQSFAAARAHAQPLPLPRSRA---MVART
Pt_YODA1 ----------
Pt_YODA2 SRRCCKDTLSERVSLSRVPSRSPSPSTHVSRCQSF-AERPQAQPLPLPLPGVPHTKIGRC
Pp_YODA NSPYSLDLASEAGSGSGVGSRSPSPSSPIHNSKRR---HSQGQPLPLPCAPPTKSVGRAY
At1g53570 TPRCSREFAGASSAFSGFDSDSTE----------KKGHPLPRPLLSP--------
At5g66850 RHLTDDDVLLGERRASTSSSTFDSGLTRSPSAFTAVPRSPSAVPLPLPLPLP--------
At1g09000 ----------
At1g54960 ----------
At3g06030 ----------
At_YODA DSGMNGSQRPGLDANLKP-SWLPLPKPHGATSIPDNTGAEPDFATASVSSGSSVG--DIP
Os_YODA1 VSDITES-KPILEKRGKPPLLLPLPKPNRPPRRHGNSEVVSEIVVASPSSN-CSDSDDHG
Os_YODA2 ASDITES-KVVLEKRGKG-QQLPLPTTNWVKERPETTEPVAELSTASISSHGSIDSDDPG
Pt_YODA1 ----------
Pt_YODA2 DSGISASVKPGLDGGGKPLHLLPLPRPGHVLNRLDQADTAGDLATASVSSDSSIDSDDLP
Pp_YODA SEVQLPNVQNQWGSLLHHAASMPLPSPAQVGLMDNGDYPSGSVSSASSLGSVEVDPRQQG
At1g53570 ----------VSIHHQDHVSGSTSGSTSVSSVS-SSG---------
At5g66850 ----------EVAGIRNAANARGLDDRDRDPERLISDRTSSG---------
At1g09000 ----------
At1g54960 ----------
At3g06030 ----------
At_YODA SDSLLSPLAS--DCENGNRTPVN----ISSRDQS-MHSNKNSA-EMFKP--VPNKNRILS
Os_YODA1 DSQLQSPVGN--DAENATLVTLKNKSSNARKECPGPITAKNMK-EIHRPANQVHGSHILS
Os_YODA2 DLRLQGPVAN--DTDNVAKVATTGNSSVVHKECSSAITRKGTK-EVTMPTNAFLSNQILS
Pt_YODA1 ----------MQQDQSPIVNKKNSI-ETLKPANLPVNNQILP
Pt_YODA2 DSRVLSPLTS--DYENGNRTAVNSPPSVMRQDQSPIINRKNSR-ETLKHANLPANNQTLS
Pp_YODA RASHLRPLQEPEPPEQLRPVPISRMSHVSAASRGGVLRHEPVRSENSRSTSLFPGSYPPV
At1g53570 ----------SADDQSQLVASRGRG----------DVKFNVA
At5g66850 ----------PPLTSVNGGFARDSRK--------ATENSSYQD
At1g09000 ----------
At1g54960 ----------
At3g06030 ----------
At_YODA ASPRR-RPLGTHVKNLQIPQRDLVLCSAPDSLLSSPSRSPMRSFIPDQVSNHGLLISKPY
Os_YODA1 TSPRG-VAADSYQSNLQNPRP-LVLDSAPNSLMSSPSRSPRR-ICPDHIPTSAFWAVKPH
Os_YODA2 TSPRGTVVADSYQSNLQNSRK-VVLDSAPNSVMSSPSRSPRI-LCPDQIPSSAFWAVKPH
Pt_YODA1 TPPKR-AIFSSQVQNLQIPHR-GAFFSAPDSSLSSP-RSPMRAFGTEQVINNSFWTGKTY
Pt_YODA2 TPPKR-AIFSSQVQNLQIPHR-VAFFSAPDSSMSSPSRSPMRAFGTEQVINNGFWAGKTY

```
Pp_YODA         TSPTSRVWPPKWPSPLKISQS------QPDSSKSSPVRSPRY------VSQHQRNGMEPT
At1g53570       AAPRS-------------PER----VSPKAATITTRPTSPRH------------------
At5g66850       FSPRNRNGYWVNIPTMSAPTS------PYMSPVPSPQRKSTGHDLPFFYLPPKSNQAWSA
At1g09000       ----------------------------MQDFFGSVRR----------------------
At1g54960       ------------------------------------------------------------
At3g06030       ----------------------------MQDILGSVRR----------------------

At_YODA         SDVSLLGSGQCSSPGSGYNSGNNSIGGDMATQLFWPQSRCSPECSPVPSPRMTSPGPSSR
Os_YODA1        TDVTFVGSGQCSSPGSGQTSGHNSVGGDMLAQLFWQPSRSSPECSPIPSPRMTSPGPSSR
Os_YODA2        TDVTFVGSAQCSSPGSGQTSGHNSVGGDMLAQLFWQPSRGSPECSPIPSPRMTSPGPSSR
Pt_YODA1        SDIGLLGSGQCSSPGSGYNSGQNSIGGDMSGQLLWPNSRCSPECSPLPSPRMTSPGPSSR
Pt_YODA2        SDIGLLGSGQCSSPGSGYNSGQNSIGGDMSGQLLWPNSRCSPECSPLPSPRVISPGPSSR
Pp_YODA         VSGSVTG-----SPHSGNTSAYNSWQGDLAPLMAPQHHE---DRSLSPSPRLRSPGPSSR
At1g53570       ------------QRLSGVVSLESSTG-------RNDDGRSSSECHPLPRP-PTSPTSPSA
At5g66850       PDMPLDTSGLPPPAFYDITAFSTDNSPIHSPQPRSPRKQIRSPQPSRPSSPLHSVDSSAP
At1g09000       -----------SLVFRPSSDDDNQENQPPFPGVLADKITSCIRKSKIFIKPSFSPPPPAN
At1g54960       ------------MVFAKSQS-----------------------------------PPNNS
At3g06030       -----------SLVFRSSLAGDDGTSGGGLSG-FVGKINSSIRSSRIGLFSKPPPGLPAP

At_YODA         IQSGAVTPLHPRAGGSTTGSPTRRLDDNR-QQSHRLPLPPLLISNTCPFS-PTYSAATS-
Os_YODA1        VHSGSVSPLHPRSGGMAPESPTNRHDDGKKKQTHKLPLPPLSIS-HSSFH-PNNSTPTSP
Os_YODA2        VHSGSVSPLHPRAGGMAPESPTRRLDEGKRKQTHRLPLPPLSICNNSTFL-PNNSTPTSP
Pt_YODA1        IHSGAVTPLHHRAVGVTIESPTSCPDDGK-QQSHRLPLPPITTSNTCPFS-PTYSTTTS-
Pt_YODA2        IHSGAVTPLHPRAAGVTIESPTSRPDDGK-QQSHRLPLPPITISNTHPFS-PTYSASTS-
Pp_YODA         TTSAAVSPLHPRAVNRTSGGP-------------LPLPPLPTGNSSPYTSPSTPTMSRH
At1g53570       VHGSRIG-----------------------------------------------------
At5g66850       PRDSVSSPLHPRLSTDVTNGRRDCCNVHP------LPLPPGATCSSSSAA---------S
At1g09000       TVD-------------------------------------------------------
At1g54960       TVQ-------------------------------------------------------
At3g06030       RKE-------------------------------------------------------

At_YODA         PSVPRSPARAEATVSPGSRWKKGRLLGMGSFGHVYLGFNSESGEMCAMKEVTLCS---DD
Os_YODA1        ISVPRSPGRTENPPSPGSRWKKGKLIGRGTFGHVYVGFNSDSGEMCAMKEVTLFL---DD
Os_YODA2        ISH--SPGRVENPTSPGSRWKKGKLVGRGTFGHVYIGFNSDKGEMCAMKEVTLFS---DD
Pt_YODA1        PSVPRSPNRMENPTSPGSRWKKGRLLGRGSFGDVYLGLNSESGELCTMKEVTLFS---DD
Pt_YODA2        PSVPRSPSRMENPTSSGTRWQKGRMLGRGSFGDVYLGFNRERGEMCAMKEVTLFS---DD
Pp_YODA         LSSTLERSRSGESPVSSTKWQKGKLLGSGTFGNVYVGFNNDNGGFCAMKEVLLVS---DD
At1g53570       -------GGYETSPSGFSTWKKGKFLGSGTFGQVYLGFNSEKGKMCAIKEVKVIS---DD
At5g66850       VPSPQAPLKLDSFPMN-SQWKKGKLIGRGTFGSVYVASNSETGALCAMKEVELFP---DD
At1g09000       -------------MAPPISWRKGQLIGRGAFGTVYMGMNLDSGELLAVKQVLIAANFASK
At1g54960       -------------IKPPIRWRKGQLIGRGAFGTVYMGMNLDSGELLAVKQVLITSNCASK
At3g06030       -------------EAPSIRWRKGELIGCGAFGRVYMGMNLDSGELLAIKQVLIAPSSASK

At_YODA         PKSRESAQQLGQEISVLSRLRHQNIVQYYGSETVDDKLYIYLEYVSGGSIYKLLQEYG-Q
Os_YODA1        PKSKESAKQLGQEISLLSRLQHPNIVQYYGSETVDDKLYIYLEYVSGGSIHKLLQEYG-Q
Os_YODA2        PKSKESAKQLCQEILLLNRLQHPNIVRYYGSEMVDDKLYIYLEYVSGGSIHKLLQEYG-Q
Pt_YODA1        AKSKESAQQLGQEIMLLSRLRHPNIVQYYGSETVEDKLYIYLEYVSGGSIYKLLQEYG-Q
Pt_YODA2        AKSKESAQQLGQEIGLLSRLRHPNIVQYYGSETVDDKLYIYLEYVSGGSIYKLLQEYG-Q
Pp_YODA         HKSKESVKQLGQEISLLSKLRHENIVQYIGTETLEDRLYIYLEYVSGGSIHKLLQEYG-A
At1g53570       QTSKECLKQLNQEINLLNQLCHPNIVQYYGSELSEETLSVYLEYVSGGSIHKLLKDYG-S
At5g66850       PKSAECIKQLEQEIKLLSNLQHPNIVQYFGSETVEDRFFIYLEYVHPGSINKYIRDHCGT
At1g09000       EKTQAHIQELEEEVKLLKNLSHPNIVRYLGTVREDDTLNILLEFVPGGSISSLLEKFG-P
At1g54960       EKTQAHIQELEEEVKLLKNLSHPNIVRYLGTVREDETLNILLEFVPGGSISSLLEKFG-A
At3g06030       EKTQGHIRELEEEVQLLKNLSHPNIVRYLGTVRESDSLNILMEFVPGGSISSLLEKFG-S
```

```
At_YODA          FGENAIRNYTQQILSGLAYLHAKNTVHRDIKGANILVDPHGRVKVADFGMAKHITAQSGP
Os_YODA1         LGEQAIRSYTQQILSGLAYLHAKNTVHRDIKGANILVDPSGRVKLADFGMAKHINGQQCP
Os_YODA2         FGEPAIRSYTKQILLGLAYLHAKNTVHRDIKGANILVDPNGRVKLADFGMAKHINGQQCA
Pt_YODA1         FGEIAIRSYTQQILSGLAYLHAKKTVHRDIKGANILVDPTGRVKLADFGMAKHISGQSCP
Pt_YODA2         FGEIAIRSYTQQILRGLAYLHAKKTVHRDIKGANILVDPTGRVKLADFGMAKHISGQSCP
Pp_YODA          FKEPVVRNYTRQILSGLAYLHNQNTVHRDIKGANILVDTNGMVKLADFGMAKHISAQSFL
At1g53570        FTEPVIQNYTRQILAGLAYLHGRNTVHRDIKGANILVDPNGEIKLADFGMAKHVTAFSTM
At5g66850        MTESVVRNFTRHILSGLAYLHNKKTVHRDIKGANLLVDASGVVKLADFGMAKHLTGQRAD
At1g09000        FPESVVRTYTRQLLLGLEYLHNHAIMHRDIKGANILVDNKGCIKLADFGASKQVAELATM
At1g54960        FPESVVRTYTNQLLLGLEYLHNHAIMHRDIKGANILVDNQGCIKLADFGASKQVAELATI
At3g06030        FPEPVIIMYTKQLLLGLEYLHNNGIMHRDIKGANILVDNKGCIRLADFGASKKVVELATV

At_YODA          ---LSFKGSPYWMAPEV-----IKNSN-GSNLAVDIWSLGCTVLEMATTKPPWS-QYEGV
Os_YODA1         ---FSFKGSPYWMAPEV-----IKNSN-GCNLAVDIWSLGCTVLEMATSKPPWS-QYEGI
Os_YODA2         ---FSFKGSPYWMAPEV-----IKNSN-GCNLAVDIWSLGCTVLEMATSKPPWS-QYEGI
Pt_YODA1         ---FSFRGSPYWMAPEV-----IKNSN-GCNLAVDIWSLGCTVLEMATTKPPWS-QYEGV
Pt_YODA2         ---LSFKGSPYWMAPEV-----IKNSN-GCNLAVDIWSLGCTVLEMATTKPPWS-QYEGV
Pp_YODA          ---QSFKGSPYWMAPEVRVNCIITSTD-WYDLAVDIWSLGCTVLEMLTTKPPWN-QYEGV
At1g53570        ---LSFKGSPYWMAPEV-----VMSQN-GYTHAVDIWSLGCTILEMATSKPPWS-QFEGV
At5g66850        ---LSLKGSPYWMAPELMQAVMQKDSNPDLAFAVDIWSLGCTIIEMFTGKPPWS-EFEGA
At1g09000        TGAKSMKGTPYWMAPEV-----ILQTG--HSFSADIWSVGCTVIEMVTGKAPWSQQYKEV
At1g54960        SGAKSMKGTPYWMAPEV-----ILQTG--HSFSADIWSVGCTVIEMVTGKAPWSQQYKEI
At3g06030        NGAKSMKGTPYWMAPEV-----ILQTG--HSFSADIWSVGCTVIEMATGKPPWSEQYQQF

At_YODA          PAMFKIGNSKELPDIPDHLSEEGKDFVRKCLQRNPANRPTAAQLLDHAFV
Os_YODA1         AAMFKIGNSKELPPIPDHLSEPGKDFIRKCLQRDPSQRPTAMELLQHPFV
Os_YODA2         AAVFKIGNSKELPPIPDHLSEEGRDFIRQCLQRNPSSRPTAVDLLQHSFI
Pt_YODA1         PAMFKIGNSKELPEIPDHLSDDGKDFVRQCLQRNPSHRPTAAQLLDHPFV
Pt_YODA2         PAMFKIGNSKELPEIPDNLSDDGKDFVRQCLQRNLSHRPTAAQLLEHPFV
Pp_YODA          AAMFKIGNSKELPVIPNTLSRTGREFVRLCLQRDPAQRPTAAQLLEHPFV
At1g53570        AAIFKIGNSKDTPEIPDHLSNDAKNFIRLCLQRNPTVRPTASQLLEHPFL
At5g66850        AAMFKVMR--DSPPIPESMSPEGKDFLRLCFQRNPAERPTASMLLEHRFL
At1g09000        AAIFFIGTTKSHPPIPDTLSSDAKDFLLKCLQEVPNLRPTASELLKHPFV
At1g54960        AAIFHIGTTKSHPPIPDNISSDANDFLLKCLQQEPNLRPTASELLKHPFV
At3g06030        AAVLHIGRTKAHPPIPEDLSPEAKDFLMKCLHKEPSLRLSATELLQHPFV
```

Figure 1-S3. Amino acid alignment of the N-terminal end of *YODA* and related genes presented in Figure 1-6C. Sequence alignment is of the N-terminal end, including the regulatory region unique to *YODA* (yellow; Figure 4C) though the conserved kinase domain (gray). YODA is a member of the A2 clade of MAPK kinase kinase proteins in *Arabidopsis* (Ichimura et al. 2002). This clade also contains At1g53570 and At1g66850. A3 is the clade sister to clade A2 in the MKKK-wide phylogeny, and it is composed of At1g09000, At1g54960, and At3g06030 (Ichimura et al. 2002). None of the *Arabidopsis* genes closely related to *YODA* contain complete sequence in the *YODA* N-terminal regulatory domain.

Table 1-S1.

SPCH-MUTE-FAMA tree	Gene name/ID
At SPCH	At5g53210
At MUTE	At3g06120
At FAMA	At3g24140
Pt SPCH1	fgenesh4_pm.C_LG_XII000063
Pt SPCH2	gw1.197.14.1
Os SPCH1	Os06g33450
Os SPCH2	Os02g15760
Pt MUTE	gw1.VIII.81.1
Os MUTE	Os05g51820
Pt FAMA1	gw1.I.3509.1
Pt FAMA2	gw1.XVII.460.1
Os FAMA	Os05g50900
Pp FAMA1	estExt_gwp_gw1.C_710125
Pp FAMA2	e_gw1.519.9.1
outgroup	At4g01460
outgroup	At5g46690

TMM tree	Gene name/ID
At TMM	At1g80080
Pt TMM	eugene3.00970008
Os TMM	Os01g43440
Pp TMM	gw1.1.550.1
At outgroup	At2g42800
Pt outgroup	gw1.77.132.1
Os outgroup	Os01g52880
Os outgroup	Os05g45430
At outgroup	At4g28560
Pt outgroup	gw1.V.1354.1
Os outgroup	Os06g42220
Pt outgroup	gw1.VIII.2940.1
Os outgroup	Os01g02060
Pp outgroup	gw1.49.40.1
Pp outgroup	gw1.387.50.1
Pp outgroup	gw1.184.16.1

YODA N-term regulatory region	Gene name/ID
At YODA	At1g63700
Pt YODA1	estExt_Genewise1_v1.C_LG_I8903
Pt YODA2	FGENESH4_PM.C_LG_III000419
Os YODA1	Os04g47240
Os YODA2	Os02g44642
Pp YODA	gw1.44.79.1
At YODA outgroup (supplement)	At1g53570
At YODA outgroup (supplement)	At5g66850
At YODA outgroup (supplement)	At1g09000
At YODA outgroup (supplement)	At1g54960
At YODA outgroup (supplement)	At3g06030

Genome sources	
Populus trichocarpa v1.1	Poplar
TAIR	Arabidopsis
GRAMENE	Rice
Physcomitrella patens v1.1	Moss

Table 1-S1. Gene names and IDs used for phylogenetic analyses.

Chapter 2.

Molecular profiling of meristemoids reveals new component of asymmetric cell division and commonalities among stem-cell populations in *Arabidopsis*.

This chapter was previously published as:

Lynn Jo Pillitteri*, Kylee M. Peterson*, Robin J. Horst*, and Keiko U. Torii (2011) *Plant Cell* 23: 3260-3275.

* These authors contributed equally to this work.

Introduction

Plant stem cells continuously produce new organs to sustain indeterminate growth. To this end, these cell populations must ensure stem-cell survival through self-renewal and produce daughter cells destined for specialization. The two major stem-cell niches, the root and shoot apical meristems (RAMs and SAMs, respectively), have provided the majority of our current knowledge of plant stem cells (Clark et al., 1993; Di Laurenzio et al., 1996; Leibfried et al., 2005; Sarkar et al., 2007). However, other cell types outside the meristems, such as procambium cells and stomatal meristemoids, share distinct properties with meristematic stem cells and are likely to provide additional insight into cell self-renewal in plants (Fisher and Turner, 2007; Pillitteri et al., 2007a). Meristemoids represent a transitional state in the cell lineage producing stomata, valves on the plant epidermis (Figure 2-1A) (Sachs, 1991). Specifically, meristemoids have transient self-renewing capacity, and their differentiation can be blocked in the absence of a key regulator (Bergmann and Sack, 2007; Pillitteri and Torii, 2007).

A central mechanism for stem-cell maintenance and the generation of cellular diversity in both plants and animals is through asymmetric cell division, which ensures that the two daughter cells maintain separate identities (Abrash and Bergmann, 2009; Fichelson et al., 2009; Menke and Scheres, 2009). Asymmetric cell division during development can occur through signals from surrounding neighbors (extrinsic control) or, alternatively, intrinsic polarity within the cell can trigger partitioning of cell fate determinants (intrinsic control) (Abrash and Bergmann, 2009). Due to the tractability and accessibility of the epidermis, stomatal development has emerged as a model to study asymmetric division and cellular self-renewal. In *Arabidopsis thaliana*, stomatal development initiates from a subset of protodermal cells, termed meristemoid mother cells (MMCs) (Figure 2-1A). An MMC undergoes an asymmetric cell division that creates a stomatal precursor called a meristemoid. The meristemoid reiterates several rounds of asymmetric cell divisions, producing neighboring nonstomatal cells (stomatal lineage ground cells, SLGCs) prior to differentiating into a guard mother cell (GMC). The GMC undergoes a single symmetric division and terminally differentiates into a set of paired guard cells (GCs) that constitute a stoma (Bergmann and Sack, 2007; Peterson et al., 2010) (Figure 2-1A).

Progression through the stomatal lineage is specified through the combinatorial activities and heterodimerization of the basic helix-loop-helix (bHLH) transcription factors SPEECHLESS (SPCH), MUTE, FAMA, SCREAM (SCRM), and SCRM2 (Ohashi-Ito and Bergmann, 2006; MacAlister et al., 2007; Pillitteri et al., 2007a; Kanaoka et al., 2008). *SPCH*, *MUTE*, and *FAMA* act as molecular switches for cell state transitions: *SPCH* drives the initial asymmetric division of MMCs and entry into the stomatal lineage, whereas *MUTE* triggers the differentiation of meristemoids into GMCs (MacAlister et al., 2007; Pillitteri et al., 2007a, 2008). *FAMA* is required for the proper

differentiation of GCs (Ohashi-Ito and Bergmann, 2006). Knockout mutations in *SPCH*, *MUTE*, or *FAMA* result in a complete block of stomatal lineage progression at their respective points of action (Figure 2-1A). SCRM and SCRM2, two paralogous and partially redundant bHLH proteins, integrate the three cell state transitional events for stomatal differentiation by forming heterodimers with SPCH, MUTE, and FAMA (Kanaoka et al., 2008) (Figure 2-1A).

The unique ability of a meristemoid to reiterate asymmetric cell divisions makes this cell type an attractive system to understand the molecular and cellular basis of self-maintenance and differentiation in plants. To date, the molecular character of meristemoids remains unclear due to the transient nature and low density of meristemoids on the epidermal surface. Here, we report genome-wide transcriptional profiling of meristemoids through targeted enrichment of stomatal cell states enabled through combinatorial, synthetic mutants of stomatal differentiation. Pairwise comparisons of transcriptomes among mutant seedlings that develop an epidermis composed solely of pavement cells, meristemoids and SLGCs, or GCs identified new genes that are highly and specifically expressed in the stomatal lineage (Figure 2-1B). Meristemoid-enriched gene products include those involved in cell–cell signaling, hormone metabolism and signaling, and the cell cycle, as well as cell polarity components associated with the meristemoid state. A survey of commonly upregulated genes in meristemoids and other stem-cell populations emphasizes a broad role for *ERECTA* family genes in cell proliferation and differentiation within the SAM and stomatal cell lineages. Through this transcriptome analysis, we identify a protein, POLAR, that is polarly localized and unequally segregated during asymmetric cell divisions of the stomatal lineage. Combined, our study enables the isolation of a unique gene set actively expressed during the transient meristemoid cell state and provides high-quality data for building

a gene regulatory network. Furthermore, our study reveals the molecular character of the meristemoid and illuminates common themes in gene expression among plant stem cells.

Results

Transcriptional profiling of stomatal lineage cells enabled by combinatorial mutations in stomatal differentiation genes

To obtain gene expression profiles reflecting cell-state transitions in the stomatal lineage, we took advantage of mutant resources identified in our previous studies, which enable targeted enrichment of pavement cells, meristemoids, and GCs. To assess the transcriptional profile of meristemoids, we used a synthetic mutant carrying both *scrm-D* and *mute*. Meristemoids of *mute* mutants never differentiate into GMCs or transdifferentiate into pavement cells (Pillitteri et al., 2007a). In contrast, the gain-of-function *scrm-D* mutation triggers constitutive initiation and differentiation of the stomatal lineage, resulting in the production of stomata over the entire epidermis (Kanaoka et al., 2008) (Figure 2-1B). Therefore, the combined *scrm-D mute* mutations produce a synthetic phenotype of an epidermis composed almost entirely of meristemoids (Figure 2-1B). The *scrm-D* mutant was used to represent the guard-cell state (Figure 2-1B). Finally, the *spch* mutant, characterized by an epidermis composed solely of pavement cells, was used as a genotype devoid of any stomatal-cell-lineage-derived transcripts (MacAlister et al., 2007; Pillitteri et al., 2007a) (Figure 2-1B). Pairwise comparisons of transcriptomes among three biological replicates of *spch*, *scrm-D mute* double mutants, *scrm-D*, and the wild type were performed using Affymetrix ATH1 expression arrays in an effort to maximize the signal-to-noise ratio between molecular signatures associated with each cell state in a way that was not previously possible

(Figure 2-1B). To minimize secondary effects due to the absence or constitutive differentiation of stomata, very young seedlings five days after germination were used for profiling (Figure 2-S1). To minimize diurnal effects on gene expression (Harmer et al., 2000), all seedlings were harvested in the middle of the light period.

Gene expression clustering reveals dramatic changes in transcript accumulation during cell state transition

To assess genotype- or cell-type-specific trends in gene expression, we performed a K-means cluster analysis using normalized expression values from each individual replicate of four different genotypes (Figure 2-2). Genes were prefiltered using the variance filter implemented in the MeV software to remove those that showed no differential expression across the genotypes (Saeed et al., 2003). This yielded 2000 genes with the highest variance among the different arrays. Five clusters (I to V; Figure 2-2) were built from a maximum of 200 iterations using a Pearson correlation distance matrix. A suite of genes specifically enriched in one or two genotypes was identified (see Supplementary Datasets 2-1 to 2-3).

Gene cluster I showed robust change associated with the meristemoid state: 552 genes are represented in this cluster, with significant upregulation in the *scrm-D mute* background, but less in *scrm-D* (Figure 2-2; see Supplementary Dataset 2-1). Within this group are all genes known to be highly expressed in meristemoids, including *TOO MANY MOUTHS (TMM)*, *ERECTA- LIKE1 (ERL1)*, *ERL2*, *STOMATAL DENSITY AND DISTRIBUTION1 (SDD1)*, *SPCH*, *MUTE*, *SCRM*, *SCRM2*, *BREAKING OF ASYMMETRY IN THE STOMATAL LINEAGE (BASL)*, and *EPIDERMAL PATTERNING FACTOR2 (EPF2)*.

Gene cluster V includes genes highly upregulated in GCs (*scrm-D*) and to a lesser extent in *scrm-D mute*: 407 genes are represented in this cluster (Figure 2-2; see

Supplementary Dataset 2-2). These likely represent genes implicated in guard cell development or function, as well as GMC differentiation. For instance, *KAT1* and *KAT2*, two genes encoding GC-specific potassium channel proteins (Schachtman et al., 1992; Nakamura et al., 1995), are represented in this cluster (Figure 2-2). *FAMA* and *EPF1*, a signaling peptide expressed in late meristemoids and GMCs to enforce stomatal spacing (Hara et al., 2007), are also in this cluster (Figure 2-2). No *FAMA* transcripts were detected in the *scrm-D mute* population, supporting the previous finding that in the absence of *MUTE*, stomatal precursor cells will not progress to GMC and express *FAMA* (Pillitteri et al., 2007a). In contrast, the presence of *EPF1* transcripts in *scrm-D mute* implies that *EPF1* is not directly controlled by *MUTE*. Genes in clusters I and V show lower expression in the wild type and fall under the detection limit in *spch* (Figure 2-2), consistent with the absence of stomatal lineage cells in *spch*. Combined, these results emphasize that our strategy enabled us to document robust changes in gene expression associated with specific stomatal precursor cell states, which are normally asynchronously transient and cannot be effectively captured in the wild-type background.

Gene categories overrepresented in each epidermal cell state

To identify genes with statistically significant enrichment in each mutant background, we used a significance of microarrays analysis using a Q-value cutoff of <0.005 and a two- or threefold-change cutoff as a threshold for deregulation (see Methods). Using these stringent criteria, we identified gene signatures for each background compared with the wild type (see Supplementary Datasets 2-4 to 2-6). We first performed overrepresentation analysis (Usadel et al., 2006) to determine the gene categories enriched in each background (Figure 2-S2). No overrepresentation of genes involved in photosynthesis, carbohydrate, or central tricarboxylic acid metabolism was

observed. A moderate overrepresentation of stress-related genes was found in *spch* and *scrm-D*, but not in *scrm-D mute*. This indicates that the alteration of epidermal phenotype did not significantly affect overall metabolism or stress level of the seedlings. Whereas cell-wall biosynthesis genes were mainly upregulated in *scrm-D*, they were downregulated in *scrm-D mute*. This may reflect the increased cell-wall deposition at the stomatal pore in *scrm-D* and the low degree of cell expansion in *scrm-D mute*.

We also found that signaling components, such as receptor kinases, were greatly enriched in the *scrm-D mute* background but not in *spch* or *scrm-D*. This is in accordance with the well-documented findings that cell-cell signals from meristemoids via secreted ligands, membrane receptors, and mitogen-activated kinase cascades are critical for proper stomatal patterning and spacing (Geisler et al., 2000; Bergmann et al., 2004; Shpak et al., 2005; Hara et al., 2007, 2009; Hunt and Gray, 2009).

Molecular profiling reveals genes highly enriched in meristemoids

The significance of microarray analysis identified 527 genes significantly deregulated (277 up-/250 downregulated) in the *scrm-D mute* background. The 2-kb upstream promoter regions of this group of genes were enriched in the E-box (CANNTG) cis-regulatory sequence known to be bound by bHLH DNA-binding domains ($P = 10^{-10}$; Athena data mining tool; O'Connor et al., 2005). This is in contrast with the 74 genes deregulated (39 up-/35 downregulated) in the *spch* background, where no significant enrichment of the E-box element was detected. The finding is consistent with the notion that combinatorial actions of master regulatory bHLH proteins specify meristemoid gene expression.

It is important to address whether the arrested meristemoid-like cells constituting the epidermis of *scrm-D mute* share molecular characteristics of

meristemoids, as indicated by the cluster analysis (Figure 2-2). We first introduced representative markers of early and late meristemoids, *EPF2pro::erGFP* (endoplasmic-reticulum-bound Green Fluorescent Protein) and *MUTEpro::GFP*, respectively, to *scrm-D mute* (Figure 2-3). Indeed, both reporter constructs showed high and nearly constitutive expression in the epidermis of *scrm-D mute* (Figure 2-3B,D).

Consistent with the reporter activity, among the top signatures (based on fold increase) in *scrm-D mute* are genes known to be highly expressed in meristemoids and regulate meristemoid development: *MUTE* (79-fold increase), *EPF2* (39-fold increase), *TMM* (22-fold increase), *ERL1* (22-fold increase), *SPCH* (20-fold increase), *BASL* (17-fold increase), and *ERL2* (7.5-fold increase) (Figure 2-3E). A quantitative RT-PCR (qRT-PCR) analysis of these genes further verified the microarray results (Figure 2-3F). It is worth noting that *mute* displays the same phenotype as a transcriptional null allele, *mute-2* (Pillitteri et al., 2008), but accumulates mis-spliced transcripts with premature termination codons (Figure 2-S3) (Pillitteri et al., 2008). Overall, the *scrm-D mute* epidermal cell population has the molecular characteristics of meristemoids, which validates our approach and provides the basis for further characterization of genes enriched in this cell population.

Reporter expression and cellular localization of meristemoid-enriched genes

To gain insight into the roles of meristemoid-enriched genes, we chose 14 genes previously uncharacterized in relation to stomatal development and produced reporter constructs (Figure 2-4). We designed constructs containing up to 2.5 kb of the upstream regulatory sequence for each gene driving the expression of the full-length coding region fused to GFP. Alternatively, for those genes whose translational fusions did not result in detectable GFP signals, transcriptional fusion constructs were produced. They include the following: At1g26600 (*CLE9*; 30-fold), At4g31805 (misannotated as WRKY

family transcription factor [see below]; 21-fold), At5g13220 (*JAZ10*; 15-fold), At2g40670 (*ARR16*; 14-fold), At3g17640 (leucine-rich repeat protein; 9-fold), and At5g62210 (*EMBRYO-SPECIFIC PROTEIN3 [ESP3]*; >20-fold in *scrm-D mute* compared with *spch*) (Table 2-S1).

Of the 14 reporter constructs, 10 showed expression patterns in the stomatal lineage (Figure 2-4A to K). Some are components of signaling pathways that have not been assigned to stomatal development. JAZ10-GFP was found in the nucleus through all transitional states of the stomatal lineage but not in fully differentiated GCs (Figure 2-4A). *JAZ10* acts as a downstream repressor of jasmonate signaling (Chung and Howe, 2009). ARR16-GFP showed strong and specific expression in meristemoids (Figure 2-4B). *ARR16* belongs to the A-type Arabidopsis response regulator (ARR) family acting downstream in cytokinin signaling (Kiba et al., 2002). Interestingly, a gene coding for the cytokinin catabolic enzyme At4g29740 (*CKX4*) (Werner et al., 2006) is also enriched in *scrm-D mute*, and its transcriptional GFP fusion exhibited expression in meristemoids (Figure 2-4C). GFP translational fusions of At1g33930, a protein with a GTP binding motif, and At3g17640 are also strongly detected in the cytoplasm of meristemoids (Figure 2-4D,E). *CLE9*, one of the CLE peptide genes, and At5g07280 (*EMS1*), an LRR–receptor-like kinase (RLK) regulating microsporogenesis (Zhao et al., 2002; Jun et al., 2010), showed strong expression in meristemoids and GMCs as GFP transcriptional fusions (Figure 2-4F,G).

Three proteins displayed unique expression patterns. ESP3, which is predicted to be membrane-anchored via glycosylphosphatidylinositol (Borner et al., 2002), displayed patchy plasma-membrane localization across the epidermis, including pavement cells, but showed strong stomatal-lineage-specific expression on internal membranes (Figure 2-4H). At1g52910 (unknown protein) exhibited a characteristic, vesicular localization

and strong accumulation at the cell plate of dividing meristemoids and GMCs (Figure 2-4I, arrow), suggesting that this protein may be involved in the cytokinesis of stomatal lineage cells. Lastly, At4g31805 exhibited an intriguing localization, with strong GFP signal detected in the cytoplasm of meristemoids and the cell periphery in SLGCs (Figure 2-4J). Strikingly, At4g31805-GFP was localized polarly in SLGCs, distal to the newly divided meristemoid and parallel to the division plane (Figure 2-4K, asterisks and arrowheads). Based on the dynamic behavior of this protein during asymmetric divisions revealed by time-lapse imaging (see below), we named At4g31805 *POLAR* (for *POLAR LOCALIZATION DURING ASYMMETRIC DIVISION AND REDISTRIBUTION*).

Cellular dynamics of POLAR protein, which is polarly localized and unequally segregated during asymmetric cell division of the stomatal lineage

The polar localization of POLAR in SLGCs (Figure 2-4J,K) resembles that of BASL (Dong et al., 2009). Unlike GFP-BASL, which remains in the nucleus in meristemoids (Dong et al., 2009), however, POLAR-GFP displays no nuclear localization in any stage of stomatal development. Based on publicly available data sets, *POLAR* and *BASL* display little or no detectable expression in other asymmetrically dividing tissues, such as roots (Figure 2-S4). To resolve the cellular dynamics of POLAR during asymmetric divisions of the stomatal lineage, we developed a time-lapse imaging technique (Figure 2-5 and Supplementary Movie 2-1). Mature embryos were dissected at germination, and cotyledons were subjected to time-lapse imaging at 30-minute intervals over 72 hours to capture the real-time movement of this protein in relation to cell division during stomatal development (see Methods). During the process of imaging, cotyledons developed normally and eventually formed stomata.

The time-lapse imaging revealed that during germination, POLAR-GFP initially appears in a subset of protodermal cells, which are likely MMCs, both in the cytosol and at the cell periphery (Figure 2-5A). Approximately 2 hours before each asymmetric division occurs, POLAR-GFP becomes dynamically localized at the cell cortex distal to the division plane (Figure 2-5B,E,H). Clear polar localization of POLAR distal to the division plane was observed in all asymmetrically dividing cells of the stomatal lineage. After asymmetric cell division, POLAR-GFP is upregulated in only one of the daughter cells, which is predictive of the cell that will continue to divide asymmetrically. When the larger daughter cell (SLGC) retains POLAR expression, localization is directed to the periphery of the cell adjoining the adjacent meristemoid or GMC, correlated with a new satellite meristemoid being placed away from the existing one (see Supplementary Movie 2-1). Expression of POLAR appears to mark stomatal lineage cells with asymmetric division potential. Those protodermal cells that did not undergo asymmetric entry division but underwent pavement cell differentiation or, alternatively, differentiated into GMCs without asymmetric divisions immediately lost POLAR-GFP expression (Figure 2-S5). Likewise, once a meristemoid terminated asymmetric division and differentiated into a GMC, POLAR-GFP expression rapidly diminished (Figure 2-5I to K).

Asymmetric localization of POLAR is disrupted by loss of function in *BASL*

To understand whether a known regulator of asymmetric cell divisions in the stomatal cell lineage influences the polar localization of POLAR, we next investigated the cellular dynamics of POLAR-GFP in *basl* using time-lapse imaging (Figure 2-6 and Supplementary Movie 2-2). The *basl* mutant is characterized by a loss of physical asymmetry during division and concomitant loss of cell fate asymmetries, resulting in clustered stomata (Dong et al., 2009). Asymmetric localization of POLAR-GFP was

disrupted in the *basl* mutant background, which correlated with nonasymmetric divisions in GFP-expressing cells (Figure 2-6D,H). During the division of stomatal lineage cells in *basl*, both daughters retain expression of POLAR-GFP and both continue to divide (Figure 2-6E,H). This is in sharp contrast with the wild type, where only one cell from a division retains expression of POLAR-GFP and continues asymmetric division (Figure 2-5E to G). This supports the previous finding that cell fate asymmetries after division are absent in the *basl* background (Dong et al., 2009). The observed perturbation of POLAR-GFP localization in the *basl* background suggests that a functional BASL protein is required for POLAR-GFP asymmetry. Some cells in the *basl* background may also express POLAR-GFP without dividing for over 24 hours, a much longer duration than we observed in the wild type, which may indicate that the regulation of this protein is perturbed by the loss of *BASL*. The asymmetric division defects in *basl* are not fully penetrant; when meristemoids divide normally in *basl*, POLAR-GFP exhibits normal expression dynamics and polar localization (see Supplementary Movie 2-2).

Protein structure and loss-of-function allele of *POLAR*

POLAR codes for a 344-amino-acid protein of unknown function. Although it is annotated as a WRKY transcription factor (The Arabidopsis Information Resource: http://www.arabidopsis.org/), sequence analysis does not reveal any similarity to the WRKY domain or any other recognizable structural motifs or domains aside from its moderate similarity to an Arabidopsis myosin heavy chain-like protein (At5g10890: 29% identity/49% similarity). Secondary structure analysis predicts a predominantly alpha-helical structure with a coiled-coil region at the C terminus (Lupas et al., 1991). We did not identify homologs of *POLAR* outside of plants. To unravel the function of *POLAR*, we analyzed two T-DNA insertion alleles (SALK_112914 and SALK_142820). The RT-

PCR analysis detected no *POLAR* mRNA accumulation (Figure 2-S6), indicating that the insertions led to the loss or reduction of *POLAR* transcripts. These insertion lines did not reveal discernable growth or developmental phenotypes, suggesting that *POLAR* may be functionally redundant.

Meristemoid-enriched transcriptome includes specific members of core cell-cycle regulators

Meristemoids exhibit a rare capacity for reiterating asymmetric cell divisions, which is lost with the cell-state transition to GMC: GMCs instead undergo a symmetric division and differentiation. We therefore investigated whether cell-cycle regulatory genes were enriched in meristemoids. Our analysis revealed that seventeen core cell-cycle genes were represented in gene cluster I (Figures 2-2 and 2-7). Among them, we observed the reporter β-glucuronidase (GUS) expression of *CYCLIN DEPENDENT KINASE B2;1 (CDKB2;1), CYCB1;2*, and *CYCA2;3* (Figure 2-7A to C), all of which showed high promoter activity in stomatal lineage cells and higher levels of transcript accumulation in the *scrm-D mute* background compared with *spch* or *scrm-D* (Figure 2-7E,F). It has been reported that *ERECTA* family receptor kinases, which regulate organ growth and stomatal patterning, affect expression of cyclin A2 genes (Pillitteri et al., 2007b). We further examined the reporter activity of the remaining cyclin A2 genes and found that *CYCA2;2* is upregulated in *scrm-D mute* (Figure 2-7F) and *CYCA2;2pro:GUS* is expressed in a subset of meristemoids (Figure 2-7D), potentially those making a transition from asymmetric to symmetric division. Consistent with this result, *CYCA2;2* transcript is slightly enriched in *scrm-D mute* (Figure 2-7F).

Molecular signatures of meristemoids and meristems: commonalities and distinctions

Comparison of transcriptome profiles among different microarray experiments can identify unexpected molecular commonalities among specific cell types. Our meristemoid-enriched transcriptome provides the opportunity to compare the meristemoid state with the true stem-cell population in the SAM and RAM. For this purpose, we obtained publicly available transcriptome data derived from isolated cells expressing the SAM-specific markers *CLAVATA3* (*CLV3*; stem-cell domain), *WUSCHEL* (*WUS*; rib meristem), and *FILAMENTOUS FLOWER* (*FIL*; peripheral zone of the floral meristem) (Yadav et al., 2009) as well as data from two independent root-tip microarrays, termed root tips zone1 (Dinneny et al., 2008) and root tips (Sena et al., 2009). These data sets were compared with our meristemoid-enriched transcript (*scrm-D mute*) profile (Figure 2-8). Overrepresentation analysis showed that genes classified as transcription factors (RNA regulation of transcription) and receptor kinases are equally overrepresented in all stem-cell populations included in this meta-analysis (Figure 2-S7), highlighting the importance of transcriptional regulation and cell-cell signaling in maintenance and function of meristematic cells. Alternatively, these trends may reflect general characteristics of actively dividing cells.

We identified relatively little overlap in upregulated genes between *scrm-D mute* and individual SAM populations: 11, 3, and 6 genes for the FILpro, CLV3pro, and WUSpro arrays, respectively. This is far less than the overlap of the CLV3pro array compared with the WUSpro (149 genes) or FILpro arrays (119 genes) (Figure 2-8A). However, we identified 19 genes that are commonly upregulated in the all three SAM populations and *scrm-D mute* (Table 2-S2). As expected, several cell-division-related proteins were commonly upregulated: At3g19590 (*BUDDING UNINHIBITED BY*

BENZYMIDAZOL [BUB3.1]), At2g06510 (*REPLICATION PROTEIN A [RPA1A*]), and At1g20930 (*CDKB2;2*). The most notable among the commonly upregulated genes are two members of the *ERECTA* family of genes, *ERL1* and *ERL2*, which encode closely related LRR-RLKs (Shpak et al., 2004). Among all arrays, *ERL1* (At5g62230) shows eightfold to 22-fold upregulation and *ERL2* (At5g07180) a sevenfold to 35-fold increase. The finding underscores the known roles of the *ERECTA* family in promoting cell proliferation during shoot and flower development as well as restricting stomatal differentiation (Torii et al., 1996; Shpak et al., 2004, 2005; Pillitteri et al., 2007b; Hord et al., 2008). The LRR-RLK gene *EMS1* was also transcriptionally enriched in *scrm-D mute*, WUSpro, and FILpro (19-, 14-, and 24-fold, respectively). Enrichment in CLV3pro was statistically insignificant (Q = 0.017) but likely a false negative due to our stringent cutoffs. *EMS1* regulates anther cell differentiation (Zhao et al., 2002) by physical interaction with the small peptide *TAPETUM DETERMINANT1 (TPD1)* (Jia et al., 2008). A role for the TPD1/EMS1 ligand receptor system in the SAM has not been documented.

Comparison of *scrm-D mute* to transcriptional profiles from root tips revealed 12 genes commonly enriched in both root tip experiments and *scrm-D mute* (Figure 2-8B). Among them is *SCHIZORIZA (SCZ),* a heat-shock family transcription factor, which regulates asymmetric cell divisions in the root stem-cell niche (ten Hove et al., 2010). *SCZ* is also transcriptionally enriched in *scrm-D mute* and specifically expressed in meristemoids during stomatal development (Figure 2-8C to E), suggesting an exciting potential role for *SCZ* in asymmetric cell division in the stomatal lineage.

Discussion

Here, we present the transcriptome profile of the meristemoid, a transient and low-density proliferating stomatal precursor cell with stem-cell-like properties. Producing plants highly enriched in specific epidermal cell types (pavement cells, meristemoids, or GCs forming stomata) is possible using loss- and gain-of-function mutations in transcription factor genes (*SPCH*, *MUTE*, and *SCRM*) that act in a combinatorial and sequential manner to direct key cell-state transitional events within the stomatal cell lineages (MacAlister et al., 2007; Pillitteri et al., 2007a; Kanaoka et al., 2008). Cell-type-specific profiling has been widely performed using physical isolation of cells via laser capture microdissection (Kerk et al., 2003; Nakazono et al., 2003; Day et al., 2007) or protoplasting followed by fluorescence-activated cell sorting (FACS) (Iyer-Pascuzzi and Benfey, 2010). Protoplasting and FACS were used for the cell-type-specific expression profiling of the SAM used in our meta-analysis (Yadav et al., 2009). These approaches, while accurately isolating specific cell types of interest from wild-type plants, are expensive and include experimental procedures, such as tissue sectioning and protoplast isolation, that could affect gene expression. In addition, the FACS process relies on the availability of GFP markers that specifically and exclusively mark the cell types of interest.

Using *spch*, *scrm-D mute*, and *scrm-D* mutants, we were able to isolate RNA directly without further manipulations to perform pairwise transcriptome comparisons. Using mutant backgrounds invites two major questions: whether the identity of the cells enriched in these mutants is truly normal and whether physiological consequences due to the mutations deregulate transcripts unrelated to the enriched cell types. We did not find significant alterations in stress- or photosynthesis-related transcripts in the *scrm-D mute* population. Top signatures from this population include genes known to

be highly or specifically expressed in meristemoids and regulate stomatal development, including *TMM, SPCH, MUTE, EPF2, BASL, ERL1, ERL2, SCRM*, and *SCRM2* (Figure 2-3E). Furthermore, a new gene associated with cellular asymmetry of meristemoids was identified (Figures 2-4 and 2-5). Combined, we conclude that our approach is a simple yet powerful alternative to unravel molecular signatures associated with this transient cell state.

Coexpression programs (ATTED-II: http://atted.jp/) (Obayashi et al., 2011) cluster only two of our ten more closely analyzed genes, *POLAR* and *CLE9*, with known stomatal regulatory networks; the other eight genes identified in this analysis with stomatal-specific expression (Figure 2-4) were not predicted to be coexpressed with stomatal regulators. This inconsistency demonstrates the importance of continued transcriptome profiling in identifying coregulated genes and ultimately regulatory networks controlling developmental processes. Such discoveries provide a framework of how the molecular mechanisms driving cell division maintenance may be partially conserved among stem-cell populations.

Cytokinin signaling may play a universal role in stem-cell regulation

We found that the cytokinin response genes *ARR16*, *CKX4*, and *CLE9* are strongly upregulated in *scrm-D mute*, and their expression in stomatal lineage cells was confirmed by reporter analysis (Figure 2-4B,C,F). *CKX4* encodes a cytokinin oxidase, which degrades and thus inactivates cytokinin (Schmülling et al., 2003), and *ARR16* is a member of the A-type ARR family, which function as negative regulators of cytokinin signaling (Ren et al., 2009). ARR proteins act as downstream components of cytokinin signaling via a two-component signaling pathway, where cytokinin receptors, *ARABIDOPSIS HISTIDINE KINASE* (*AHK2*, *3*, and *4*) (Inoue et al., 2001; Yamada et al., 2001), relay a phosphate group to ARR proteins. *AHK3* has been shown to display

stomatal-lineage-specific epidermal expression (Stolz et al., 2011). Although T-DNA insertion lines of *ARR16* and *CKX4* did not confer discernible phenotypes, our finding implies that downregulation of cytokinin signaling and/or metabolism may play a role in the transition from proliferation to differentiation states of stomatal cell lineages.

Recently, a molecular link between cytokinin signaling and CLE peptide-mediated stem-cell maintenance has been reported. In the SAM, the *CLV3-WUS*–mediated signaling pathway directly represses the transcription of Type-A ARRs (*ARR5, ARR6, ARR7*, and *ARR15*), which are required for proper meristem function (Leibfried et al., 2005). Similarly, during root vascular development, one of the CLE peptides, CLE10, inhibits protoxylem differentiation by repressing Type-A ARRs (*ARR5* and *ARR6*) (Kondo et al., 2011). Based on the known roles of the CLE peptides in stem-cell maintenance and high expression of *CLE9* in meristemoids (Figure 2-4F) (Jun et al., 2010), it is attractive to hypothesize that CLE9 may act as a signaling molecule for meristemoid maintenance. Our results, in combination with the expanding evidence of crosstalk between cytokinin and CLE signaling in other stem-cell populations, suggest that integration of these two pathways might be a general mechanism to mediate broad hormone signals into local changes of gene expression and, ultimately, cell behavior.

Asymmetric division and segregation of cell fate in the stomatal lineage

Stomatal development serves as a model to study *de novo* asymmetric cell division in plants. BASL was the only known protein to be distributed unevenly during the asymmetric division of stomatal lineages (Dong et al., 2009). Our result adds a potential component to the BASL-mediated pathway. POLAR exhibits a striking, transient polar distribution during asymmetric cell division, and this polar localization requires functional BASL. POLAR possesses a coiled-coil motif, which is known to function in protein–protein interactions (Burkhard et al., 2001) and shows sequence

similarity to myosin heavy chain–like protein (Figure 2-S6). Based on the protein structure and localization patterns, it is tempting to hypothesize that POLAR associates with BASL, components of a complex that includes BASL, or cytoskeletal components to mark the site for rapid cell expansion, which facilitates the cell fate separation of SLGC from meristemoid. POLAR is evenly distributed at the cell periphery in meristemoids where BASL is known to be exclusively localized in the nucleus (Dong et al., 2009). This supports the idea that sequestering BASL in the nucleus is a mechanism preventing polar cell expansion, a process involving the dynamic relocation of POLAR to the cell periphery. The null allele of *POLAR* does not exhibit stomatal patterning defects like *basl*, implying the presence of functional redundancy. Indeed, the closest homolog of *POLAR*, At5g10890, which is not on the ATH1 expression array, is also highly expressed in *scrm-D mute* (88-fold upregulation), while being downregulated in *spch* (14- fold downregulation) (Figure 2-4M). The extent of up- and down-regulation of At5g10890 in the meristemoid-enriched populations is similar to that of *POLAR* (149-fold upregulation in *scrm-D mute*; 17-fold downregulation in *spch*) (Figure 2-4M).

It is known that the bHLH transcription factor *SPCH* initiates the meristemoid state together with *SCRM*/*SCRM2* (MacAlister et al., 2007; Pillitteri et al., 2007a; Kanaoka et al., 2008). Thus, genes driving asymmetric division of meristemoids may be downstream targets of these bHLH proteins. A coexpression gene network places both *BASL* and *POLAR* directly connected to *SPCH* and closely connected to *SCRM* (Figure 2-S8). Further experiments should address the possibility of direct induction.

Cell cycle regulators and stomatal cell divisions

Stomatal cell lineages undergo stereotypical cell divisions, from entry and amplifying divisions of meristemoids to the eventual symmetric divisions of GMCs. Our analysis revealed that promoter activity of *CDKB2;1*, *CYCB1;2*, *CYCA2;2*, and

CYCA2;3 was highly specific to stomatal lineage cells (Figure 2-7). Consistent with a possible direct role in stomatal development, all three cyclin genes examined have been shown to interact directly with *CDKB1;2* and *CDKB2;1* in yeast and in planta (Van Leene et al., 2011). A direct link between *CYCA2;3* and epidermal patterning was recently reported using ectopic overexpression analysis, where the ectopic overexpression of both *CYCA2;3* and *CYCB1;1* resulted in increased epidermal cell division similar to a *SPCH* overexpression phenotype (Boudolf et al., 2009). Single overexpression of either regulator failed to confer a clear phenotype, suggesting that they function together as a complex (Boudolf et al., 2009). However, *CYCD4*, which is known to regulate stomatal cell lineage divisions specifically in the hypocotyl (Kono et al., 2007), was not represented in our *scrm-D mute* transcriptome. Plants encode significantly more cell-cycle regulatory proteins than other eukaryotic organisms, presumably due to a sessile existence and subsequent need for plasticity during development. Our profiling and expression analysis add to the growing evidence supporting the idea that overlapping expression of plant core cell-cycle regulators is a key component driving cell division and differentiation (Engler et al., 2009).

Molecular commonalities among stem-cell populations in plants

A notable feature of plant development is indeterminate growth, where organogenesis persists throughout the life cycle via continual activity of stem cells in the SAM and RAM. Our meristemoid-enriched transcriptome provides the opportunity to investigate the similarities and differences of molecular constituents between transient and permanent stem-cell populations. The comparison of meristemoid versus SAM signatures reemphasizes the diverse roles of *ERECTA* family RLKs as regulators coordinating cell proliferation and differentiation during shoot and flower development as well as stomatal patterning (Torii et al., 1996; Shpak et al., 2003, 2004, 2005; Pillitteri et

al., 2007b; Hord et al., 2008). In the epidermis, *ERECTA* restricts asymmetric cell divisions, while *ERL1* and *ERL2* inhibit the differentiation of meristemoids to GMC and enforce stomatal spacing. Genetic studies suggest that ERECTA-family RLKs act together with TMM to perceive secreted signaling ligands EPF1 and EPF2 (Shpak et al., 2005; Hara et al., 2007, 2009; Hunt and Gray, 2009). Because *TMM*, *EPF1*, and *EPF2* are not expressed in the SAM, an exciting possibility is that the *ERECTA* family perceives other EPF-LIKE (EPFL) peptides via dimerization with other receptor partner(s). There are 11 *EPFL* genes in Arabidopsis (Hara et al., 2009; Rowe and Bergmann, 2010; Rychel et al., 2010), for seven of which the expression patterns and functions are yet unknown. The highlighted roles of LRR-RLKs, the *ERECTA* family, and *EMS1* as commonly upregulated SAM-meristemoid signatures underscore the importance of cell-cell signaling in orchestrating stem-cell proliferation and differentiation.

In contrast with the SAM, where stem-cell divisions are less organized, Arabidopsis RAM cells undergo stereotypical asymmetric divisions. Based on their positions relative to the quiescent center, each stem cell in the RAM gives rise to a defined cell file (Scheres, 2002). Our RAM versus meristemoid comparison revealed *SCZ* as a commonly upregulated gene. *SCZ* executes proper separation of cell fate in stem cells generating differentiated root tissues, including epidermis, cortex, endodermis, and root cap (ten Hove et al., 2010). The molecular mechanism by which *SCZ* mediates cell fate separation is unknown, but it appears to involve both cell-autonomous and non-cell-autonomous effects (ten Hove et al., 2010). Our study highlights specific components involved in asymmetric division within the stomatal lineage, *BASL* and *POLAR*, as well as a potential common component, *SCZ*. Future mechanistic studies of these proteins may molecularly define the commonalities and uniqueness of asymmetric divisions in plants.

Methods

Plant material and growth conditions

Arabidopsis thaliana Columbia (Col) accession was used as the wild type. Mutants used in the study were in the Col background. Identification numbers for T-DNA insertional mutants of genes identified in this study are given in Table 2-S3. All insertion lines and the plasma membrane mCherry reporter, pm-RB (CD3-1008) (Nelson et al., 2007), were obtained from ABRC. The following mutants and reporter lines were described previously: *spch* and *mute* (Pillitteri et al., 2007a), *scrm-D* and *scrm-D mute* (Kanaoka et al., 2008), *basl-2* (Dong et al., 2009), *MUTEpro:GFP* (Pillitteri et al., 2008), *EPF2pro:erGFP* (Hara et al., 2009), *CDKB2;1pro:GUS, CycB1;2pro:GUS, CycA2;2pro:GUS*, and *CycA2;3pro:GUS* (gifts from Lieven de Veylder and Steffen Vanneste; Vanneste et al., 2011). Selected reporter lines were crossed with stomatal differentiation mutants. For microarray preparation, seeds of the wild type, *scrm-D, scrm-D mute*, and *spch* were sterilized using 33% bleach solution (bleach and 0.05% Tween 20) for 12 minutes and washed five times with sterile water. Seeds were plated on Murashige and Skoog media and placed at 4° C for five days. Plates were then placed under standard growth conditions (16-hour day/eight-hour night, 21° C), and whole seedlings were collected five days after germination.

Microarray material preparation

Five-day-old seedlings of the wild type, *scrm-D, scrm-D mute*, and *spch* were collected and immediately placed in liquid nitrogen. To avoid any complications due to circadian gene regulation, all genotypes were harvested at the same time (1 PM) within one hour of each other for each replicate. *spch* plants were from a segregating population and were identified by the smooth appearance of the cotyledons compared with wild-type plants. Total RNA was extracted using the Qiagen RNeasy kit (Qiagen) following

manufacturer's instructions. RNA purity and yield were confirmed using the Agilent 2100 Bioanalyzer (Brolingen). Probe preparation and hybridization to the ATH1 GeneChip (Affymetrix) were performed according to the manufacturer's instructions by the University of Washington Center for Array Technologies. Signal intensities from each array were converted to raw expression .CEL data using GeneChip operating software (Affymetrix).

Statistical analysis of microarrays

Our own and publicly available ATH1 microarray data were normalized using Robin software (Lohse et al., 2010). Quality control was performed with standard settings (Limma package using RMA normalized data with Benjamini-Hochberg P-value correction) (Benjamini and Hochberg, 1995). Further statistical analysis was performed using the MultiExperiment Viewer (MeV) software (Saeed et al., 2003). Genes with statistically significant deregulation compared with the wild type were identified using the significance of microarrays module with a Q-value cutoff of <0.005 and a twofold- or threefold-change cutoff. K-means clustering was performed on our own microarrays using log2-transformed normalized expression values as input. Genes were prefiltered using the variance filter implemented in the MeV software to remove genes that showed little variance in expression across the genotypes, to yield 2000 genes with the highest variance among the different arrays. Five clusters were built from maximal 200 iterations using a Pearson correlation distance matrix. Annotation of genes and classification into functional pathways (BINs) was performed with the Mapman software (Thimm et al., 2004). Pageman software (Usadel et al., 2006) was used for overrepresentation analysis.

Verification of microarray results by RT-PCR and qRT-PCR

The RNA isolated from Col-0 wild type, *spch*, *scrm-D mute*, and *scrm-D* seedlings used for the microarray experiment was treated with DNase I (Invitrogen). RNA was converted to cDNA using iScript reverse transcriptase (Bio-Rad) and random hexamers. First-strand cDNA was used directly or diluted 1:10 in double distilled water and used as template for PCR. Real-time PCR was performed using a CFX96 real-time PCR detection system (Bio-Rad) using iTaq SYBR Green Supermix with ROX (Bio-Rad) and standard qPCR conditions in three technical and three biological replicates. Amplification of the *ACT2* gene was used to verify equal loading of cDNA (RT-PCR) or as an internal control in reactions with relative expression calculated using the cycle threshold (Ct) $2^{-\Delta\Delta Ct}$ method (qRT-PCR) (Livak and Schmittgen, 2001). For RT-PCR, all genes were amplified over 31 cycles, except *ACT2* (30 cycles) and *POLAR* (35 cycles), using standard PCR conditions. DNA sequences of primer pairs used for RT-PCR and qRT-PCR can be found in Table 2-S4.

Molecular cloning and generation of transgenic plants

See Table 2-S5 for a list of plasmid constructs generated in this study and Table 2-S6 for a list of primer DNA sequences used for molecular cloning. Briefly, the promoter region (up to ~2.5 kb) upstream of the translation start site was cloned into the pENTR 5' TOPO (Invitrogen) vector according to the manufacturer's instructions. The full-length genomic coding region was amplified and cloned into the pENTR D-TOPO vector (Invitrogen). These vectors were combined in a three-way Gateway cloning reaction using LR Clonase II Plus (Invitrogen) to produce a C-terminal GFP fusion driven by the endogenous promoter. Each genomic plasmid was used in a two-way Gateway reaction to produce an overexpression construct driven by the cauliflower mosaic virus 35S

promoter. The Gateway-based destination vectors were provided by Tsuyoshi Nakagawa (Shimane University) (Nakagawa et al., 2007, 2008). Generation and selection of transgenic plants and their phenotypic analyses were performed as described by Pillitteri et al. (2007a).

Microscopy

Confocal microscopy images were taken on either the Zeiss LSM700 or the Olympus FV1000 microscope. Cell borders were visualized with either propidium iodide (Molecular Probes) or FM4-64 (Invitrogen). The confocal images were false-colored, and their brightness and contrast settings were adjusted using Photoshop CS4 (Adobe). Histochemical staining for GUS activity was performed as previously described (Sessions et al., 1999).

Time-lapse cotyledon imaging

Seeds expressing GFP markers in the desired backgrounds were washed as described above and held at 4° C for two days. All seeds were imbibed and placed under germinating conditions within one hour of the same time (10 AM). Seeds were moved into room temperature and immediately dissected: the seed coat was removed and the hypocotyl severed. Removing the hypocotyl shows no noticeable effect on cotyledon development for the first four days under our conditions. A solution of 0.5% Bacto Agar was prepoured into chamber slides (Lab Tek II chambered #1.5 German cover glass system; Fisher Scientific) at room temperature, and dissected cotyledons were placed beneath it for immobilization and protection from desiccation. The Zeiss LSM700 (inverted, 20x objective, 1x zoom) was used to image the GFP reporter with excitation at 488 nm and a band-pass emission filter of 470 to 500 nm at intervals of 30 min. Z-stacks of 50 to 70 slices at 0.96 μm were captured for each time point and projected for

maximum intensity. For *POLAR-GFP*, stable lines were crossed with pm-RB and the F1 generation imaged as above, with additional excitation at 555 nm and collection with a short-pass 555-nm filter.

Accession numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under the accession numbers in Table 2-S1 and the following: *POLAR* (At4g31805), *CDKB2;2* (At1g20930), *RPA1A* (At2g06510), *BUB3.1* (At3g19590), *ERECTA* (At2g26330), *ERL1* (At5g62230), *ERL2* (At5g07180), *TPD1* (At4g24977), *CDKB2;1* (At1g76540), *CYCB1;2* (At5g06150), *CYCA2;3* (At1g15570), *CYCA2;2* (At5g11300), *IAA7* (At3g23050), Integral membrane family protein (At5g44550), Myosin heavy chain protein (At5g10890), Orp4C (At5g57240), *MUTE* (At3g06120), *SPCH* (At5g53210), *FAMA* (At3g24140), *ICE1*/*SCRM* (At3g26744), *SCRM2* (At1g12860), *SDD1* (At1g04110), *TMM* (At1g80080), *YODA* (At1g63700), *BASL* (At5g60880), *EPF1* (At2g20875), *EPF2* (At1g34245), *KAT1* (At5g46240), *KAT2* (At4g18290), *FIL* (At2g45190), *WUS* (At2g17950), *CLV3* (At2g27250), *AHK2* (At5g35750), *AHK3* (At1g27320), *AHK4* (At2g08130), *ARR5* (At3g48100), *ARR6* (At5g62920), *ARR7* (At1g19050), *ARR15* (At1g74890), *CLE10* (At1g69320), and *SCZ* (At1g46264). The GenBank accession number for *POLAR* cDNA is JN663804. The complete expression data set is available in the National Center for Biotechnology Information Gene Expression Omnibus (Edgar et al., 2002) under accession number GSE29814 (http://www.ncbi.nlm.nih.gov/geo).

Acknowledgements

We thank the ABRC and SIGnAL for providing clones and insertion lines used in this study; S. Vanneste, P. Doerner, J. Murray, D. Inzé, and L. de Vylder for kindly providing transgenic plants expressing various cell-cycle GUS reporter lines; R. Heidstra for the *SCZpro:CFP* line; J. Nemhauser for advice on designing the microarray experiments; D. Bergmann for discussion about stomatal cell lineage profiling; and T. Kuroha and A. Rychel for assisting construction of reporter plasmids. The work was supported by the University of Washington Royalty Research Fund (RRF-4098), the National Science Foundation (MCB- 0855659), and the Precursory Research for Embryonic Science and Technology award from the Japan Science and Technology Agency to K.U.T. L.J.P.'s research was in part supported by startup funds from Western Washington University. K.M.P. is supported by the National Science Foundation Graduate Research Fellowship (DGE-0718124), and R.J.H. is supported by the Deutsche Forschungssgemeinschaft research fellowship. K.U.T. is a Howard Hughes Medical Institute–Gordon and Betty Moore Foundation investigator.

Author Contributions

K.U.T. conceived and directed the project. L.J.P. conducted microarray experiments and analyzed the results with R.J.H. R.J.H. performed bioinformatic analysis. L.J.P. and K.M.P. generated expression constructs and Arabidopsis reporter lines and performed analysis with R.J.H. K.M.P. developed and performed time-lapse imaging. Thus, L.J.P., K.M.P., and R.J.H. made uniquely equal contributions. All authors contributed to writing the article.

References

Abrash, E.B., and Bergmann, D.C. (2009). Asymmetric cell divisions: A view from plant development. *Dev. Cell* 16: 783–796.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289–300.

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

Bergmann, D.C., and Sack, F.D. (2007). Stomatal development. *Annu. Rev. Plant Biol.* 58: 163–181.

Borner, G.H., Sherrier, D.J., Stevens, T.J., Arkin, I.T., and Dupree, P. (2002). Prediction of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A genomic analysis. *Plant Physiol.* 129: 486–499.

Boudolf, V., Lammens, T., Boruc, J., Van Leene, J., Van Den Daele, H., Maes, S., Van Isterdael, G., Russinova, E., Kondorosi, E., Witters, E., De Jaeger, G., Inzé, D., De Veylder, L. (2009). CDKB1;1 forms a functional complex with CYCA2;3 to suppress endocycle onset. *Plant Physiol.* 150: 1482–1493.

Burkhard, P., Stetefeld, J., and Strelkov, S.V. (2001). Coiled coils: A highly versatile protein folding motif. *Trends Cell Biol.* 11: 82–88.

Chung, H.S., and Howe, G.A. (2009). A critical role for the TIFY motif in repression of jasmonate signaling by a stabilized splice variant of the JASMONATE ZIM-domain protein JAZ10 in Arabidopsis. *Plant Cell* 21: 131–145.

Clark, S.E., Running, M.P., and Meyerowitz, E.M. (1993). CLAVATA1, a regulator of meristem and flower development in Arabidopsis. *Development* 119: 397–418.

Day, R.C., McNoe, L., and Macknight, R.C. (2007). Evaluation of global RNA amplification and its use for high-throughput transcript analysis of laser-microdissected endosperm. *Int. J. Plant Genomics* 2007: 61028.

De Almeida Engler, J., De Veylder, L., De Groodt, R., Rombauts, S., Boudolf, V., De Meyer, B., Hemerly, A., Ferreira, P., Beeckman, T., Karimi, M., Hilson, P., Inzé, D., Engler, G. (2009). Systematic analysis of cell-cycle gene expression during *Arabidopsis* development. *Plant J.* 59: 645–660.

Di Laurenzio, L., Wysocka-Diller, J., Malamy, J.E., Pysh, L., Helariutta, Y., Freshour, G., Hahn, M.G., Feldmann, K.A., and Benfey, P.N. (1996). The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. *Cell* 86: 423–433.

Dinneny, J.R., Long, T.A., Wang, J.Y., Jung, J.W., Mace, D., Pointer, S., Barron, C., Brady, S.M., Schiefelbein, J., and Benfey, P.N. (2008). Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. *Science* 320: 942–945.

Dong, J., MacAlister, C.A., and Bergmann, D.C. (2009). BASL controls asymmetric cell division in *Arabidopsis*. *Cell* 137: 1320–1330.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30: 207–210.

Fichelson, P., Moch, C., Ivanovitch, K., Martin, C., Sidor, C.M., Lepesant, J.A., Bellaiche, Y., and Huynh, J.R. (2009). Live-imaging of single stem cells within their niche reveals that a U3snoRNP component segregates asymmetrically and is required for self-renewal in *Drosophila*. *Nat. Cell Biol.* 11: 685–693.

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.

Geisler, M., Nadeau, J., and Sack, F.D. (2000). Oriented asymmetric divisions that generate the stomatal spacing pattern in *Arabidopsis* are disrupted by the *too many mouths* mutation. *Plant Cell* 12: 2075– 2086.

Hara, K., Kajita, R., Torii, K.U., Bergmann, D.C., 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, K.M., Torii, K.U., 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.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A., and Kay, S.A. (2000). Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* 290: 2110–2113.

Hord, C.L., Sun, Y.J., Pillitteri, L.J., Torii, K.U., Wang, H., Zhang, S., and Ma, H. (2008). Regulation of *Arabidopsis* early anther development by the mitogen-activated protein kinases, MPK3 and MPK6, and the ERECTA and related receptor-like kinases. *Mol. Plant* 1: 645–658.

Hunt, L., and Gray, J.E. (2009). The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Curr. Biol.* 19: 864–869.

Inoue, T., Higuchi, M., Hashimoto, Y., Seki, M., Kobayashi, M., Kato, T., Tabata, S., Shinozaki, K., and Kakimoto, T. (2001). Identification of CRE1 as a cytokinin receptor from *Arabidopsis*. *Nature* 409: 1060– 1063.

Iyer-Pascuzzi, A.S., and Benfey, P.N. (2010). Fluorescence-activated cell sorting in plant developmental biology. *Methods Mol. Biol.* 655: 313–319.

Jia, G., Liu, X., Owen, H.A., and Zhao, D. (2008). Signaling of cell fate determination by the TPD1 small protein and EMS1 receptor kinase. *Proc. Natl. Acad. Sci. USA* 105: 2220–2225.

Jun, J., Fiume, E., Roeder, A.H., Meng, L., Sharma, V.K., Osmont, K.S., Baker, C., Ha, C.M., Meyerowitz, E.M., Feldman, L.J., and Fletcher, J.C. (2010). Comprehensive analysis of CLE polypeptide signaling gene expression and overexpression activity in *Arabidopsis*. *Plant Physiol.* 154: 1721–1736.

Kanaoka, M.M., Pillitteri, L.J., Fujii, H., Yoshida, Y., Bogenschutz, N.L., Takabayashi, J., Zhu, J.K., and Torii, K.U. (2008). *SCREAM/ ICE1* and *SCREAM2* specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* 20: 1775–1785.

Kerk, N.M., Ceserani, T., Tausta, S.L., Sussex, I.M., and Nelson, T.M. (2003). Laser capture microdissection of cells from plant tissues. *Plant Physiol.* 132: 27–35.

Kiba, T., Yamada, H., and Mizuno, T. (2002). Characterization of the *ARR15* and *ARR16* response regulators with special reference to the cytokinin signaling pathway mediated by the AHK4 histidine kinase in roots of *Arabidopsis thaliana*. *Plant Cell Physiol.* 43: 1059–1066.

Kondo, Y., Hirakawa, Y., Kieber, J.J., and Fukuda, H. (2011). CLE peptides can negatively regulate protoxylem vessel formation via cytokinin signaling. *Plant Cell Physiol.* 52: 37–48.

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.

Leibfried, A., To, J.P., Busch, W., Stehling, S., Kehle, A., Demar, M., Kieber, J.J., and Lohmann, J.U. (2005). *WUSCHEL* controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* 438: 1172–1175.
Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402–408.
Lohse, M., et al. (2010). Robin: an intuitive wizard application for R-based expression microarray quality assessment and analysis. *Plant Physiol.* 153: 642–651.
Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* 252: 1162–1164.
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.
Menke, F.L., and Scheres, B. (2009). Plant asymmetric cell division, vive la différence! *Cell* 137: 1189–1192.
Nakamura, R.L., McKendree, W.L., Jr., Hirsch, R.E., Sedbrook, J.C., Gaber, R.F., and Sussman, M.R. (1995). Expression of an *Arabidopsis* potassium channel gene in guard cells. *Plant Physiol.* 109: 371–374.
Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., and Kimura, T. (2007). Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* 104: 34–41.
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.
Nakazono, M., Qiu, F., Borsuk, L.A., and Schnable, P.S. (2003). Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: Identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *Plant Cell* 15: 583–596.
Nelson, B.K., Cai, X., and Nebenführ, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. *Plant J.* 51: 1126–1136.
Obayashi, T., Nishida, K., Kasahara, K., and Kinoshita, K. (2011). ATTED-II updates: Condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. *Plant Cell Physiol.* 52: 213–219.
O'Connor, T.R., Dyreson, C., and Wyrick, J.J. (2005). Athena: A resource for rapid visualization and systematic analysis of Arabidopsis promoter sequences. *Bioinformatics* 21: 4411–4413.
Ohashi-Ito, K., and Bergmann, D.C. (2006). Arabidopsis *FAMA* controls the final proliferation/differentiation switch during stomatal development. *Plant Cell* 18: 2493–2505.
Peterson, K.M., Rychel, A.L., and Torii, K.U. (2010). Out of the mouths of plants: The molecular basis of the evolution and diversity of stomatal development. *Plant Cell* 22: 296–306.
Pillitteri, L.J., Bemis, S.M., Shpak, E.D., and Torii, K.U. (2007b). Haploinsufficiency after successive loss of signaling reveals a role for *ERECTA*-family genes in Arabidopsis ovule development. *Development* 134: 3099–3109.

Pillitteri, L.J., Bogenschutz, N.L., and Torii, K.U. (2008). The bHLH protein MUTE controls differentiation of stomata and the hydathode pore in *Arabidopsis*. *Plant Cell Physiol.* 49: 934–943.
Pillitteri, L.J., Sloan, D.B., Bogenschutz, N.L., and Torii, K.U. (2007a). Termination of asymmetric cell division and differentiation of stomata. *Nature* 445: 501–505.
Pillitteri, L.J., and Torii, K.U. (2007). Breaking the silence: Three bHLH proteins direct cell-fate decisions during stomatal development. *Bioessays* 29: 861–870.
Ren, B., Liang, Y., Deng, Y., Chen, Q., Zhang, J., Yang, X., and Zuo, J. (2009). Genome-wide comparative analysis of type-A *Arabidopsis* response regulator genes by overexpression studies reveals their diverse roles and regulatory mechanisms in cytokinin signaling. *Cell Res.* 19: 1178–1190.
Rowe, M.H., and Bergmann, D.C. (2010). Complex signals for simple cells: The expanding ranks of signals and receptors guiding stomatal development. *Curr. Opin. Plant Biol.* 13: 548–555.
Rychel, A.L., Peterson, K.M., and Torii, K.U. (2010). Plant Twitter: Ligands under 140 amino acids enforcing stomatal patterning. *J. Plant Res.* 123: 275–280.
Sachs, T. (1991). *Pattern Formation in Plant Tissues.* (Cambridge, UK: Cambridge University Press).
Saeed, A.I., et al. (2003). TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* 34: 374–378.
Sarkar, A.K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., Scheres, B., Heidstra, R., and Laux, T. (2007). Conserved factors regulate signalling in *Arabidopsis thaliana* shoot and root stem cell organizers. *Nature* 446: 811–814.
Schachtman, D.P., Schroeder, J.I., Lucas, W.J., Anderson, J.A., and Gaber, R.F. (1992). Expression of an inward-rectifying potassium channel by the Arabidopsis *KAT1* cDNA. *Science* 258: 1654–1658.
Scheres, B. (2002). Plant patterning: TRY to inhibit your neighbors. *Curr. Biol.* 12: R804–R806.
Schmülling, T., Werner, T., Riefler, M., Krupková, E., and Bartrina y Manns, I. (2003). Structure and function of cytokinin oxidase/dehydrogenase genes of maize, rice, *Arabidopsis* and other species. *J. Plant Res.* 116: 241–252.
Sena, G., Wang, X., Liu, H.Y., Hofhuis, H., and Birnbaum, K.D. (2009). Organ regeneration does not require a functional stem cell niche in plants. *Nature* 457: 1150–1153.
Sessions, A., Weigel, D., and Yanofsky, M.F. (1999). The *Arabidopsis thaliana MERISTEM LAYER 1* promoter specifies epidermal expression in meristems and young primordia. *Plant J.* 20: 259–263.
Shpak, E.D., Berthiaume, C.T., Hill, E.J., and Torii, K.U. (2004). Synergistic interaction of three ERECTA-family receptor-like kinases controls *Arabidopsis* organ growth and flower development by promoting cell proliferation. *Development* 131: 1491–1501.
Shpak, E.D., Lakeman, M.B., and Torii, K.U. (2003). Dominant-negative receptor uncovers redundancy in the Arabidopsis ERECTA Leucine-rich repeat receptor-like kinase signaling pathway that regulates organ shape. *Plant Cell* 15: 1095–1110.
Shpak, E.D., McAbee, J.M., Pillitteri, L.J., and Torii, K.U. (2005). Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science* 309: 290–293.

Stolz, A., Riefler, M., Lomin, S.N., Achazi, K., Romanov, G.A., and Schmülling, T. (2011). The specificity of cytokinin signalling in *Arabidopsis thaliana* is mediated by differing ligand affinities and expression profiles of the receptors. *Plant J.* 67: 157–168.

ten Hove, C.A., Willemsen, V., de Vries, W.J., van Dijken, A., Scheres, B., and Heidstra, R. (2010). *SCHIZORIZA* encodes a nuclear factor regulating asymmetry of stem cell divisions in the *Arabidopsis* root. *Curr. Biol.* 20: 452–457.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y., and Stitt, M. (2004). MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37: 914–939.

Torii, K.U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R.F., and Komeda, Y. (1996). The Arabidopsis *ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* 8: 735–746.

Usadel, B., Nagel, A., Steinhauser, D., Gibon, Y., Bla¨sing, O.E., Redestig, H., Sreenivasulu, N., Krall, L., Hannah, M.A., Poree, F., Fernie, A.R., and Stitt, M. (2006). PageMan: An interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. *BMC Bioinformatics* 7: 535.

Vanneste, S., et al. (2011). Developmental regulation of CYCA2s contributes to tissue-specific proliferation in *Arabidopsis*. *EMBO J.* 30: 3430–3441.

Van Leene, J., Boruc, J., De Jaeger, G., Russinova, E., and De Veylder, L. (2011). A kaleidoscopic view of the *Arabidopsis* core cell cycle interactome. *Trends Plant Sci.* 16: 141–150.

Werner, T., Köllmer, I., Bartrina, I., Holst, K., and Schmülling, T. (2006). New insights into the biology of cytokinin degradation. *Plant Biol.* 8: 371–381.

Winter, D., Vinegar, B., Nahal, H. Ammar, R., Wilson, G.V., and Povart, M. (2001). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2: e718.

Yadav, R., 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.

Yamada, H., Suzuki, T., Terada, K., Takei, K., Ishikawa, K., Miwa, K., Yamashino, T., and Mizuno, T. (2001). The *Arabidopsis* AHK4 histidine kinase is a cytokinin-binding receptor that transduces cytokinin signals across the membrane. *Plant Cell Physiol.* 42: 1017–1023.

Zhao, D.Z., Wang, G.F., Speal, B., and Ma, H. (2002). The *excess microsporocytes1* gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the Arabidopsis anther. *Genes Dev.* 16: 2021–2031.

Figure 2-1.

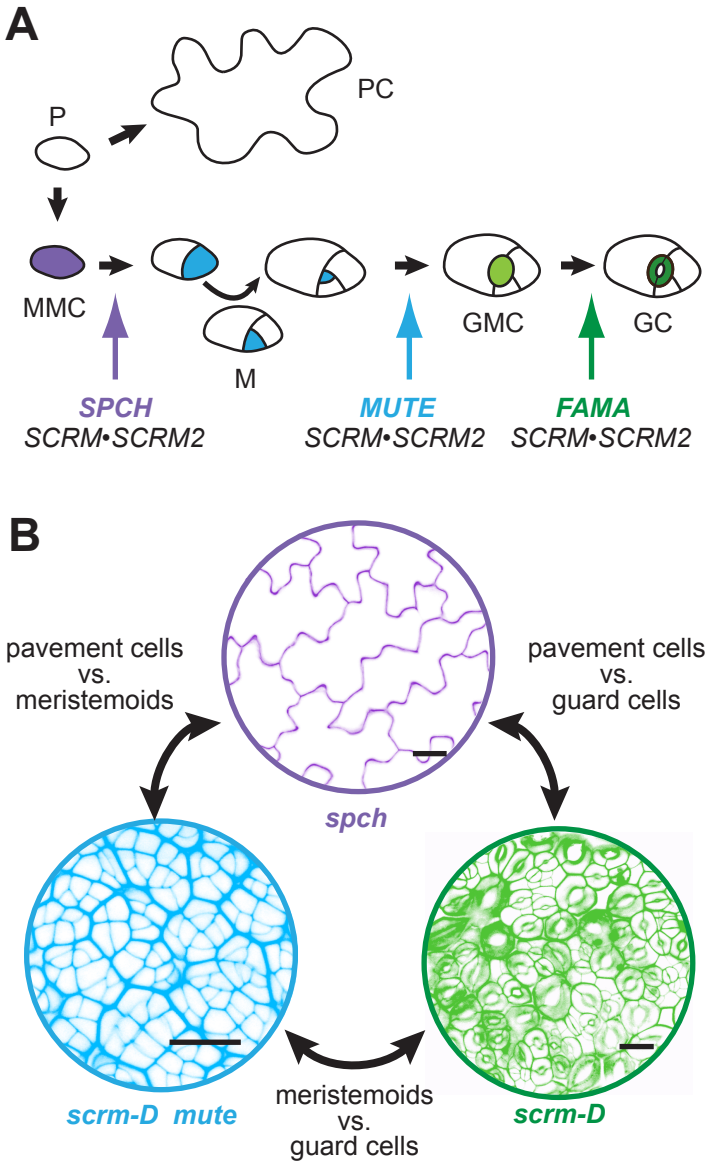


Figure 2-1. Stomatal cell lineage transitions and cell enrichment strategy.
(A) Diagram of the sequential cell-state transitions during stomatal development specified by the combinatorial and sequential actions of the stomatal bHLH genes. A protodermal cell (P) can differentiate into a pavement cell (PC) or undergo a transition to become an MMC (purple). An MMC enters the stomatal lineage through an asymmetric division to create a meristemoid (M; blue). Meristemoids have transient stem-cell-like properties and can undergo several rounds of amplifying divisions before differentiating into a GMC (light green). GMCs divide symmetrically to produce two GCs (dark green), which form a mature stoma. The point of action is indicated for each bHLH gene (colored arrows).
(B) Strategy for comparing molecular signatures associated with each epidermal cell state/type. Images are of abaxial cotyledon epidermis of *spch*, *scrm-D mute*, and *scrm-D* mutant seedlings at 5 d after germination. Colors coincide with those described in (A). Bars = 10 mm.

Figure 2-2.

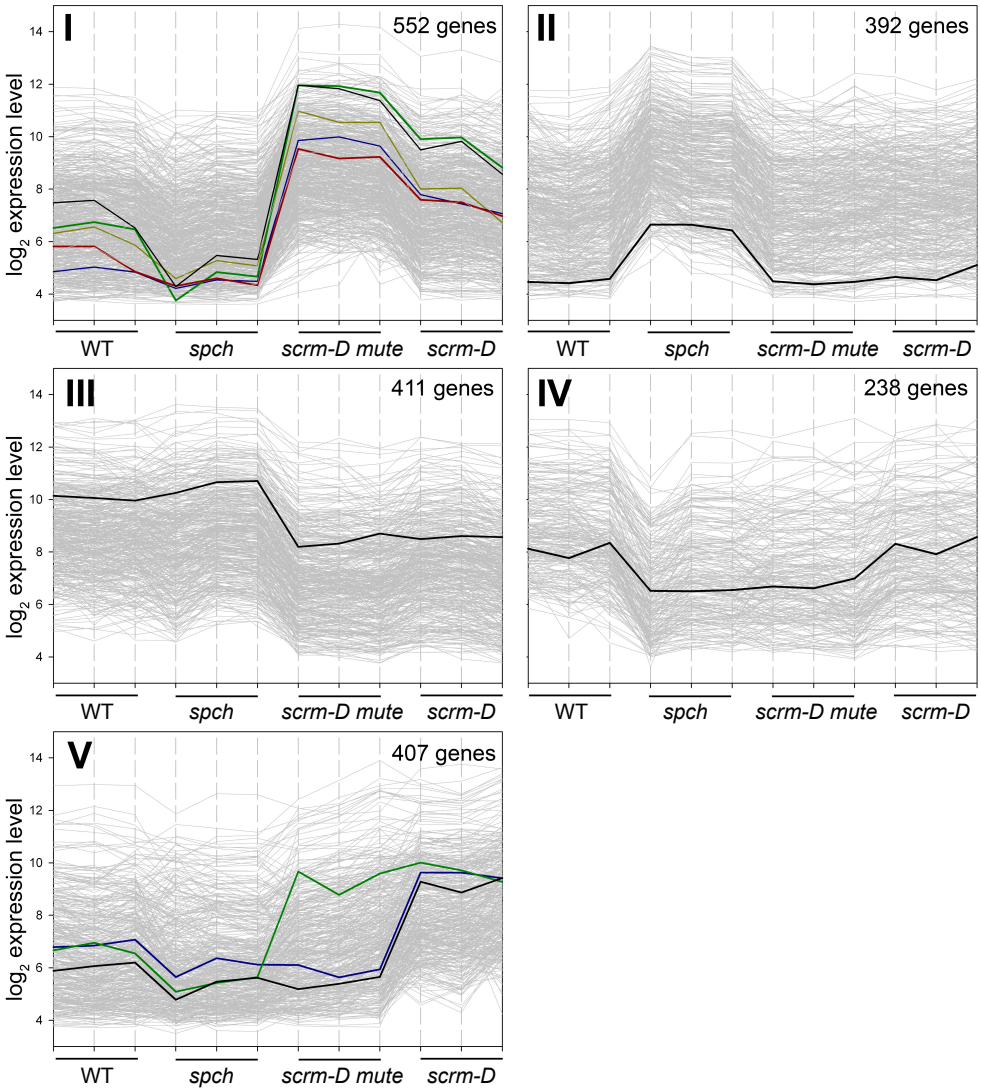


Figure 2-2. K-means clustering of genes showing differential expression in stomatal differentiation mutants.
The log2 expression levels are plotted against the three replicates of each genotype (the wild type, *spch*, *scrm-D mute*, and *scrm-D*). Genes were prefiltered using the variance filter implemented in the MeV software to yield the 2000 genes showing highest variance across the 12 arrays. Cluster I: genes upregulated in *scrm-D mute* and to a lesser extent in *scrm-D* (meristemoid and GMC/GC enriched, respectively). Black, *TMM*; green, *EPF2*; red, *ARR16*; blue, *CLE9*; yellow, *POLAR* (At4g31805). WT, wild type. Cluster II: genes upregulated in *spch* (pavement-cell-only epidermis). Highlighted in black is a representative gene (*Orp4C*, At5g57240) of this cluster. Cluster III: genes with reduced expression in *scrm-D mute* and *scrm-D*. Highlighted in black is *IAA7* (At3g23050), which shows the representative expression pattern in this cluster. Cluster IV: genes expressed most strongly in the wild type and *scrm-D*. Highlighted in black is a representative gene (integral membrane family protein, At5g44550) of this cluster. Cluster V: genes upregulated in *scrm-D* (GMC/GC-enriched). Black, *CHX20*, a cation/H+ exchanger of GCs; blue, *FAMA*; green, *EPF1*.

Figure 2-3.

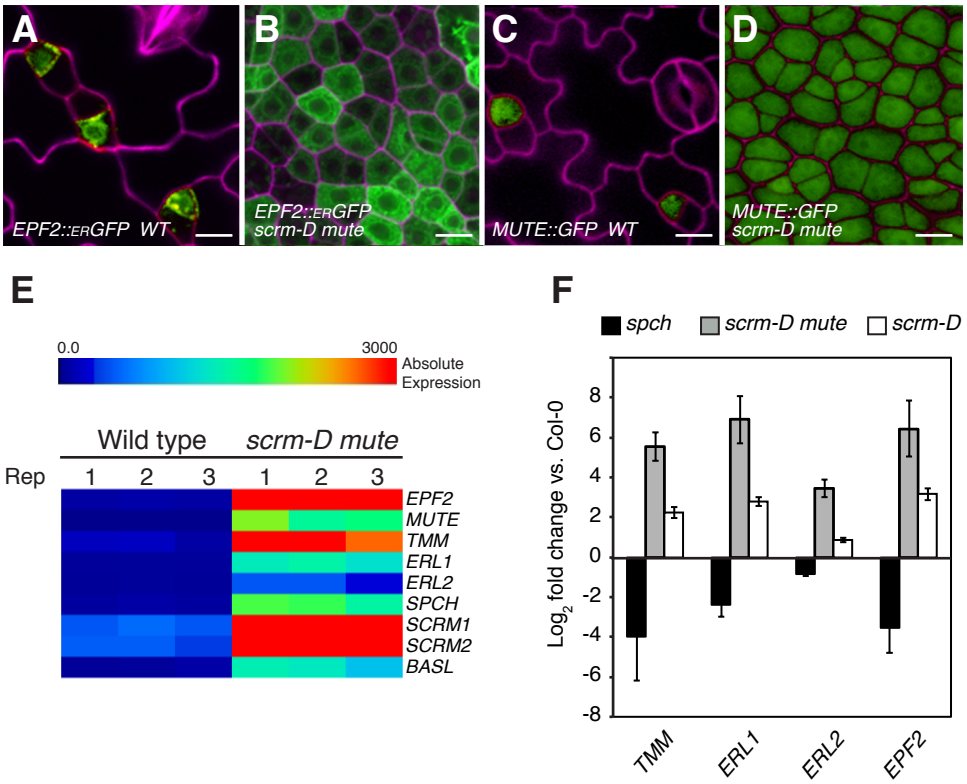


Figure 2-3. Arrested meristemoids constituting the epidermis of *scrm-D mute* express meristemoid markers.
EPF2 and *MUTE* promoter activities (green) are detected only in early stomatal lineage cells and not in mature stomata. Bars = 10 μm. (A) and (B) *EPF2pro:erGFP* (green) in the wild type (WT) (A) and *scrm-D mute* (B). (C) and (D) *MUTEpro:GFP* in the wild type (C) and *scrm-D mute* (D). Cell peripheries were highlighted with either propidium iodide or FM4-64 (magenta). (E) Known regulators of meristemoid development are highly expressed in *scrm-D mute*. Shown is a heat map showing the absolute expression levels in *spch* (pavement-cell-only epidermis) and *scrm-D mute* (meristemoid-enriched). All three replicates are shown to demonstrate consistency. (F) qRT-PCR verification of a subset of stomatal regulators known to be expressed in meristemoids. Error bars represent the SE (n = 3).

Figure 2-4.

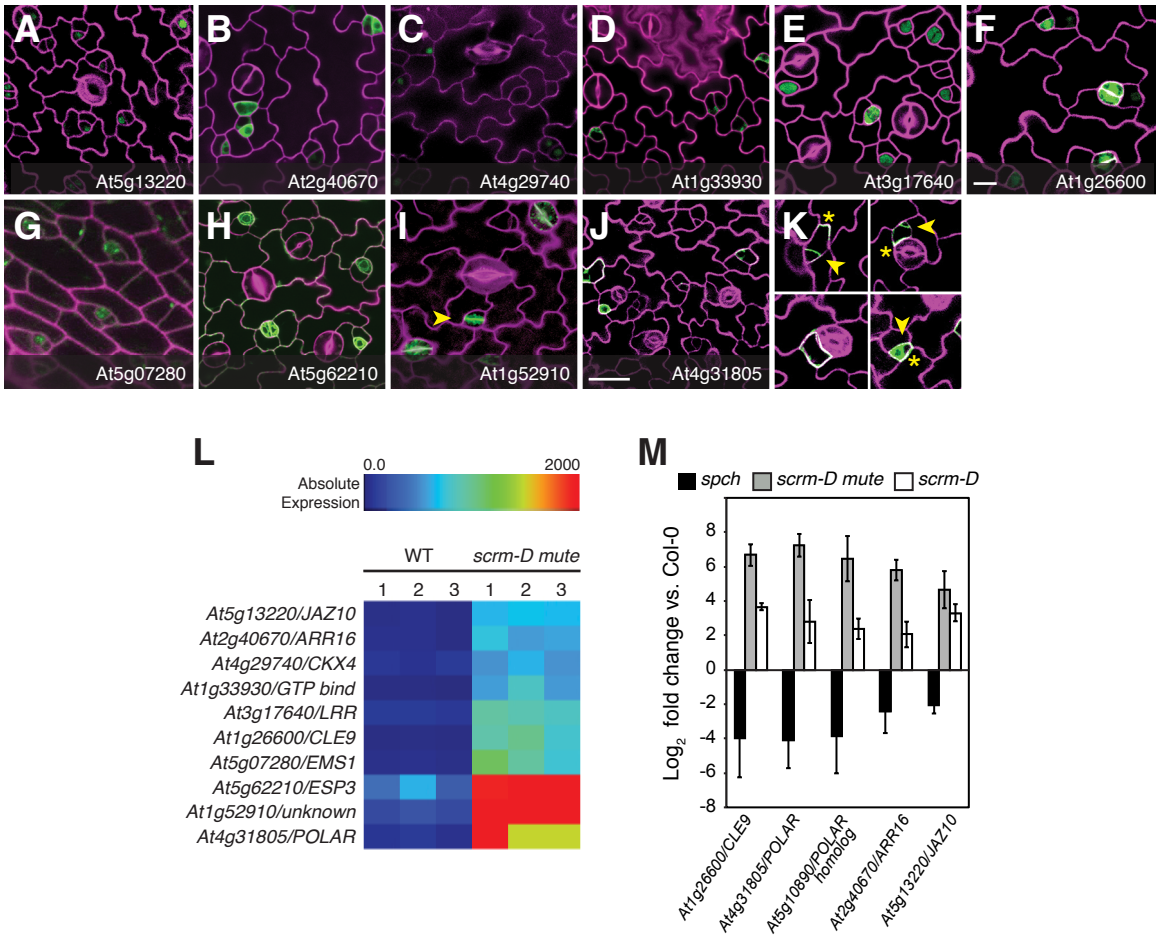


Figure 2-4. Localization in stomatal cell lineages and expression levels.
Genes represent those upregulated in the *scrm-D mute* background. All lines carry a C-terminal GFP protein fusion (green) of the indicated gene driven by its native promoter, except for *CLE9* and *EMS1*, which carry the native promoter driving GFP. (A) to (K) Confocal images of wild-type seedling leaf epidermis. Images are of the abaxial surface, except (E) (adaxial surface). Cell peripheries were highlighted with either propidium iodide or FM4-64 (magenta). Arrow in (I) indicates cell plate. Asterisks in (K) indicate asymmetrical localization of At4g31805-GFP; arrowheads indicate location of division. Bars = 10 mm in (A) to (I) and 25 mm in (J). (L) Heat map representing the degree of upregulation of each indicated gene in *scrm-D mute* compared with the wild type. Three replicates are shown to demonstrate consistency. Scale represents absolute expression of each gene. (M) qRT-PCR verification of selected meristemoid-enriched genes identified in this study. Values are relative to expression of corresponding genes in the wild type. Error bars represent the SE (n = 3).

Figure 2-5.

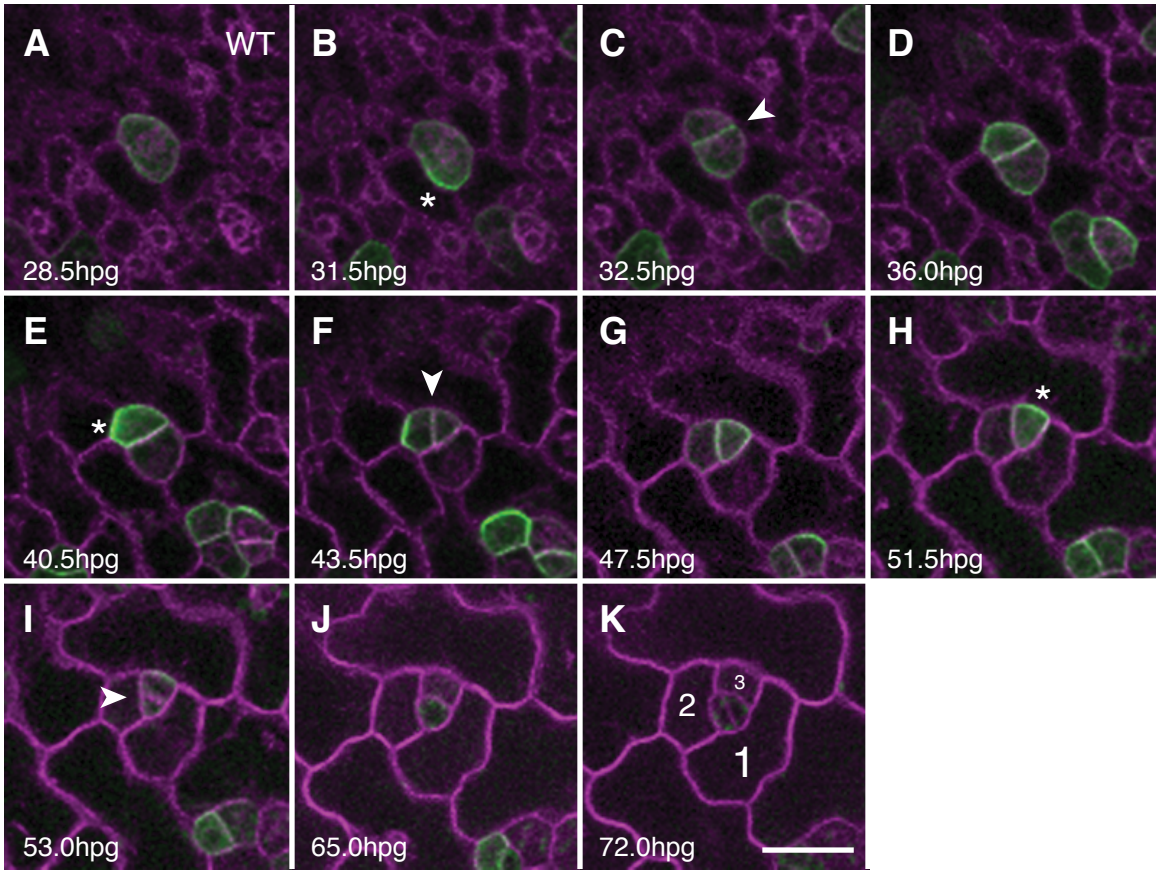


Figure 2-5. Polar localization of POLAR precedes asymmetric division in stomatal cell lineages. Real-time imaging was performed using wild-type cotyledons expressing a C-terminal GFP protein fusion of POLAR driven by the native promoter (*POLARpro:POLAR-GFP*). (A-K) Individual fields from imaging are shown over time as indicated. Cell membranes (magenta) are visualized with the mCherry plasma membrane marker pm-RB. POLAR-GFP displays localized, transient accumulation at the cell periphery distal to the plane of division prior to asymmetric divisions of stomatal lineage cells (MMC and meristemoids), then is redistributed after asymmetric division. Asterisk indicates polar localization. Arrowhead indicates the plane of division. SLGCs resulting from asymmetric divisions are numbered (K). hpg, hours post germination. Scale bar, 20 μm. See also Supplemental Movie 2-1.

Figure 2-6.

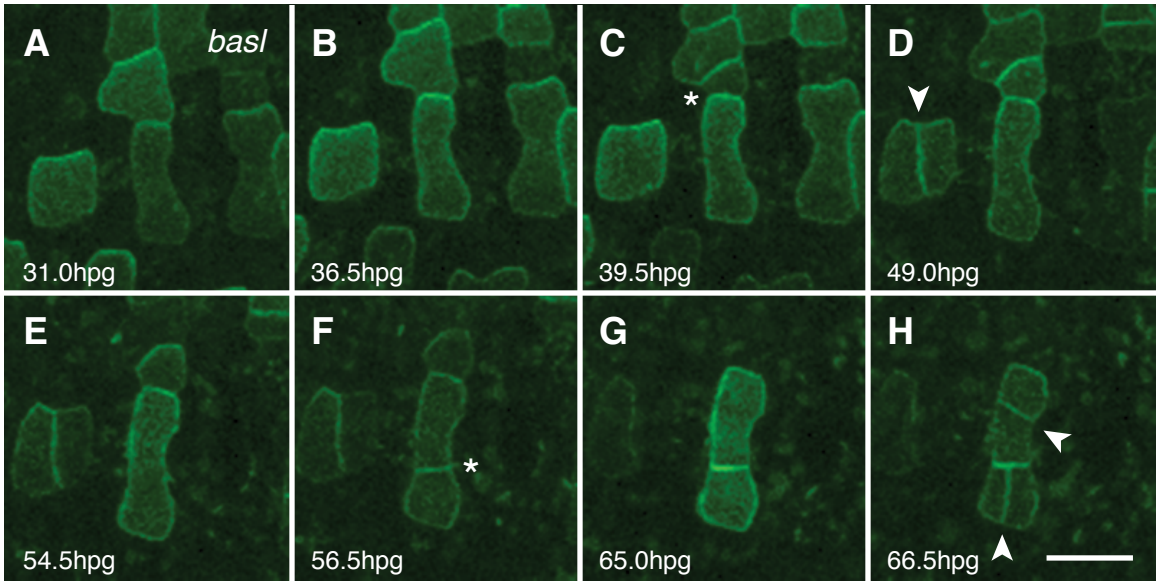


Figure 2-6. Localization of POLAR in the *basl* mutant defective in asymmetric division.
Real-time imaging was performed using *basl* cotyledons expressing *POLARpro:POLAR-GFP*. (A-H) Individual fields from imaging are shown over time as indicated. Loss-of-function *basl* mutants have diffuse peripheral POLAR-GFP that is not strongly polarized. Arrowheads indicate symmetrical division planes (D, H) relative to the parent cell. Asterisks indicate morphologically asymmetrical divisions. hpg, hours post germination. Scale bar, 20 μm. See also Supplementary Movie 2-2.

Figure 2-7.

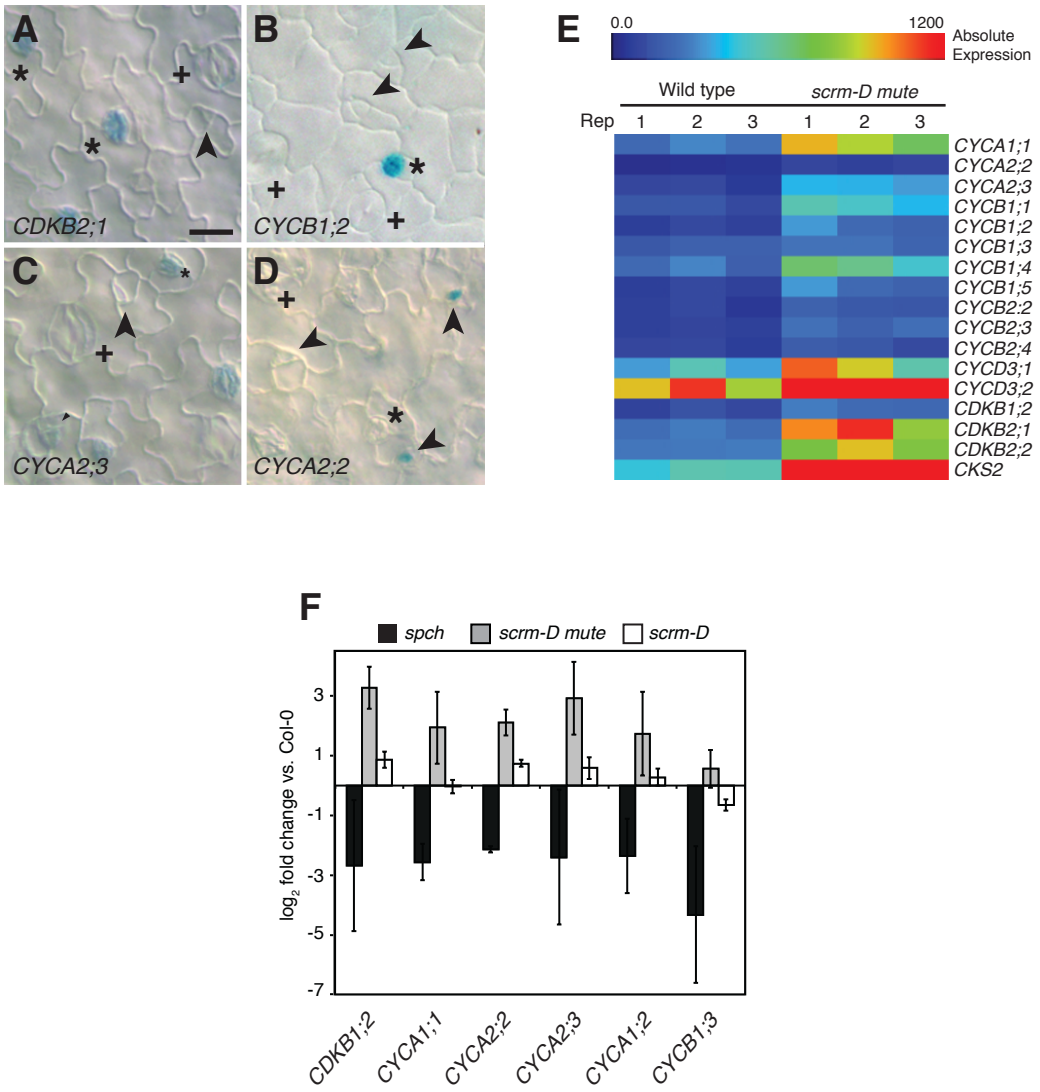


Figure 2-7. Analysis of cell-cycle regulatory gene expression in the stomatal lineage.
All lines carry the β-glucuronidase (GUS) coding sequence driven by the native promoter of each indicated gene. Shown are DIC microscope images of abaxial leaf epidermis from wild-type seedlings expressing the following reporter lines: (A) *CDKB1;2*, At1g76540. (B) *CYCB1;2*, At5g06150. (C) *CYCA2;3*, At1g15570. (D) *CYCA2;2*, At5g11300. Scale bar, 10 μm. Arrowheads indicate meristemoids; asterisks indicate GMCs; (+) indicate stomata. (E) Heat map representing the degree of upregulation of each indicated gene in the *scrm-D mute* background compared to wild type. All three replicates of both genotypes are shown to demonstrate consistency. Scale represents absolute expression of each gene. (F) qRT-PCR of selected core cell-cycle regulators in *spch*, *scrm-D mute*, and *scrm-D*. Gene names are indicated. Values are relative to the wild-type background. Error bars represent the SE (n = 3).

Figure 2-8.

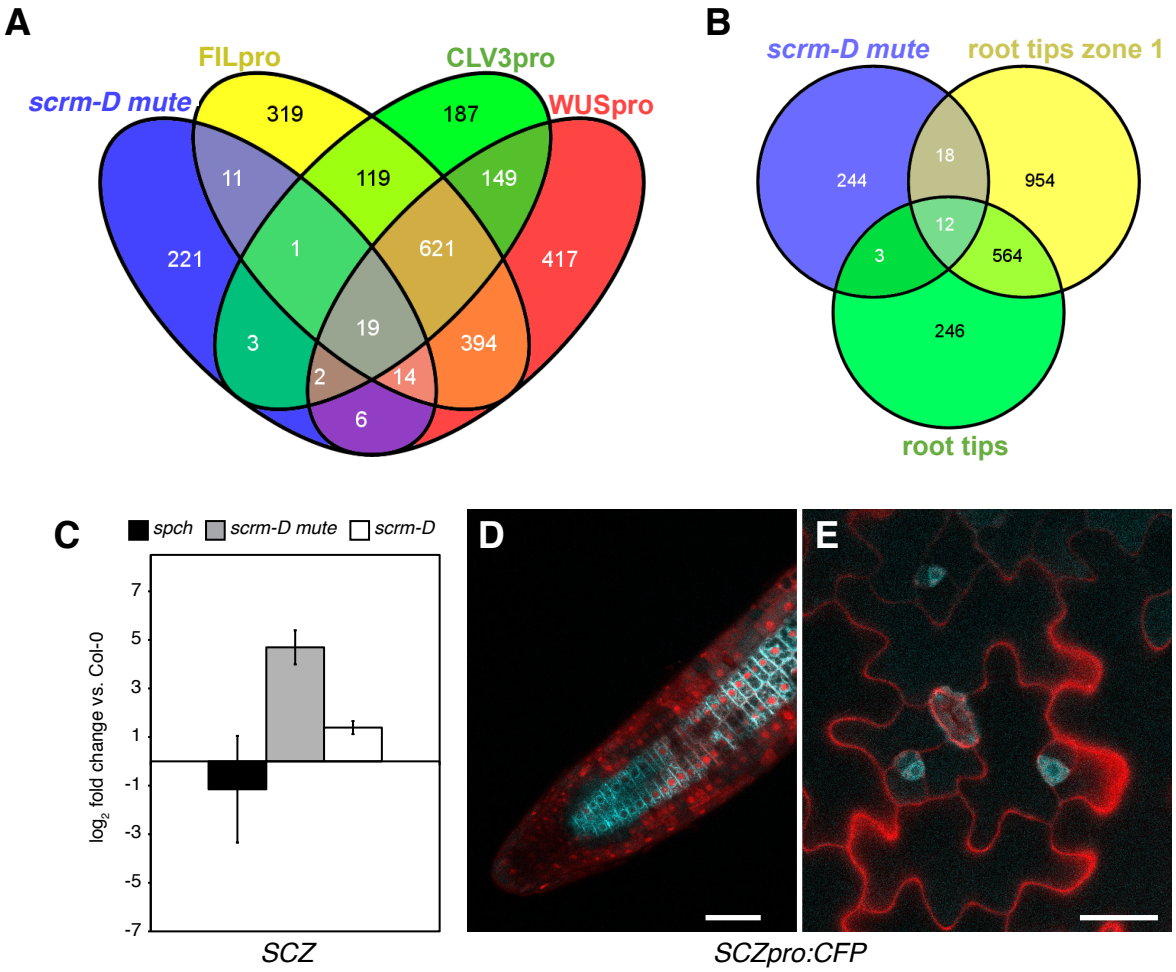


Figure 2-8. Integrative analysis of transient and permanent stem-cell populations.
(A, B) Venn diagrams showing the distribution of unique and shared genes upregulated among *scrm-D mute*, *CLV3-*, *FIL-*, or *WUS*-expressing cells (A) as well as among *scrm-D mute*, root tips zone1 and root tips (B). (C) qPCR analysis showing that *SCZ*, a known regulator of asymmetric cell division within root meristems, is highly upregulated in *scrm-D mute* and downregulated in *spch*. (D, E) Expression of *SCZ* promoter driving CFP. *SCZ* promoter is active in stelar tissue within the root meristem (D). *SCZ* promoter is also highly active in meristemoids (E). Scale bars, 10 μm (D), 20 μm (E).

Figure 2-S1.

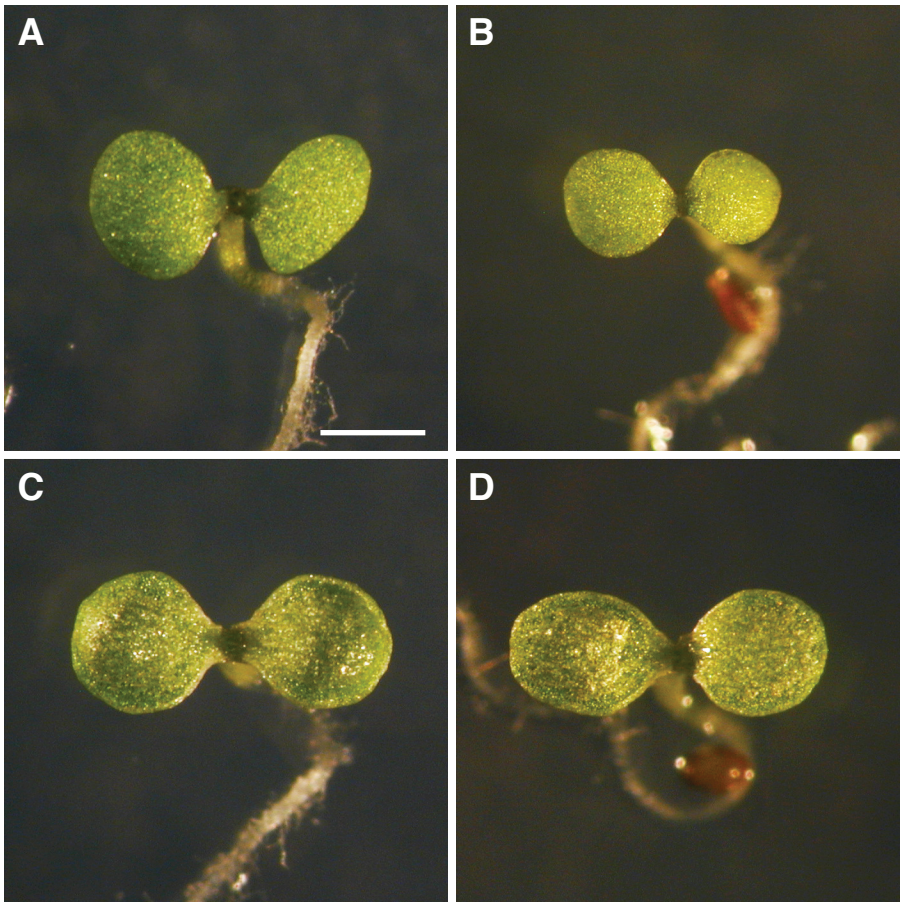


Figure 2-S1. Images of representative 5-dag seedlings used in this study.
(A) WT; (B) *spch*; (C) *scrm-D*; (D) *scrm-D mute*. Images were taken under the same magnificaiton. Scale bar, 1 mm.

Figure 2-S2.

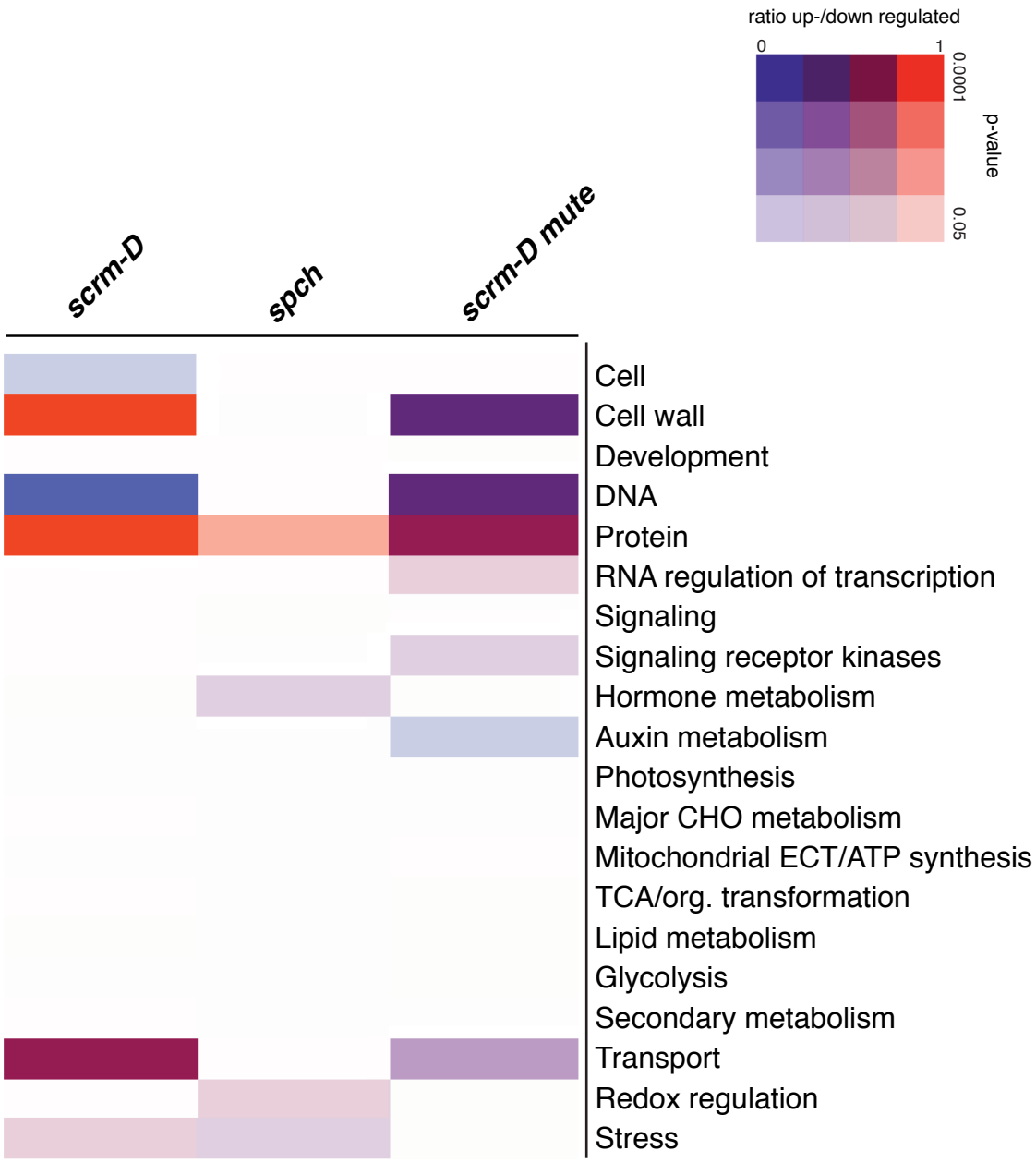


Figure 2-S2. Summary of overrepresented functional gene categories associated with cell-enriched genotypes.
Columns represent *scrm-D*, *spch*, or *scrm-D mute* as specified. Gene functional categories overrepresented in each genetic background are indicated. Gene upregulation was determined through comparison with wild-type seedling expression. Overrepresentation analysis was performed using Pageman software (Usadel et al., 2006).

Figure 2-S3.

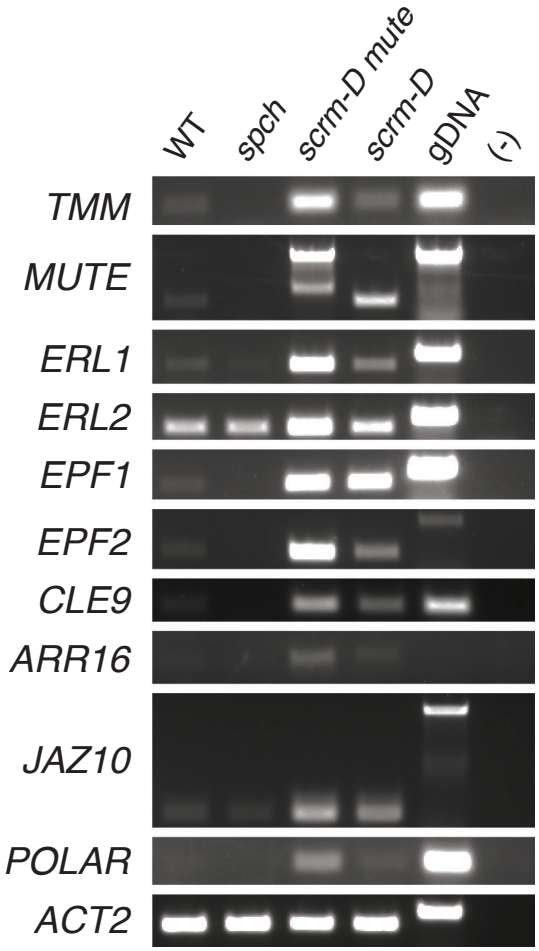


Figure 2-S3. RT-PCR data for selected genes.
Expression of selected stomatal lineage markers and candidate genes that were enriched in the *scrm-D mute* microarray was verified by RT-PCR in Col-0, *spch*, *scrm-D mute*, and *scrm-D* genotypes as indicated. *ACTIN2 (ACT2)* was used as loading control. All genes tested showed the same expression pattern as determined by microarrays. gDNA, genomic DNA. (-), H_2O no template control.

Figure 2-S4.

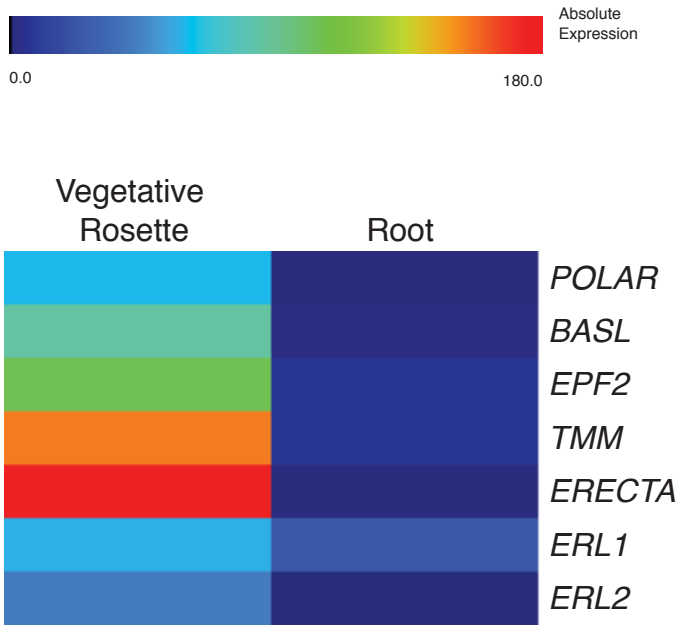


Figure 2-S4. Heatmap showing absolute expression of selected stomatal regulatory genes in rosettes and roots.
Data was retrieved from publicly available datasets (eFP Browser; Winter et al., 2007).

Figure 2-S5.

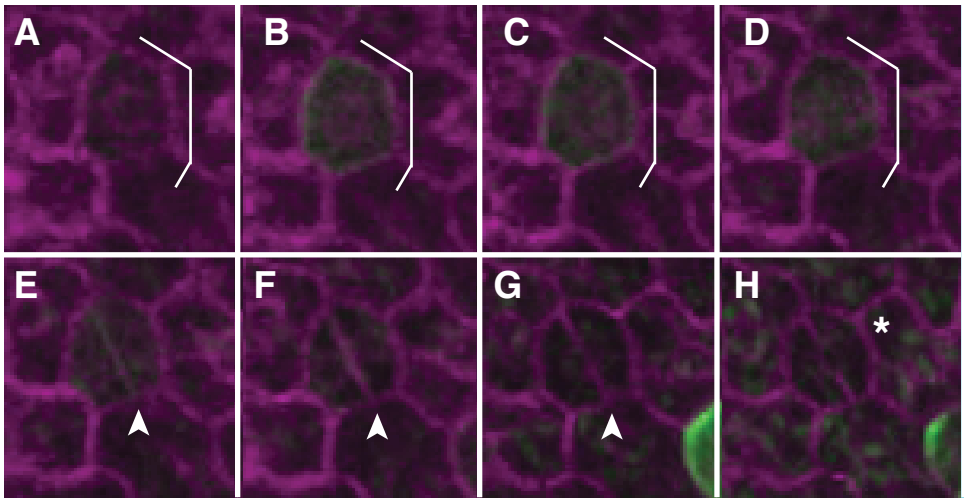


Figure 2-S5. Expression of *POLAR* in a GMC differentiating without asymmetric divisions.
(A-H) Real-time imaging was performed on cotyledons expressing *POLARpro::POLAR-GFP*. Individual fields are shown. Cell membranes are visualized with the pm-RB mCherry marker. (A-D) Brackets indicate a GMC with very weak, diffuse POLAR-GFP localization. (E-H) GMC divides symmetrically to produce a stoma. Arrowhead indicates division plane. Asterisk indicates a stoma with clear central pore.

Figure 2-S6.

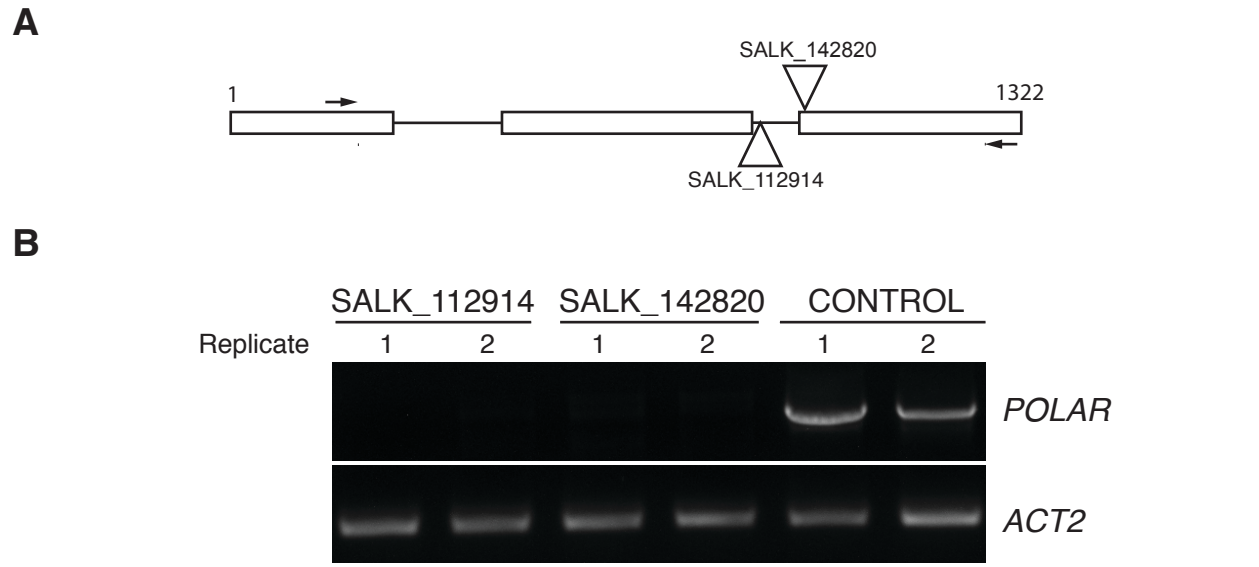


Figure 2-S6. RT-PCR of *POLAR* T-DNA insertion lines and sequence comparison of POLAR coiled-coil domain.

A) Genomic structure of *POLAR*. Exons are shown as boxes, introns as lines and T-DNA insertions as triangles. Relative location of primers used for RT-PCR are shown as arrows. B) Reverse-transcriptase PCR of two biological replicates of SALK_112914, SALK_142820, and Col DNA with Q8 GFP plasma membrane marker. Forward and reverse primers used for analysis: 5′-AATCAGCTGGTAATGCAGGAC-3′ and 5′-CATCTAGAAAAACTGGTGCGG-3′. *ACTIN (ACT2)* was used as loading control. T-DNA insertion lines show no detectable transcript. B) POLAR has a predicted C-terminal coiled-coil domain (http://smart.embl-heidelberg.de/). Sequence alignment of POLAR deduced amino acid sequence with the myosin heavy chain-related protein, At5g10890. Alignment was performed using Clustal W. * indicates an identical amino acid; : indicates a similar amino acid substitution.

Figure 2-S7.

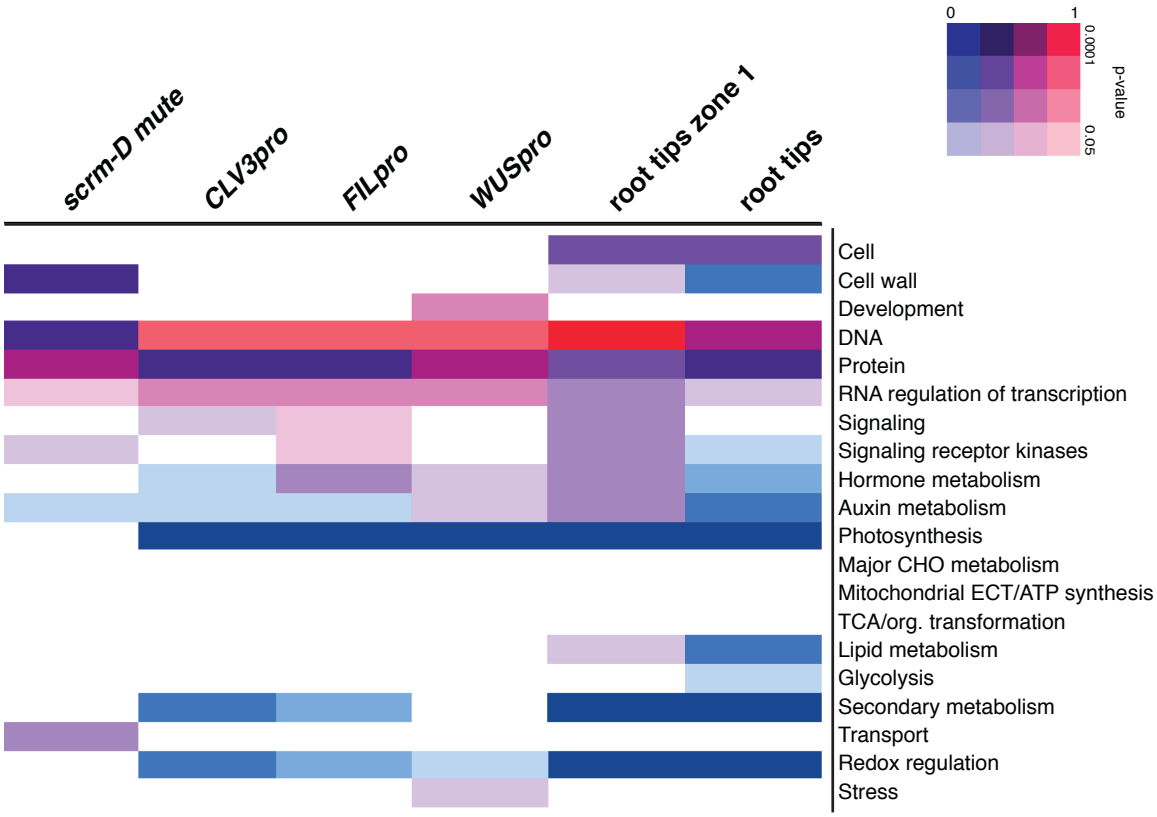


Figure 2-S7. Overrepresentation analysis of transient and permanent stem-cell populations. Columns represent cells expressing the *CLV3*, *FIL*, or *WUS* promoters, respectively. Root tip zone 1 represents all cells up to 150 μm from the root tip of 5-day-old seedlings (Dinneny et al., 2008) and root tip represents cells collected from 130 μm from the root tip, which includes the QC and surrounding stem cells (Sena et al., 2009). Up- and downregulated genes ($p \leq 0.005$) were determined through comparison with wild-type seedling expression. All comparisons are made to wild-type seedlings.

Figure 2-S8.

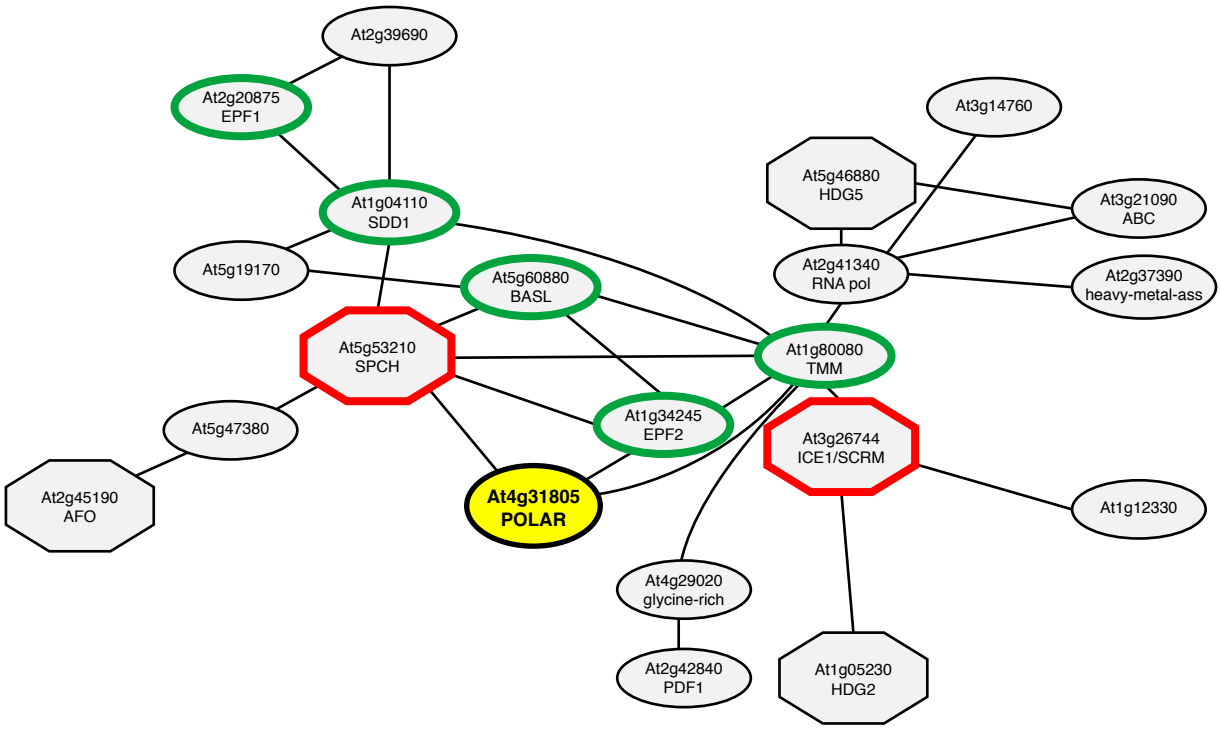


Figure 2-S8. *POLAR* co-expression network.
Identification of genes co-expressed with *POLAR* was performed using ATTED-II (Obayashi et al., 2011). Red octagonal blocks indicate bHLH proteins that are known regulators of stomatal differentiation (see Figure 2-1). Green outlines indicate non-bHLH genes involved in stomatal patterning and differentiation.

Table 2-S1.

ATH1 Array identifier	Gene identifier	Gene description
252280_at	AT3G49260	IQD21 (IQ-DOMAIN 21, IQ-domain 21); calmodulin binding
247056_at	AT5G66750	CHR01/DDM1 (DECREASED DNA METHYLATION 1); helicase
266802_at	AT2G22900	galactosyl transferase GMA12/MNN10 family protein
249832_at	AT5G23400	disease resistance family protein / LRR family protein
266508_at	AT2G47920	kinase interacting family protein
247049_at	AT5G66440	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G34560.1); similar to hypothetical protein *Thellungiella halophila* (GB:ABB45849.1)
260683_at	AT1G17560	HLL (HUELLENLOS); structural constituent of ribosome
261780_at	AT1G76310	CYCB2;4 (CYCLIN B2;4); cyclin-dependent protein kinase regulator
253340_s_at	AT4G33270; AT4G33260	AT4G33270, CDC20.1; signal transducer; AT4G33260, CDC20.2; signal transducer
260771_at	AT1G49160	WNK7 (Arabidopsis WNK kinase 7); kinase
261907_at	AT1G65060	4CL3 (4-coumarate:CoA ligase 3); 4-coumarate-CoA ligase
246683_at	AT5G33300	chromosome-associated kinesin-related
261089_at	AT1G07570	APK1A (Arabidopsis protein kinase 1A); kinase
267280_at	AT2G19450	TAG1 (TRIACYLGLYCEROL BIOSYNTHESIS DEFECT 1); diacylglycerol O-acyltransferase
266401_s_at	AT2G38620; AT3G54180	AT2G38620, CDKB1;2 (cyclin-dependent kinase B1;2); kinase; AT3G54180, CDC2B (CDC2-LIKE GENE); kinase/ protein binding
266921_at	AT2G45970	CYP86A8 (LACERATA); fatty acid (omega-1)-hydroxylase/ oxygen binding
250433_at	AT5G10400	histone H3
255790_at	AT2G33560	spindle checkpoint protein-related
266009_at	AT2G37420	kinesin motor protein-related
254608_at	AT4G18910	NIP1;2/NLM2 (NOD26-like intrinsic protein 1;2); water channel
249427_at	AT5G39850	40S ribosomal protein S9 (RPS9C)
260301_at	AT1G80290	glycosyltransferase family protein 47
258702_at	AT3G09730	similar to unnamed protein product *Vitis vinifera* (GB:CAO18124.1)
252319_at	AT3G48710	GTP binding / RNA binding
256766_at	AT3G22231	PCC1 (PATHOGEN AND CIRCADIAN CONTROLLED 1)
266688_at	AT2G19660	DC1 domain-containing protein
260262_at	AT1G68470	exostosin family protein
264767_at	AT1G61380	S-locus protein kinase, putative
248729_at	AT5G48010	pentacyclic triterpene synthase, putative
265171_at	AT1G23790	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G70340.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO40888.1); contains InterPro domain Protein of unknown function DUF936, plant (InterPro:IPR010341)

ATH1 Array identifier	Gene identifier	Gene description
247134_at	AT5G66230	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G51230.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO48217.1)
245965_at	AT5G19730	pectinesterase family protein
254190_at	AT4G23885	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G24165.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO69543.1)
260727_at	AT1G48100	glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein
262802_at	AT1G20930	CDKB2;2 (CYCLIN-DEPENDENT KINASE B2;2); kinase
251477_at	AT3G59680	similar to hypothetical protein *Vitis vinifera* (GB:CAN69676.1)
261323_at	AT1G44760	universal stress protein (USP) family protein
257714_at	AT3G27360	histone H3
251535_at	AT3G58540	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G06190.1)
263329_at	AT2G15260	zinc finger (C3HC4-type RING finger) family protein
260002_at	AT1G67940	ATNAP3 (Arabidopsis thaliana non-intrinsic ABC protein 3)
267271_at	AT2G02540	ATHB21/ZFHD4 (ZINC FINGER HOMEODOMAIN 4); DNA binding / transcription factor
258707_at	AT3G09480	histone H2B, putative
251140_at	AT5G01090	legume lectin family protein
253665_at	AT4G30230	unknown protein
261521_at	AT1G71830	SERK1 (SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1); kinase
249851_at	AT5G23260	TT16 (TRANSPARENT TESTA16); transcription factor
258952_at	AT3G01410	RNase H domain-containing protein
256389_at	AT3G06220	DNA binding / transcription factor
261146_at	AT1G19620	unknown protein
253978_at	AT4G26660	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G55520.1); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G55520.2); similar to unnamed protein product *Vitis vinifera* (GB:CAO38858.1); contains InterPro domain Kinesin-related (InterPro:IPR010544)
256352_at	AT1G54970	ATPRP1 (PROLINE-RICH PROTEIN 1); structural constituent of cell wall
248526_at	AT5G50740	metal ion binding
245689_at	AT5G04120	phosphoglycerate/bisphosphoglycerate mutase family protein
255410_at	AT4G03100	rac GTPase activating protein, putative
264652_at	AT1G08920	sugar transporter, putative
256622_at	AT3G28920	ATHB34 (ARABIDOPSIS THALIANA HOMEOBOX PROTEIN 34); DNA binding / transcription factor
249361_at	AT5G40540	protein kinase, putative
253453_at	AT4G31860	protein phosphatase 2C, putative / PP2C, putative
253051_at	AT4G37490	CYC1 (CYCLIN 1); cyclin-dependent protein kinase regulator
253696_at	AT4G29740	CKX4 (CYTOKININ OXIDASE 4); cytokinin dehydrogenase
247471_at	AT5G62230	ERL1 (ERECTA-LIKE 1); kinase

ATH1 Array identifier	Gene identifier	Gene description
262201_at	AT2G01120	ATORC4/ORC4 (ORIGIN RECOGNITION COMPLEX SUBUNIT 4); protein binding
257020_at	AT3G19590	WD-40 repeat family protein / mitotic checkpoint protein, putative
263979_at	AT2G42840	PDF1 (PROTODERMAL FACTOR 1)
262558_at	AT1G31335	unknown protein
253840_at	AT4G27780	ACBP2 (ACYL-COA BINDING PROTEIN ACBP 2)
263567_at	AT2G15440	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G67210.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO71903.1); contains InterPro domain Protein of unknown function DUF579, plant (InterPro:IPR006514)
255283_at	AT4G04620	ATG8B (AUTOPHAGY 8B); microtubule binding
249599_at	AT5G37990	S-adenosylmethionine-dependent methyltransferase
267436_at	AT2G19190	FRK1 (FLG22-INDUCED RECEPTOR-LIKE KINASE 1); kinase
259789_at	AT1G29395	COR414-TM1 (cold regulated 414 thylakoid membrane 1)
250026_at	AT5G18090	transcriptional factor B3 family protein
253722_at	AT4G29190	zinc finger (CCCH-type) family protein
259987_at	AT1G75030	ATLP-3 (Arabidopsis thaumatin-like protein 3)
264061_at	AT2G27970	CKS2 (CDK-SUBUNIT 2); cyclin-dependent protein kinase
259694_at	AT1G63180	UGE3 (UDP-D-GLUCOSE/UDP-D-GALACTOSE 4-EPIMERASE 3); UDP-glucose 4-epimerase/ protein dimerization
267550_at	AT2G32800	AP4.3A; ATP binding / protein kinase
253065_at	AT4G37740	AtGRF2 (GROWTH REGULATING FACTOR 2)
257793_at	AT3G26960	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G41050.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO49052.1)
261859_at	AT1G50490	UBC20 (UBIQUITIN-CONJUGATING ENZYME 20); ubiquitin-protein ligase
257929_at	AT3G16980	DNA-directed RNA polymerase II, putative
246889_at	AT5G25470	DNA binding
265924_at	AT2G18620	geranylgeranyl pyrophosphate synthase, putative / GGPP synthetase, putative / farnesyltranstransferase, putative
259979_at	AT1G76600	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G21010.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN67638.1)
252077_at	AT3G51720	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G38370.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO48842.1); contains InterPro domain Protein of unknown function DUF827, plant (InterPro:IPR008545)
265085_at	AT1G03780	targeting protein-related
251705_at	AT3G56400	WRKY70 (WRKY DNA-binding protein 70); transcription factor
265277_at	AT2G28410	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G60650.1)
256210_at	AT1G50950	thioredoxin-related
261247_at	AT1G20070	unknown protein

ATH1 Array identifier	Gene identifier	Gene description
263944_at	AT2G36040	transposable element gene
262507_at	AT1G11330	S-locus lectin protein kinase family protein
264240_at	AT1G54820	protein kinase family protein
254689_at	AT4G13710	pectate lyase family protein
256780_at	AT3G13640	ATRLI1 (Arabidopsis thaliana RNase L inhibitor protein 1)
248889_at	AT5G46230	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G09310.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO14438.1); contains InterPro domain Protein of unknown function DUF538 (InterPro:IPR007493)
258796_at	AT3G04630	WDL1 (WVD2-LIKE 1)
247069_at	AT5G66920	SKS17 (SKU5 Similar 17); copper ion binding / oxidoreductase
257809_at	AT3G27060	TSO2 (TSO MEANING 'UGLY' IN CHINESE); ribonucleoside-diphosphate reductase
248395_at	AT5G52120	ATPP2-A14 (Phloem protein 2-A14); carbohydrate binding
265979_at	AT2G11150	transposable element gene
258471_at	AT3G06030	ANP3 (Arabidopsis NPK1-related protein kinase 3); kinase
253301_at	AT4G33720	pathogenesis-related protein, putative
257005_at	AT3G14190	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G12360.1)
245038_at	AT2G26560	PLP2 (PHOSPHOLIPASE A 2A); nutrient reservoir
250821_at	AT5G05190	Identical to Uncharacterized protein At5g05190 (Y-1) *Arabidopsis thaliana* (GB:Q9FHK4); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G56410.1); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G56410.2); similar to unnamed protein product *Vitis vinifera* (GB:CAO41531.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN78033.1)
266655_at	AT2G25880	ATAUR2 (ATAURORA2); histone serine kinase(H3-S10 specific) / kinase
259388_at	AT1G13420	sulfotransferase family protein
264019_at	AT2G21130	peptidyl-prolyl cis-trans isomerase / cyclophilin (CYP2) / rotamase
250994_at	AT5G02490	heat shock cognate 70 kDa protein 2 (HSC70-2) (HSP70-2)
249727_at	AT5G35490	unknown protein
262543_at	AT1G34245	similar to EPF1 (EPIDERMAL PATTERNING FACTOR 1) *Arabidopsis thaliana* (TAIR:AT2G20875.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO64636.1); contains domain PROKAR_LIPOPROTEIN (PS51257)

ATH1 Array identifier	Gene identifier	Gene description
253737_at	AT4G28703	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G04300.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO44490.1); similar to Os04g0509400 *Oryza sativa* (*japonica* cultivar-group) (GB:NP_001053279.1); similar to OSJNBb0065L13.6 *Oryza sativa* (*japonica* cultivar-group) (GB:CAE03363.1); contains InterPro domain RmlC-like jelly roll fold (InterPro:IPR014710); contains InterPro domain Protein of unknown function DUF861, cupin-3 (InterPro:IPR008579); contains InterPro domain Cupin, RmlC-type (InterPro:IPR011051)
248656_at	AT5G48460	fimbrin-like protein, putative
252149_at	AT3G51290	proline-rich family protein
264990_at	AT1G27210	binding
250642_at	AT5G07180	ERL2 (ERECTA-LIKE 2); kinase
261159_s_at	AT1G34460; AT5G06150	AT1G34460, CYCB1;4 (CYCLIN 3); cyclin-dependent protein kinase regulator; AT5G06150, CYC1BAT (CYCLIN B 1;2); cyclin-dependent protein kinase regulator
251372_at	AT3G60520	zinc ion binding
250376_at	AT5G11550	binding
246952_at	AT5G04820	ATOFP13/OFP13 (Arabidopsis thaliana ovate family protein 13)
258098_at	AT3G23670	KINESIN-12B/PAKRP1L; microtubule motor/ plus-end-directed microtubule motor
246929_at	AT5G25210	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G32030.1)
253514_at	AT4G31805	WRKY family transcription factor
256125_at	AT1G18250	ATLP-1 (Arabidopsis thaumatin-like protein 1)
251663_at	AT3G57070	glutaredoxin family protein
246905_at	AT5G25570	similar to unnamed protein product *Vitis vinifera* (GB:CAO44135.1)
255039_at	AT4G09570	CPK4 (calcium-dependent protein kinase 4); calmodulin-dependent protein kinase/ kinase
256864_at	AT3G23890	TOPII (TOPOISOMERASE II); ATP binding / DNA binding / DNA topoisomerase (ATP-hydrolyzing)
256100_at	AT1G13750	calcineurin-like phosphoesterase family protein
256617_at	AT3G22240	unknown protein
247667_at	AT5G60150	similar to unnamed protein product *Vitis vinifera* (GB:CAO21780.1)
253148_at	AT4G35620	CYCB2;2 (CYCLIN B2;2); cyclin-dependent protein kinase regulator
259930_at	AT1G34355	forkhead-associated domain-containing protein / FHA domain-containing protein
248292_at	AT5G53030	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G27810.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN71155.1)
256522_at	AT1G66160	U-box domain-containing protein

ATH1 Array identifier	Gene identifier	Gene description
257858_at	AT3G12920	protein binding / zinc ion binding
262369_at	AT1G73010	phosphoric monoester hydrolase
255360_at	AT4G03960	tyrosine specific protein phosphatase family protein
246415_at	AT5G17160	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G03130.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO62214.1)
263202_at	AT1G05630	5PTASE13/AT5PTASE13; inositol or phosphatidylinositol phosphatase
245229_at	AT4G25620	hydroxyproline-rich glycoprotein family protein
250860_at	AT5G04770	ATCAT6/CAT6 (CATIONIC AMINO ACID TRANSPORTER 6); amino acid transmembrane transporter/ basic amino acid transmembrane transporter/ cationic amino acid transmembrane transporter
253110_at	AT4G35930	Identical to F-box protein At4g35930 *Arabidopsis thaliana* (GB:Q5XF11;GB:O65630); similar to F-box family protein *Arabidopsis thaliana* (TAIR:AT1G61340.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN71843.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO43165.1); contains InterPro domain Cyclin-like F-box (InterPro:IPR001810)
254934_at	AT4G11140	CRF1 (CYTOKININ RESPONSE FACTOR 1); DNA binding / transcription factor
259018_at	AT3G07390	AIR12 (Auxin-Induced in Root cultures 12); extracellular matrix structural constituent
267012_at	AT2G39220	PLA IIB/PLP6 (Patatin-like protein 6); nutrient reservoir
264166_at	AT1G65370	meprin and TRAF homology domain-containing protein / MATH domain-containing protein
249775_at	AT5G24160	squalene monooxygenase 1,2 / squalene epoxidase 1,2 (SQP1,2)
254954_at	AT4G10910	unknown protein
247126_at	AT5G66080	protein phosphatase 2C family protein / PP2C family protein
258530_at	AT3G06840	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G49170.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO44815.1)
248471_at	AT5G50840	similar to LINC2 (LITTLE NUCLEI2), protein binding *Arabidopsis thaliana* (TAIR:AT1G13220.2); similar to unnamed protein product *Vitis vinifera* (GB:CAO68982.1); contains domain PTHR16127 (PTHR16127); contains domain PTHR16127:SF5 (PTHR16127:SF5)
251489_at	AT3G59460	similar to F-box family protein *Arabidopsis thaliana* (TAIR:AT3G60040.1)
263960_at	AT2G36200	kinesin motor protein-related
265321_at	AT2G18280	AtTLP2 (TUBBY LIKE PROTEIN 2); phosphoric diester hydrolase/ transcription factor
265349_at	AT2G22610	kinesin motor protein-related
258867_at	AT3G03130	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G17160.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO62214.1)

ATH1 Array identifier	Gene identifier	Gene description
247115_at	AT5G65930	ZWI (ZWICHEL); calmodulin binding / microtubule motor
254341_at	AT4G22130	SRF8 (STRUBBELIG-RECEPTOR FAMILY 8); kinase
251925_at	AT3G54000	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G59050.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO15740.1); contains InterPro domain Uncharacterised conserved protein UCP022260 (InterPro:IPR016802)
263535_at	AT2G24970	similar to unnamed protein product *Vitis vinifera* (GB:CAO44373.1)
262967_at	AT1G75730	similar to hypothetical protein *Vitis vinifera* (GB:CAN68413.1)
248710_at	AT5G48480	Identical to Uncharacterized protein At5g48480 *Arabidopsis thaliana* (GB:Q9LV66); similar to unknown *Populus trichocarpa* (GB:ABK95611.1); contains domain SSF54593 (SSF54593)
263455_at	AT2G22320	unknown protein
266385_at	AT2G14610	PR1 (PATHOGENESIS-RELATED GENE 1)
252412_at	AT3G47295	unknown protein
249659_s_at	AT5G36710; AT5G36800	AT5G36710, similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G36800.1); similar to unknown *Brassica rapa* (GB:ABL97950.1); contains domain PROKAR_LIPOPROTEIN (PS51257); AT5G36800, similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G36710.1); similar to unknown *Brassica rapa* (GB:ABL97950.1); contains domain PROKAR_LIPOPROTEIN (PS51257)
261700_at	AT1G32690	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G35200.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN63850.1)
253702_at	AT4G29900	ACA10 (autoinhibited Ca2+ -ATPase 10); calcium-transporting ATPase/ calmodulin binding
253519_at	AT4G31240	electron carrier
259083_at	AT3G04810	ATNEK2; kinase
250519_at	AT5G08460	GDSL-motif lipase/hydrolase family protein
245612_at	AT4G14440	enoyl-CoA hydratase/isomerase family protein
250189_at	AT5G14410	unknown protein
258529_at	AT3G06740	zinc finger (GATA type) family protein
263441_at	AT2G28620	kinesin motor protein-related
262421_at	AT1G50290	unknown protein
261886_s_at	AT1G80700; AT1G80980	AT1G80700, similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G80980.1); similar to unknown *Picea sitchensis* (GB:ABK25947.1); AT1G80980, similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G80700.1); similar to unknown *Picea sitchensis* (GB:ABK25947.1)
249567_at	AT5G38020	S-adenosyl-L-methionine:carboxyl methyltransferase family protein
253480_at	AT4G31840	plastocyanin-like domain-containing protein
249289_at	AT5G41040	transferase family protein

ATH1 Array identifier	Gene identifier	Gene description
261400_at	AT1G79630	protein phosphatase 2C family protein / PP2C family protein
265605_at	AT2G25540	CESA10 (CELLULOSE SYNTHASE 10); transferase, transferring glycosyl groups
250292_at	AT5G13220	JAS1/JAZ10/TIFY9 (JASMONATE-ZIM-DOMAIN PROTEIN 10)
267555_at	AT2G32765	SUM5 (SUMO 5)
264377_at	AT2G25060	plastocyanin-like domain-containing protein
264746_at	AT1G62300	WRKY6 (WRKY DNA-binding protein 6); transcription factor
258380_at	AT3G16650	PP1/PP2A phosphatases pleiotropic regulator 2 (PRL2)
248020_at	AT5G56490	FAD-binding domain-containing protein
247034_at	AT5G67260	CYCD3;2 (CYCLIN D3;2); cyclin-dependent protein kinase
248697_at	AT5G48370	thioesterase family protein
263919_at	AT2G36470	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G27770.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO15960.1); contains InterPro domain Protein of unknown function DUF868, plant (InterPro:IPR008586)
252148_at	AT3G51280	male sterility MS5, putative
264582_at	AT1G05230	homeobox-leucine zipper family protein / lipid-binding START domain-containing protein
260806_at	AT1G78260	RNA recognition motif (RRM)-containing protein
247437_at	AT5G62490	ATHVA22B (Arabidopsis thaliana HVA22 homologue B)
254964_at	AT4G11080	high mobility group (HMG1/2) family protein
256900_at	AT3G24670	pectate lyase family protein
258218_at	AT3G17998; AT3G18000	AT3G17998, CPuORF30 (Conserved peptide upstream open reading frame 30); AT3G18000, NMT1 (N-METHYLTRANSFERASE 1); phosphoethanolamine N-methyltransferase
261363_at	AT1G41830	SKS6 (SKU5 Similar 6); pectinesterase
259978_at	AT1G76540	CDKB2;1 (CYCLIN-DEPENDENT KINASE B2;1); kinase
249843_at	AT5G23570	SGS3 (SUPPRESSOR OF GENE SILENCING 3)
250794_at	AT5G05270	chalcone-flavanone isomerase family protein
262081_at	AT1G59540	ZCF125; microtubule motor
253804_at	AT4G28230	similar to unnamed protein product *Vitis vinifera* (GB:CAO41238.1)
261712_at	AT1G32780	alcohol dehydrogenase, putative
246505_at	AT5G16250	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G02640.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO50168.1)
249547_at	AT5G38160	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
250859_at	AT5G04660	CYP77A4 (cytochrome P450, family 77, subfamily A, polypeptide 4); oxygen binding
245259_at	AT4G14150	PAKRP1 (PHRAGMOPLAST-ASSOCIATED KINESIN-RELATED PROTEIN 1); microtubule motor/ plus-end-directed microtubule motor
253776_at	AT4G28390	AAC3 (ADP/ATP CARRIER 3); ATP:ADP antiporter/ binding

ATH1 Array identifier	Gene identifier	Gene description
257625_at	AT3G26230	CYP71B24 (cytochrome P450, family 71, subfamily B, polypeptide 24); oxygen binding
251078_at	AT5G01990	auxin efflux carrier family protein
264114_at	AT2G31270	ATCDT1A/CDT1/CDT1A (ARABIDOPSIS HOMOLOG OF YEAST CDT1 A); cyclin-dependent protein kinase/ protein binding
260213_at	AT1G74490	protein kinase, putative
253571_at	AT4G31000	calmodulin-binding protein
256595_x_at	AT3G28530; AT1G36770	AT3G28530, UDP-glucose 4-epimerase; AT1G36770, transposable element gene
247671_at	AT5G60210	cytoplasmic linker protein-related
248568_at	AT5G49760	leucine-rich repeat family protein / protein kinase family protein
255675_at	AT4G00480	ATMYC1 (Arabidopsis thaliana myc-related transcription factor 1); DNA binding / transcription factor
248843_at	AT5G46880	HB-7 (homeobox-7); DNA binding / transcription factor
261055_at	AT1G01300	aspartyl protease family protein
257978_at	AT3G20860	ATNEK5; kinase
252549_at	AT3G45860	receptor-like protein kinase, putative
252646_at	AT3G44610	protein kinase family protein
245401_at	AT4G17670	senescence-associated protein-related
260264_at	AT1G68500	similar to hypothetical protein *Vitis vinifera* (GB:CAN66643.1)
255597_at	AT4G01730	zinc finger (DHHC type) family protein
253805_at	AT4G28260	similar to unnamed protein product *Vitis vinifera* (GB:CAO69726.1)
247962_at	AT5G56580	ATMKK6 (ARABIDOPSIS NQK1); kinase
246071_at	AT5G20150	SPX (SYG1/Pho81/XPR1) domain-containing protein
259563_s_at	AT1G20590; AT1G20610	AT1G20590, cyclin, putative; AT1G20610, CYCB2;3 (CYCLIN B2;3); cyclin-dependent protein kinase regulator
251521_at	AT3G59420	ACR4 (ARABIDOPSIS CRINKLY4); kinase
248413_at	AT5G51600	PLE (PLEIADE)
247599_at	AT5G60880	unknown protein
256259_at	AT3G12460	3'-5' exonuclease/ nucleic acid binding
248054_at	AT5G55820	similar to unnamed protein product *Vitis vinifera* (GB:CAO45135.1)
255644_at	AT4G00870	basic helix-loop-helix (bHLH) family protein
245265_at	AT4G14400	ACD6 (ACCELERATED CELL DEATH 6); protein binding
250614_at	AT5G07260	homeobox protein-related
261196_at	AT1G12860	basic helix-loop-helix (bHLH) family protein / F-box family protein
261502_at	AT1G14440	ATHB31 (ARABIDOPSIS THALIANA HOMEOBOX PROTEIN 31); transcription factor

ATH1 Array identifier	Gene identifier	Gene description
260932_s_at	AT1G02530; AT1G02520	AT1G02530, PGP12 (P-GLYCOPROTEIN 12); ATPase, coupled to transmembrane movement of substances; AT1G02520, PGP11 (P-GLYCOPROTEIN 11); ATPase, coupled to transmembrane movement of substances
255236_at	AT4G05520	calcium-binding EF hand family protein
259290_at	AT3G11520	CYCB1;3 (CYCLIN B1;3); cyclin-dependent protein kinase regulator
256626_at	AT3G20015	pepsin A
249644_at	AT5G37010	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G65710.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN76723.1)
252533_at	AT3G46110	signal transducer
253403_at	AT4G32830	ATAUR1 (ATAURORA1); histone serine kinase(H3-S10 specific) / kinase/ protein serine/threonine kinase
266824_at	AT2G22800	HAT9 (homeobox-leucine zipper protein 9); DNA binding / transcription factor
249494_at	AT5G39050	transferase family protein
264769_at	AT1G61350	armadillo/beta-catenin repeat family protein
247753_at	AT5G59070	glycosyl transferase family 1 protein
248820_at	AT5G47060	senescence-associated protein-related
253791_at	AT4G28640	IAA11 (indoleacetic acid-induced protein 11); transcription factor
256122_at	AT1G18180	oxidoreductase, acting on the CH-CH group of donors
261454_at	AT1G21090	hydroxyproline-rich glycoprotein family protein
261459_at	AT1G21100	O-methyltransferase, putative
247039_at	AT5G67270	ATEB1C (MICROTUBULE END BINDING PROTEIN 1); microtubule binding
264802_at	AT1G08560	SYP111 (syntaxin 111); SNAP receptor
246842_at	AT5G26731	unknown protein
257951_at	AT3G21700	GTP binding
259912_at	AT1G72670	IQD8 (IQ-domain 8); calmodulin binding
261834_at	AT1G10640	polygalacturonase, putative / pectinase, putative
266735_at	AT2G46930	pectinacetylesterase, putative
253109_at	AT4G35920	MCA1 (MID1-COMPLEMENTING ACTIVITY 1)
265586_at	AT2G19990	PR-1-LIKE (PATHOGENESIS-RELATED PROTEIN-1-LIKE)
265902_at	AT2G25590	agenet domain-containing protein
257409_at	AT2G17470	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G25480.1); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G68600.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO42118.1); contains InterPro domain Protein of unknown function UPF0005 (InterPro:IPR006214)
262539_at	AT1G17200	integral membrane family protein
256337_at	AT1G72060	serine-type endopeptidase inhibitor
249780_at	AT5G24240	phosphatidylinositol 3- and 4-kinase family protein / ubiquitin family protein

ATH1 Array identifier	Gene identifier	Gene description
247463_at	AT5G62210	embryo-specific protein-related
262171_at	AT1G74950	JAZ2/TIFY10B (JASMONATE-ZIM-DOMAIN PROTEIN 2)
262615_at	AT1G13950	EIF-5A (eukaryotic translation initiation factor 5A-1); translation initiation factor
264893_at	AT1G23140	C2 domain-containing protein
260876_at	AT1G21460	nodulin MtN3 family protein
245343_at	AT4G15830	binding
257624_at	AT3G26220	CYP71B3 (cytochrome P450, family 71, subfamily B, polypeptide 3); oxygen binding
257382_at	AT2G40750	WRKY54 (WRKY DNA-binding protein 54); transcription factor
248435_at	AT5G51210	OLEO3 (OLEOSIN3)
248819_at	AT5G47050	ATP binding / protein binding / shikimate kinase/ zinc ion binding
260135_at	AT1G66400	calmodulin-related protein, putative
262682_at	AT1G75900	family II extracellular lipase 3 (EXL3)
254524_at	AT4G20000	VQ motif-containing protein
264343_at	AT1G11850	unknown protein
246557_at	AT5G15510	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G01015.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO15061.1); contains InterPro domain Targeting for Xklp2 (InterPro:IPR009675)
252045_at	AT3G52450	U-box domain-containing protein
252089_at	AT3G52110	similar to hypothetical protein OsI_009975 *Oryza sativa* (*indica* cultivar-group) (GB:EAY88742.1); similar to Os03g0174200 *Oryza sativa* (*japonica* cultivar-group) (GB:NP_001049125.1)
253795_at	AT4G28420	aminotransferase, putative
267618_at	AT2G26760	CYCB1;4; cyclin-dependent protein kinase regulator
260608_at	AT2G43870	polygalacturonase, putative / pectinase, putative
251039_at	AT5G02020	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G55646.1); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G59080.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO15731.1)
257722_at	AT3G18490	aspartyl protease family protein
245800_at	AT1G46264	AT-HSFB4 (Arabidopsis thaliana heat shock transcription factor B4); DNA binding / transcription factor
266226_at	AT2G28740	HIS4 (Histone H4)
258160_at	AT3G17820	ATGSKB6 (Arabidopsis thaliana glutamine synthase clone KB6); glutamate-ammonia ligase
252347_at	AT3G48130	ribosomal protein L13 homolog
245194_at	AT1G67820	protein phosphatase 2C, putative / PP2C, putative
257964_at	AT3G19850	phototropic-responsive NPH3 family protein
261832_at	AT1G10650	protein binding / zinc ion binding
258409_at	AT3G17640	leucine-rich repeat family protein
250694_at	AT5G06710	HAT14 (homeobox-leucine zipper protein 14); DNA binding / transcription factor

ATH1 Array identifier	Gene identifier	Gene description
252821_at	AT4G39860	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G22270.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO39818.1)
248204_at	AT5G54280	ATM2 (ARABIDOPSIS THALIANA MYOSIN 4)
254188_at	AT4G23920	UGE2 (UDP-D-GLUCOSE/UDP-D-GALACTOSE 4-EPIMERASE 2); UDP-glucose 4-epimerase/ protein dimerization
249809_at	AT5G23910	microtubule motor
257735_at	AT3G27400	pectate lyase family protein
264001_at	AT2G22420	peroxidase 17 (PER17) (P17)
252608_at	AT3G45090	2-phosphoglycerate kinase-related
245584_at	AT4G14940	ATAO1 (Arabidosis thaliana amine oxidase 1); copper ion binding
258082_at	AT3G25905	CLE27 (CLAVATA3/ESR-RELATED 27); receptor binding
265672_at	AT2G31980	cysteine proteinase inhibitor-related
248091_at	AT5G55120	VTC5; galactose-1-phosphate guanylyltransferase (GDP)
263787_at	AT2G46420	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G61700.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO39750.1); contains InterPro domain Conserved hypothetical protein CHP01589, plant (InterPro:IPR006476)
267008_at	AT2G39350	ABC transporter family protein
248625_at	AT5G48880	KAT5/PKT1/PKT2 (PEROXISOMAL 3-KETO-ACYL-COA THIOLASE 1, PEROXISOMAL 3-KETO-ACYL-COA THIOLASE 2); acetyl-CoA C-acyltransferase
249916_at	AT5G22880	H2B/HTB2 (HISTONE H2B); DNA binding
257290_at	AT3G15560	similar to unnamed protein product *Vitis vinifera* (GB:CAO46019.1)
246814_at	AT5G27200	ACP5 (ACYL CARRIER PROTEIN 5); acyl carrier
260673_at	AT1G19330	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G75060.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO61216.1)
264574_at	AT1G05300	ZIP5 (ZINC TRANSPORTER 5 PRECURSOR); cation transmembrane transporter
263098_at	AT2G16005	MD-2-related lipid recognition domain-containing protein / ML domain-containing protein
267080_at	AT2G41190	amino acid transporter family protein
252351_at	AT3G48210	similar to unnamed protein product *Vitis vinifera* (GB:CAO64278.1); contains InterPro domain Kinetochore-Ndc80 subunit Spc25 (InterPro:IPR013255)
251130_at	AT5G01180	proton-dependent oligopeptide transport (POT) family protein
248812_at	AT5G47330	palmitoyl protein thioesterase family protein
263607_at	AT2G16270	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G16630.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN75064.1)
248597_at	AT5G49160	MET1 (DECREASED METHYLATION 2DNA)

ATH1 Array identifier	Gene identifier	Gene description
245976_at	AT5G13080	WRKY75 (WRKY DNA-BINDING PROTEIN 75); transcription factor
264958_at	AT1G76960	unknown protein
261660_at	AT1G18370	HIK (HINKEL); microtubule motor
245576_at	AT4G14770	tesmin/TSO1-like CXC domain-containing protein
266738_at	AT2G47010	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G17030.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN67131.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO40236.1)
258067_at	AT3G25980	mitotic spindle checkpoint protein, putative (MAD2)
264465_at	AT1G10230	ASK18 (ARABIDOPSIS SKP1-LIKE 18); protein binding / ubiquitin-protein ligase
252051_at	AT3G52570	similar to hydrolase, alpha/beta fold family protein *Arabidopsis thaliana* (TAIR:AT4G10030.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO15880.1); contains domain PTHR10992:SF17 (PTHR10992:SF17); contains domain PTHR10992 (PTHR10992); contains domain SSF53474 (SSF53474); contains domain G3DSA:3.40.50.1820 (G3DSA:3.40.50.1820)
258197_at	AT3G14000	ATBRXL2/BRX-LIKE2
259513_at	AT1G12430	PAK (PHOSPHATIDIC ACID KINASE); microtubule motor
265771_at	AT2G48030	endonuclease/exonuclease/phosphatase family protein
259561_at	AT1G21250	WAK1 (CELL WALL-ASSOCIATED KINASE); kinase
263352_at	AT2G22080	unknown protein
256402_at	AT3G06130	heavy-metal-associated domain-containing protein
248021_at	AT5G56500	ATP binding / protein binding / unfolded protein binding
259976_at	AT1G76560	CP12-3
258323_at	AT3G22750	protein kinase, putative
260276_at	AT1G80450	VQ motif-containing protein
248752_at	AT5G47600	heat shock protein-related
259373_at	AT1G69160	unknown protein
247020_at	AT5G67020	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G50340.1); similar to unknown protein *Oryza sativa* (*japonica* cultivar-group) (GB:BAD09363.1); contains domain DHS-like NAD/FAD-binding domain (SSF52467)
248807_at	AT5G47500	pectinesterase family protein
257643_at	AT3G25730	AP2 domain-containing transcription factor, putative
262801_at	AT1G21010	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G76600.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO66314.1)
247603_at	AT5G60930	chromosome-associated kinesin, putative
246349_at	AT1G51915	cryptdin protein-related
266175_at	AT2G02450	ANAC034/ANAC035 (Arabidopsis NAC domain containing protein 34, Arabidopsis NAC domain containing protein 35); transcription factor
256012_at	AT1G19250	FMO1 (FLAVIN-DEPENDENT MONOOXYGENASE 1); monooxygenase

ATH1 Array identifier	Gene identifier	Gene description
246184_at	AT5G20950	glycosyl hydrolase family 3 protein
262275_at	AT1G68710	haloacid dehalogenase-like hydrolase family protein
257740_at	AT3G27330	zinc finger (C3HC4-type RING finger) family protein
262758_at	AT1G10780	F-box family protein
262376_at	AT1G72970	HTH (HOTHEAD); aldehyde-lyase
249012_at	AT5G44620	CYP706A3 (cytochrome P450, family 706, subfamily A, polypeptide 3); oxygen binding
258310_at	AT3G26744	ICE1 (INDUCER OF CBF EXPRESSION 1); DNA binding / transcription factor
251388_at	AT3G60840	microtubule associated protein (MAP65/ASE1) family protein
249060_at	AT5G44560	VPS2.2
262350_at	AT2G48150	ATGPX4 (GLUTATHIONE PEROXIDASE 4); glutathione peroxidase
249726_at	AT5G35480	unknown protein
262128_at	AT1G52690	late embryogenesis abundant protein, putative / LEA protein, putative
264725_at	AT1G22885	unknown protein
247116_at	AT5G65970	MLO10 (MILDEW RESISTANCE LOCUS O 10); calmodulin binding
264307_at	AT1G61900	Identical to Uncharacterized GPI-anchored protein At1g61900 precursor *Arabidopsis thaliana* (GB:Q8GUI4;GB:O80696;GB:Q3ECK6); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G30700.1); similar to Os07g0102300 *Oryza sativa* (*japonica* cultivar-group) (GB:NP_001058682.1); similar to unknown protein *Oryza sativa* (*japonica* cultivar-group) (GB:BAD31837.1)
263981_at	AT2G42870	HLH1/PAR1 (PHY RAPIDLY REGULATED 1); transcription regulator
248169_at	AT5G54610	ANK (ANKYRIN); protein binding
265323_at	AT2G18260	SYP112 (syntaxin 112); SNAP receptor
264616_at	AT2G17740	DC1 domain-containing protein
261413_at	AT1G07630	PLL5 (POL-like 5); protein serine/threonine phosphatase
265935_at	AT2G19580	TET2 (TETRASPANIN2)
259661_at	AT1G55265	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G19860.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN64354.1); contains InterPro domain Protein of unknown function DUF538 (InterPro:IPR007493)
248247_at	AT5G53210	SPCH (SPEECHLESS); DNA binding / transcription factor
259528_at	AT1G12330	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G12900.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO16971.1)
256397_at	AT3G06110	MKP2; protein tyrosine/serine/threonine phosphatase
259151_at	AT3G10310	kinesin motor protein-related
262176_at	AT1G74960	FAB1 (FATTY ACID BIOSYNTHESIS 1); fatty-acid synthase
247924_at	AT5G57655	xylose isomerase family protein
266078_at	AT2G40670	ARR16 (response regulator 16); transcription regulator/ two-component response regulator

ATH1 Array identifier	Gene identifier	Gene description
262702_at	AT1G16220	protein phosphatase 2C family protein / PP2C family protein
258326_at	AT3G22760	SOL1 (TSO1-Like); transcription factor
263016_at	AT1G23410	ubiquitin extension protein, putative / 40S ribosomal protein S27A (RPS27aA)
256597_at	AT3G28500	60S acidic ribosomal protein P2 (RPP2C)
259851_at	AT1G72250	kinesin motor protein-related
255637_at	AT4G00750	dehydration-responsive family protein
258091_at	AT3G14560	unknown protein
248596_at	AT5G49330	AtMYB111 (myb domain protein 111); DNA binding / transcription factor
245141_at	AT2G45400	BEN1; oxidoreductase, acting on CH-OH group of donors
253038_at	AT4G37790	HAT22 (homeobox-leucine zipper protein 22); transcription factor
253226_at	AT4G35010	BGAL11 (beta-galactosidase 11); beta-galactosidase
258387_at	AT3G15550	similar to unnamed protein product *Vitis vinifera* (GB:CAO46019.1)
247590_at	AT5G60720	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G39690.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO21880.1); contains InterPro domain Protein of unknown function DUF547 (InterPro:IPR006869)
267613_at	AT2G26700	protein kinase family protein
253720_at	AT4G29270	acid phosphatase class B family protein
250158_at	AT5G15190	unknown protein
261765_at	AT1G15570	CYCA2;3 (CYCLIN A2;3); cyclin-dependent protein kinase regulator
246682_at	AT5G33290	XGD1 (XYLOGALACTURONAN DEFICIENT 1); catalytic
257260_at	AT3G22104	phototropic-responsive NPH3 protein-related
266239_at	AT2G29530	TIM10 (Arabidopsis thaliana translocase inner membrane subunit 10); P-P-bond-hydrolysis-driven protein transmembrane transporter
258623_at	AT3G02790	zinc finger (C2H2 type) family protein
262109_at	AT1G02730	ATCSLD5 (CELLULOSE SYNTHASE-LIKE D5); 1,4-beta-D-xylan synthase/ cellulose synthase
248623_at	AT5G49170	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G06840.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO44815.1)
264763_at	AT1G61450	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G61415.1); similar to Os05g0275600 *Oryza sativa* (*japonica* cultivar-group) (GB:NP_001055074.1)
255228_at	AT4G05470	F-box family protein (FBL21)
261427_at	AT1G18670	IBS1 (IMPAIRED IN BABA-INDUCED STERILITY 1); kinase
249934_at	AT5G22410	peroxidase, putative
263528_at	AT2G24800	peroxidase, putative

ATH1 Array identifier	Gene identifier	Gene description
266956_at	AT2G34510	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G29980.1); similar to unknown *Populus trichocarpa* (GB:ABK95079.1); contains InterPro domain Protein of unknown function DUF642 (InterPro:IPR006946); contains InterPro domain Galactose-binding like (InterPro:IPR008979)
263584_at	AT2G17040	ANAC036 (Arabidopsis NAC domain containing protein 36); transcription factor
265330_at	AT2G18440	GUT15 (GENE WITH UNSTABLE TRANSCRIPT 15); other RNA
253239_at	AT4G34500	protein kinase family protein
265208_at	AT2G36690	oxidoreductase, 2OG-Fe(II) oxygenase family protein
245560_at	AT4G15480	UGT84A1; UDP-glycosyltransferase/ sinapate 1-glucosyltransferase/ transferase, transferring glycosyl groups
258962_at	AT3G10570	CYP77A6 (cytochrome P450, family 77, subfamily A, polypeptide 6); oxygen binding
261893_at	AT1G80690	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G25170.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO45044.1); contains InterPro domain Protein of unknown function DUF862, eukaryotic (InterPro:IPR008580)
252078_at	AT3G51740	IMK2 (INFLORESCENCE MERISTEM RECEPTOR-LIKE KINASE 2); ATP binding / kinase/ protein serine/threonine kinase
256832_at	AT3G22880	ATDMC1 (RECA-LIKE GENE); ATP binding / DNA-dependent ATPase/ damaged DNA binding
252983_at	AT4G37980	ELI3-1 (ELICITOR-ACTIVATED GENE 3); binding / catalytic/ oxidoreductase/ zinc ion binding
249885_at	AT5G22940	exostosin family protein
260804_at	AT1G78410	VQ motif-containing protein
254233_at	AT4G23800	high mobility group (HMG1/2) family protein
267076_at	AT2G41090	calmodulin-like calcium-binding protein, 22 kDa (CaBP-22)
247759_at	AT5G59040	COPT3 (Copper transporter 3); copper ion transmembrane transporter
258376_at	AT3G17680	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G48405.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN67913.1); contains InterPro domain KIP1-like (InterPro:IPR011684)
257460_at	AT1G75580	auxin-responsive protein, putative
256239_at	AT3G12470	nucleic acid binding
247723_at	AT5G59220	protein phosphatase 2C, putative / PP2C, putative
259235_at	AT3G11600	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G06270.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO15841.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN79170.1)
258264_at	AT3G15790	MBD11 (methyl-CpG-binding domain 11); DNA binding
253754_at	AT4G29020	glycine-rich protein
245669_at	AT1G28300	LEC2 (LEAFY COTYLEDON 2); transcription factor

ATH1 Array identifier	Gene identifier	Gene description
265837_at	AT2G14560	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G33840.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO48825.1); contains InterPro domain Protein of unknown function DUF567 (InterPro:IPR007612)
255203_at	AT4G07510	transposable element gene
248347_at	AT5G52250	transducin family protein / WD-40 repeat family protein
250992_at	AT5G02260	ATEXPA9 (ARABIDOPSIS THALIANA EXPANSIN A9)
256570_at	AT3G19540	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G49840.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO70870.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN68554.1); contains InterPro domain Protein of unknown function DUF620 (InterPro:IPR006873)
260151_at	AT1G52910	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G15480.1); similar to unknown *Populus trichocarpa* (GB:ABK94458.1); contains InterPro domain Protein of unknown function DUF1218 (InterPro:IPR009606)
252877_at	AT4G39630	similar to unnamed protein product *Vitis vinifera* (GB:CAO71785.1)
248355_at	AT5G52340	ATEXO70A2 (exocyst subunit EXO70 family protein A2); protein binding
248484_at	AT5G51030	short-chain dehydrogenase/reductase (SDR) family protein
260997_at	AT1G26610	zinc finger (C2H2 type) family protein
254484_at	no_match	no_match
253270_at	AT4G34160	CYCD3/CYCD3;1/D3 (CYCLIN D3;1); cyclin-dependent protein kinase regulator/ protein binding
244939_at	ATCG00065	chloroplast gene encoding ribosomal protein s12. The gene is located in three distinct loci on the chloroplast genome and is transpliced to make one transcript.
254609_at	AT4G18970	GDSL-motif lipase/hydrolase family protein
251046_at	AT5G02370	kinesin motor protein-related
258859_at	AT3G02120	hydroxyproline-rich glycoprotein family protein
262494_at	AT1G21810	similar to myosin heavy chain-related *Arabidopsis thaliana* (TAIR:AT1G77580.2); similar to coiled-coil protein *Lycopersicon esculentum* (GB:AAN03605.1); contains InterPro domain Prefoldin (InterPro:IPR009053); contains InterPro domain Protein of unknown function DUF869, plant (InterPro:IPR008587)
260531_at	AT2G47240	long-chain-fatty-acid–CoA ligase family protein / long-chain acyl-CoA synthetase family protein
251952_at	AT3G53650	histone H2B, putative
256258_at	AT3G12480	transcription factor, putative
251979_at	AT3G53140; AT3G53130	AT3G53140, O-diphenol-O-methyl transferase, putative; AT3G53130, LUT1 (LUTEIN DEFICIENT 1); oxygen binding
262516_at	AT1G17190	ATGSTU26 (Arabidopsis thaliana Glutathione S-transferase (class tau) 26); glutathione transferase

ATH1 Array identifier	Gene identifier	Gene description
261451_at	AT1G21060	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G76620.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO66316.1); contains InterPro domain Protein of unknown function DUF547 (InterPro:IPR006869)
249329_at	AT5G40960	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G48660.1); similar to unknown *Populus trichocarpa* (GB:ABK93721.1)
247047_at	AT5G66650	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G23790.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO22909.1); contains InterPro domain Protein of unknown function DUF607 (InterPro:IPR006769)
251858_at	AT3G54820	PIP2;5/PIP2D (plasma membrane intrinsic protein 2;5); water channel
245739_at	AT1G44110	CYCA1;1 (CYCLIN A1;1); cyclin-dependent protein kinase regulator
254380_at	no_match	no_match
255265_at	AT4G05190	ATK5 (Arabidopsis thaliana kinesin 5); microtubule motor
247425_at	AT5G62550	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G28770.1); similar to unknown *Populus trichocarpa* (GB:ABK92769.1)
254241_at	AT4G23190	CRK11 (CYSTEINE-RICH RLK11); kinase
248527_at	AT5G50740	metal ion binding
248057_at	AT5G55520	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G26660.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO38858.1); contains InterPro domain Kinesin-related (InterPro:IPR010544)
250558_at	AT5G07990	TT7 (TRANSPARENT TESTA 7); flavonoid 3'-monooxygenase/ oxygen binding
267495_at	no_match	no_match
248037_at	AT5G55930	ATOPT1 (oligopeptide transporter 1); oligopeptide transporter
251791_at	AT3G55500	ATEXPA16 (ARABIDOPSIS THALIANA EXPANSIN A16)
262788_at	AT1G10690	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G60783.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN59885.1)
250228_at	AT5G13840	WD-40 repeat family protein
261181_at	AT1G34580	monosaccharide transporter, putative
262092_at	AT1G56150	auxin-responsive family protein
266098_at	AT2G37870	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
265660_at	AT2G25470	leucine-rich repeat family protein
262136_at	AT1G77850	ARF17 (AUXIN RESPONSE FACTOR 17); transcription factor
254605_at	AT4G18950	ankyrin protein kinase, putative
263783_at	AT2G46400	WRKY46 (WRKY DNA-binding protein 46); transcription factor
256291_at	AT3G12200	ATNEK7 (NIMA-RELATED KINASE7); kinase
252387_at	AT3G47800	aldose 1-epimerase family protein
261511_at	AT1G71770	PAB5 (POLY(A)-BINDING PROTEIN); RNA binding

ATH1 Array identifier	Gene identifier	Gene description
245513_at	AT4G15780	ATVAMP724 (Arabidopsis thaliana vesicle-associated membrane protein 724)
265587_at	AT2G19980	allergen V5/Tpx-1-related family protein
250581_at	AT5G07300	BON2 (BONZAI 2)
261949_at	AT1G64670	BDG1 (BODYGUARD1); hydrolase
262438_at	AT1G47410	unknown protein
264532_at	AT1G55740	ATSIP1 (ARABIDOPSIS THALIANA SEED IMBIBITION 1); hydrolase, hydrolyzing O-glycosyl compounds
256896_at	AT3G24630	similar to unnamed protein product *Vitis vinifera* (GB:CAO65510.1)
253815_at	AT4G28250	ATEXPB3 (ARABIDOPSIS THALIANA EXPANSIN B3)
248427_at	AT5G51750	ATSBT1.3; subtilase
263378_at	AT2G40180	ATHPP2C5; protein serine/threonine phosphatase
262783_at	AT1G10850	ATP binding / protein serine/threonine kinase
264836_at	AT1G03610	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G03420.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO24103.1); contains InterPro domain Protein of unknown function DUF789 (InterPro:IPR008507)
249548_at	AT5G38170	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
261012_at	AT1G26600	CLE9 (CLAVATA3/ESR-RELATED 9); receptor binding
261455_at	AT1G21070	transporter-related
245285_s_at	AT4G14030; AT4G14040	AT4G14030, selenium-binding protein, putative; AT4G14040, EDA38 (embryo sac development arrest 38); selenium binding
262773_at	AT1G13220	LINC2 (LITTLE NUCLEI2); protein binding
263788_at	AT2G24580	sarcosine oxidase family protein
255460_at	AT4G02800	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G30050.1); similar to Os04g0228100 *Oryza sativa* (*japonica* cultivar-group) (GB:NP_001052288.1); similar to H0209A05.2 *Oryza sativa* (*indica* cultivar-group) (GB:CAH66085.1)
247643_at	AT5G60450	ARF4 (AUXIN RESPONSE FACTOR 4); transcription factor
257243_at	AT3G24230	pectate lyase family protein
245607_at	AT4G14330	phragmoplast-associated kinesin-related protein 2 (PAKRP2)
249364_at	AT5G40590	DC1 domain-containing protein
259858_at	AT1G68400	leucine-rich repeat transmembrane protein kinase, putative
260119_at	AT1G33930	avirulence-responsive family protein / avirulence induced gene (AIG1) family protein
257203_at	AT3G23730	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative

ATH1 Array identifier	Gene identifier	Gene description
266733_at	AT2G03280	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G14020.1); similar to hypothetical protein OsJ_000519 *Oryza sativa* (*japonica* cultivar-group) (GB:EAZ10694.1); similar to hypothetical protein OsI_000541 *Oryza sativa* (*indica* cultivar-group) (GB:EAY72694.1); contains InterPro domain Protein of unknown function DUF246, plant (InterPro:IPR004348)
252164_at	AT3G50620	nodulation protein-related
256599_at	AT3G14760	similar to unnamed protein product *Vitis vinifera* (GB:CAO39487.1)
267414_at	AT2G34790	EDA28/MEE23 (MATERNAL EFFECT EMBRYO ARREST 23); electron carrier
245688_at	AT1G28290	AGP31 (ARABINOGALACTAN-PROTEIN 31); structural constituent of cell wall
262040_at	AT1G80080	TMM (TOO MANY MOUTHS); protein binding
260494_at	AT2G41820	leucine-rich repeat transmembrane protein kinase, putative
252227_at	AT3G49900	BTB/POZ domain-containing protein
259969_at	AT1G76550	pyrophosphate–fructose-6-phosphate 1-phosphotransferase alpha subunit, putative / pyrophosphate-dependent 6-phosphofructose-1-kinase, putative
248963_at	AT5G45700	NLI interacting factor (NIF) family protein
265464_at	AT2G37080	myosin heavy chain-related
253096_at	AT4G37330	CYP81D4 (cytochrome P450, family 81, subfamily D, polypeptide 4); oxygen binding
261364_at	AT1G53140	dynamin family protein
247348_at	AT5G63810	BGAL10 (beta-galactosidase 10); beta-galactosidase
245702_at	AT5G04220	ATSYTC/NTMC2T1.3/NTMC2TYPE1.3/SYTC
265184_at	AT1G23710	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G70420.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO66069.1); contains InterPro domain Protein of unknown function DUF1645 (InterPro:IPR012442)
266324_at	AT2G46710	rac GTPase activating protein, putative
246999_at	AT5G67440	phototropic-responsive NPH3 family protein
261384_at	AT1G05440	similar to unnamed protein product *Vitis vinifera* (GB:CAO71189.1)
252879_at	AT4G39390	transporter-related
258480_at	AT3G02640	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G16250.1); similar to unknown *Populus trichocarpa* (GB:ABK93711.1)
251035_at	AT5G02220	similar to unknown *Picea sitchensis* (GB:ABK23883.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN70860.1)
256395_at	AT3G06120	MUTE (MUTE); DNA binding / transcription factor
262971_at	AT1G75640	leucine-rich repeat family protein / protein kinase family protein
261203_at	AT1G12845	similar to hypothetical protein MtrDRAFT_AC149131g9v2 *Medicago truncatula* (GB:ABD32556.1)
254986_at	AT4G10640	IQD16 (IQ-domain 16); calmodulin binding

ATH1 Array identifier	Gene identifier	Gene description
252381_s_at	AT3G47760; AT3G47750	AT3G47760, ATATH4 (ABC2 homolog 4); ATPase, coupled to transmembrane movement of substances; AT3G47750, ATATH3 (ABC2 homolog 3); ATPase, coupled to transmembrane movement of substances
259145_at	AT3G10180	kinesin motor protein-related
252691_at	AT3G44050	kinesin motor protein-related
254301_at	AT4G22790	MATE efflux family protein
257115_at	AT3G20150	kinesin motor family protein
261804_at	AT1G30530	UDP-glucoronosyl/UDP-glucosyl transferase family protein
264833_at	AT1G03590	protein phosphatase 2C family protein / PP2C family protein
250616_at	AT5G07280	EMS1 (EXCESS MICROSPOROCYTES1); kinase

Table 2-S1. Genes showing high expression in *scrm-D mute*.
List of 552 genes representing gene cluster I from Figure 2-2. ATH1 array and respective gene locus identifiers are indicated. Gene descriptions for each gene were derived from the Bio-Array Resource for Plant Biology (BAR; `http://142.150.214.117/welcome.htm`). Coexpression was determined using $\log_2$ transformed normalized expression values from wild type, *spch*, *scrm-D mute*, and *scrm-D* as input. Details of K-means cluster analysis are referenced in Materials and Methods.

Table 2-S2.

ATH1 Array identifier	Gene identifier	Gene description
255732_at	AT1G25450	very-long-chain fatty acid condensing enzyme, putative
251054_at	AT5G01540	lectin protein kinase, putative
266415_at	AT2G38530	LTP2 (LIPID TRANSFER PROTEIN 2); lipid binding
252238_at	AT3G49960	peroxidase, putative
257264_at	AT3G22060	receptor protein kinase-related
258151_at	AT3G18080	glycosyl hydrolase family 1 protein
247144_at	AT5G65590	Dof-type zinc finger domain-containing protein
250793_at	AT5G05600	oxidoreductase, 2OG-Fe(II) oxygenase family protein
262480_at	AT1G11340	S-locus lectin protein kinase family protein
255110_at	AT4G08770	peroxidase, putative
245616_at	AT4G14480	protein kinase family protein
261005_at	AT1G26420	FAD-binding domain-containing protein
267169_at	AT2G37540	short-chain dehydrogenase/reductase (SDR) family protein
246389_at	AT1G77380	AAP3 (amino acid permease 3); amino acid transmembrane transporter
254914_at	AT4G11290	peroxidase, putative
266246_at	AT2G27690	CYP94C1 (cytochrome P450, family 94, subfamily C, polypeptide 1); oxygen binding
267128_at	AT2G23620	esterase, putative
246273_at	AT4G36700	cupin family protein
257890_s_at	AT3G42570; AT3G17070	AT3G42570, peroxidase-related; AT3G17070, peroxidase, putative
252677_at	AT3G44320	NIT3 (NITRILASE 3)
259514_at	AT1G12480	CDI3/OZS1/RCD3/SLAC1 (SLOW ANION CHANNEL-ASSOCIATED 1); transporter
263776_s_at	AT2G46430; AT2G46440	AT2G46430, ATCNGC3 (CYCLIC NUCLEOTIDE GATED CHANNEL 3); calmodulin binding / cyclic nucleotide binding / ion channel; AT2G46440, ATCNGC11 (cyclic nucleotide gated channel 11); calmodulin binding / cyclic nucleotide binding / ion channel
257365_x_at	AT2G26020	PDF1.2b (plant defensin 1.2b)
256922_at	AT3G19010	oxidoreductase, 2OG-Fe(II) oxygenase family protein
249481_at	AT5G38900	DSBA oxidoreductase family protein
249051_at	AT5G44390	FAD-binding domain-containing protein
257774_at	AT3G29250	oxidoreductase
250445_at	AT5G10760	aspartyl protease family protein
251106_at	AT5G01500	mitochondrial substrate carrier family protein
249459_at	AT5G39580	peroxidase, putative
254251_at	AT4G23300	protein kinase family protein
248727_at	AT5G47990	CYP705A5 (cytochrome P450, family 705, subfamily A, polypeptide 5); oxygen binding
260386_at	AT1G74010	strictosidine synthase family protein
250644_at	AT5G06750	protein phosphatase 2C family protein / PP2C family protein
247718_at	AT5G59310	LTP4 (LIPID TRANSFER PROTEIN 4); lipid binding

ATH1 Array identifier	Gene identifier	Gene description
267628_at	AT2G42280	basic helix-loop-helix (bHLH) family protein
256940_at	AT3G30720	unknown protein
259579_at	AT1G28010	PGP14 (P-GLYCOPROTEIN 14); ATPase, coupled to transmembrane movement of substances
261318_at	AT1G53035	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G15358.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO69750.1)
265957_at	AT2G37300	unknown protein
251906_at	AT3G53720	ATCHX20 (CATION/H+ EXCHANGER 20); monovalent cation:proton antiporter
264071_at	AT2G27920	SCPL51; serine carboxypeptidase
255111_at	AT4G08780	peroxidase, putative
257421_at	AT1G12030	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G62420.1); similar to Protein of unknown function DUF506, plant *Medicago truncatula* (GB:ABN05810.1); contains InterPro domain Protein of unknown function DUF506, plant (InterPro:IPR006502)
253099_s_at	AT4G37530; AT4G37520	AT4G37530, peroxidase, putative; AT4G37520, peroxidase 50 (PER50) (P50) (PRXR2)
253054_at	AT4G37580	HLS1 (HOOKLESS 1); N-acetyltransferase
266965_at	AT2G39510	nodulin MtN21 family protein
254255_at	AT4G23220	protein kinase family protein
256243_at	AT3G12500	ATHCHIB (BASIC CHITINASE); chitinase
254754_at	AT4G13210	pectate lyase family protein
258299_at	AT3G23410	alcohol oxidase-related
254321_at	AT4G22590; AT4G22592	AT4G22590, trehalose-6-phosphate phosphatase, putative; AT4G22592, CPuORF27 (Conserved peptide upstream open reading frame 27)
254832_at	AT4G12490	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
248237_at	AT5G53890	leucine-rich repeat transmembrane protein kinase, putative
247604_at	AT5G60950	COBL5 (COBRA-LIKE PROTEIN 5 PRECURSOR)
253999_at	AT4G26200	ACS7 (1-Amino-cyclopropane-1-carboxylate synthase 7); 1-aminocyclopropane-1-carboxylate synthase
267246_at	AT2G30250	WRKY25 (WRKY DNA-binding protein 25); transcription factor
250365_at	AT5G11410	protein kinase family protein
261981_at	AT1G33811	GDSL-motif lipase/hydrolase family protein
264153_at	AT1G65390	ATPP2-A5; carbohydrate binding
245052_at	AT2G26440	pectinesterase family protein
249532_at	AT5G38780	S-adenosyl-L-methionine:carboxyl methyltransferase family protein
258037_at	AT3G21230	4CL5 (4-COUMARATE:COA LIGASE 5); 4-coumarate-CoA ligase
252170_at	AT3G50480	HR4 (HOMOLOG OF RPW8 4)
249184_at	AT5G43020	leucine-rich repeat transmembrane protein kinase, putative
265002_at	AT1G24400	LHT2 (LYSINE HISTIDINE TRANSPORTER 2); amino acid transmembrane transporter

ATH1 Array identifier	Gene identifier	Gene description
253608_at	AT4G30290	ATXTH19 (XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE 19); hydrolase, acting on glycosyl bonds
256757_at	AT3G25620	ABC transporter family protein
264301_at	AT1G78780	pathogenesis-related family protein
262050_at	AT1G80130	binding
264960_at	AT1G76930	ATEXT4 (EXTENSIN 4)
255596_at	AT4G01720	WRKY47 (WRKY DNA-binding protein 47); transcription factor
263536_at	AT2G25000	WRKY60 (WRKY DNA-BINDING PROTEIN 60); transcription factor
255568_at	AT4G01250	WRKY22 (WRKY DNA-binding protein 22); transcription factor
251625_at	AT3G57260	BGL2 (PATHOGENESIS-RELATED PROTEIN 2); glucan 1,3-beta-glucosidase/ hydrolase, hydrolyzing O-glycosyl compounds
263726_at	AT2G13610	ABC transporter family protein
253998_at	AT4G26010	peroxidase, putative
252921_at	AT4G39030	EDS5 (ENHANCED DISEASE SUSCEPTIBILITY 5); antiporter/ transporter
260556_at	AT2G43620	chitinase, putative
258217_at	AT3G17998; AT3G18000	AT3G17998, CPuORF30 (Conserved peptide upstream open reading frame 30); AT3G18000, NMT1 (N-METHYLTRANSFERASE 1); phosphoethanolamine N-methyltransferase
261868_s_at	AT1G11450; AT1G11460	AT1G11450, nodulin MtN21 family protein; AT1G11460, nodulin MtN21 family protein
255032_at	AT4G09500	glycosyltransferase family protein
249619_at	AT5G37500	GORK (GATED OUTWARDLY-RECTIFYING K+ CHANNEL); cyclic nucleotide binding / inward rectifier potassium channel/ outward rectifier potassium channel
248895_at	AT5G46330	FLS2 (FLAGELLIN-SENSITIVE 2); ATP binding / kinase/ protein binding / protein serine/threonine kinase/ transmembrane receptor protein serine/threonine kinase
257247_at	AT3G24140	FMA (FAMA); DNA binding / transcription activator/ transcription factor
263478_at	AT2G31880	leucine-rich repeat transmembrane protein kinase, putative
254833_s_at	AT4G12290; AT4G12280	AT4G12290, copper amine oxidase, putative; AT4G12280, copper amine oxidase family protein
247871_at	AT5G57530	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative
254710_at	AT4G18050	PGP9 (P-GLYCOPROTEIN 9); ATPase, coupled to transmembrane movement of substances
260551_at	AT2G43510	ATTI1 (ARABIDOPSIS THALIANA TRYPSIN INHIBITOR PROTEIN 1)
255524_at	AT4G02330	ATPMEPCRB; pectinesterase

ATH1 Array identifier	Gene identifier	Gene description
253168_at	AT4G35070	similar to SBP1 (S-RIBONUCLEASE BINDING PROTEIN 1), protein binding / zinc ion binding *Arabidopsis thaliana* (TAIR:AT1G45976.1); similar to unknown *Populus trichocarpa* (GB:ABK93023.1); contains InterPro domain S-ribonuclease binding protein, SBP1, pollen (InterPro:IPR017066)
257543_at	AT3G28960	amino acid transporter family protein
265853_at	AT2G42360	zinc finger (C3HC4-type RING finger) family protein
262119_s_at	AT1G02930; AT1G02920	AT1G02930, ATGSTF6 (EARLY RESPONSIVE TO DEHYDRATION 11); glutathione transferase; AT1G02920, ATGSTF7 (GLUTATHIONE S-TRANSFERASE 11); glutathione transferase
267391_at	AT2G44480	glycosyl hydrolase family 1 protein
260541_at	AT2G43530	trypsin inhibitor, putative
265611_at	AT2G25510	unknown protein
247802_at	AT5G58580	ATL63; protein binding / zinc ion binding
254247_at	AT4G23260	protein kinase
248728_at	AT5G48000	CYP708A2 (cytochrome P450, family 708, subfamily A, polypeptide 2); oxygen binding
248611_at	AT5G49520	WRKY48 (WRKY DNA-binding protein 48); transcription factor
255148_at	AT4G08470	MAPKKK10 (Mitogen-activated protein kinase kinase kinase 10); kinase
262930_at	AT1G65690	harpin-induced protein-related / HIN1-related / harpin-responsive protein-related
245275_at	AT4G15210	ATBETA-AMY (BETA-AMYLASE); beta-amylase
267545_at	AT2G32690	pseudogene, glycine-rich protein
255462_at	AT4G02940	oxidoreductase, 2OG-Fe(II) oxygenase family protein
257066_at	AT3G18280	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
258901_at	AT3G05640	protein phosphatase 2C, putative / PP2C, putative
267089_at	AT2G38300	DNA binding / transcription factor
254044_at	AT4G25820	XTR9 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 9); hydrolase, acting on glycosyl bonds
258338_at	AT3G16150	L-asparaginase, putative / L-asparagine amidohydrolase, putative
264012_at	AT2G21080	similar to extracellular ligand-gated ion channel *Arabidopsis thaliana* (TAIR:AT3G20300.1); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G50630.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO43972.1)
247424_at	AT5G62850	ATVEX1 (VEGETATIVE CELL EXPRESSED1)
262913_at	AT1G59960	aldo/keto reductase, putative
250942_at	AT5G03350	legume lectin family protein
246880_s_at	AT5G26000; AT5G25980	AT5G26000, TGG1 (THIOGLUCOSIDE GLUCOHYDROLASE 1); hydrolase, hydrolyzing O-glycosyl compounds; AT5G25980, TGG2 (GLUCOSIDE GLUCOHYDROLASE 2); hydrolase, hydrolyzing O-glycosyl compounds
245266_at	AT4G17070	peptidyl-prolyl cis-trans isomerase
258930_at	AT3G10040	transcription factor

ATH1 Array identifier	Gene identifier	Gene description
255575_at	AT4G01430	nodulin MtN21 family protein
266895_at	AT2G26040	Bet v I allergen family protein
261099_at	AT1G62980	ATEXPA18 (ARABIDOPSIS THALIANA EXPANSIN A18)
255595_at	AT4G01700	chitinase, putative
263791_at	AT2G24520	AHA5 (ARABIDOPSIS H(+)-ATPASE 5); ATPase
261420_at	AT1G07720	beta-ketoacyl-CoA synthase family protein
255116_at	AT4G08850	leucine-rich repeat family protein / protein kinase family protein
260921_at	AT1G21540	AMP-binding protein, putative
261350_at	AT1G79770	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G25840.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO45837.1); contains InterPro domain Protein of unknown function DUF1677, plant (InterPro:IPR012876)
254667_at	AT4G18280	glycine-rich cell wall protein-related
253277_at	AT4G34230	CAD5 (CINNAMYL ALCOHOL DEHYDROGENASE 5); cinnamyl-alcohol dehydrogenase
256860_at	AT3G23840	transferase family protein
262616_at	AT1G06620	2-oxoglutarate-dependent dioxygenase, putative
249346_at	AT5G40780	LHT1 (LYSINE HISTIDINE TRANSPORTER 1); amino acid transmembrane transporter
250648_at	AT5G06760	late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing protein
254372_at	AT4G21620	glycine-rich protein
266613_at	AT2G14900	gibberellin-regulated family protein
257679_at	AT3G20470	encodes a glycine-rich protein that is expressed more abundantly in immature seed pods than in stems and leaves. Expression is not detected in roots or flowers.
266913_at	AT2G45890	ATROPGEF4/ROPGEF4 (KINASE PARTNER PROTEIN-LIKE); Rho guanyl-nucleotide exchange factor/
264947_at	AT1G77020	DNAJ heat shock N-terminal domain-containing protein
265411_at	AT2G16630	proline-rich family protein
267361_at	AT2G39920	acid phosphatase class B family protein
261059_at	AT1G01250	AP2 domain-containing transcription factor, putative
253887_at	AT4G27730	ATOPT6 (oligopeptide transporter 6); oligopeptide transporter
263539_at	AT2G24850	TAT3 (TYROSINE AMINOTRANSFERASE 3); transaminase
251013_at	AT5G02540	short-chain dehydrogenase/reductase (SDR) family protein
248749_at	AT5G47880	ERF1-1 (EUKARYOTIC RELEASE FACTOR 1-1); translation release factor
247327_at	AT5G64120	peroxidase, putative
264213_at	AT1G65390	ATPP2-A5; carbohydrate binding
262694_at	AT1G62790	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
247529_at	AT5G61520	hexose transporter, putative
252133_at	AT3G50900	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G66490.1); similar to hypothetical protein *Thellungiella halophila* (GB:ABB45855.1)

ATH1 Array identifier	Gene identifier	Gene description
267288_at	AT2G23680	stress-responsive protein, putative
265670_s_at	AT2G32190; AT2G32210	AT2G32190, similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G32210.1); similar to unknown *Populus trichocarpa* (GB:ABK92801.1); contains domain PD188784 (PD188784); AT2G32210, similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G32190.1); similar to unknown *Populus trichocarpa* (GB:ABK92801.1); contains domain PD188784 (PD188784)
256017_at	AT1G19180	JAZ1/TIFY10A (JASMONATE-ZIM-DOMAIN PROTEIN 1); protein binding
259222_at	AT3G03680	C2 domain-containing protein
249545_at	AT5G38030	MATE efflux family protein
262232_at	AT1G68600	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G25480.1); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G17470.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO42118.1); contains InterPro domain Protein of unknown function UPF0005 (InterPro:IPR006214)
256170_at	AT1G51790	kinase
265674_at	no_match	no_match
262215_at	AT1G74790	catalytic
261175_at	AT1G04800	glycine-rich protein
262838_at	AT1G14960	major latex protein-related / MLP-related
249147_at	AT5G43330	malate dehydrogenase, cytosolic, putative
261157_at	AT1G34510	peroxidase, putative
254110_at	AT4G25260	invertase/pectin methylesterase inhibitor family protein
266070_at	AT2G18660	EXLB3 (EXPANSIN-LIKE B3 PRECURSOR)
249983_at	AT5G18470	curculin-like (mannose-binding) lectin family protein
258487_at	AT3G02550	LBD41 (LOB DOMAIN-CONTAINING PROTEIN 41)
254314_at	AT4G22470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
259520_at	AT1G12320	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G62840.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO24498.1); contains InterPro domain Protein of unknown function DUF1442 (InterPro:IPR009902)
263809_at	AT2G04570	GDSL-motif lipase/hydrolase family protein
248725_at	AT5G47980	transferase family protein
260546_at	AT2G43520	ATTI2 (ARABIDOPSIS THALIANA TRYPSIN INHIBITOR PROTEIN 2); trypsin inhibitor
246238_at	AT4G36670	mannitol transporter, putative
266743_at	AT2G02990	RNS1 (RIBONUCLEASE 1); endoribonuclease
257876_at	AT3G17130	invertase/pectin methylesterase inhibitor family protein
257061_at	AT3G18250	contains domain PROKAR_LIPOPROTEIN (PS51257)
254201_at	AT4G24130	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G56580.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO62919.1); contains InterPro domain Protein of unknown function DUF538 (InterPro:IPR007493)

ATH1 Array identifier	Gene identifier	Gene description
247462_at	AT5G62080	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
246943_at	AT5G25440	protein kinase family protein
262324_at	AT1G64170	ATCHX16 (CATION/H+ EXCHANGER 16); monovalent cation:proton antiporter
261826_at	AT1G11580	ATPMEPCRA; pectinesterase
263498_at	AT2G42610	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G07090.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN66524.1); contains InterPro domain Protein of unknown function DUF640 (InterPro:IPR006936)
250435_at	AT5G10380	zinc finger (C3HC4-type RING finger) family protein
254828_at	AT4G12550	AIR1 (Auxin-Induced in Root cultures 1); lipid binding
255467_at	AT4G03010	leucine-rich repeat family protein
260353_at	AT1G69230	SP1L2
260181_at	AT1G70710	AtGH9B1 (ARABIDOPSIS THALIANA GLYCOSYL HYDROLASE 9B1); hydrolase, hydrolyzing O-glycosyl compounds
265561_s_at	AT2G05440; AT2G05510	AT2G05440, glycine-rich protein; AT2G05510, glycine-rich protein
266142_at	AT2G39030	GCN5-related N-acetyltransferase (GNAT) family protein
254818_at	AT4G12470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
263406_at	AT2G04160	AIR3 (Auxin-Induced in Root cultures 3); subtilase
256169_at	AT1G51800	leucine-rich repeat protein kinase, putative
259925_at	AT1G75040	PR5 (PATHOGENESIS-RELATED GENE 5)
246340_s_at	AT3G44860; AT3G44870	AT3G44860, FAMT (FARNESOIC ACID CARBOXYL-O-METHYLTRANSFERASE); S-adenosylmethionine-dependent methyltransferase/ farnesoic acid O-methyltransferase; AT3G44870, S-adenosyl-L-methionine:carboxyl methyltransferase family protein
263465_at	AT2G31940	oxidoreductase/ transition metal ion binding
252915_at	AT4G38810	calcium-binding EF hand family protein
267472_at	AT2G02850	ARPN (PLANTACYANIN); copper ion binding
260203_at	AT1G52890	ANAC019 (Arabidopsis NAC domain containing protein 19); transcription factor
260254_at	AT1G74210	glycerophosphoryl diester phosphodiesterase family protein
264148_at	AT1G02220	ANAC003 (Arabidopsis NAC domain containing protein 3); transcription factor
255632_at	AT4G00680	actin-depolymerizing factor, putative
261776_at	AT1G76190	auxin-responsive family protein
254869_at	AT4G11890	protein kinase family protein
265102_at	AT1G30870	cationic peroxidase, putative
262637_at	AT1G06640	2-oxoglutarate-dependent dioxygenase, putative
264939_at	AT1G60630	leucine-rich repeat family protein
267388_at	AT2G44450	glycosyl hydrolase family 1 protein

ATH1 Array identifier	Gene identifier	Gene description
266292_at	AT2G29350	SAG13 (Senescence-associated gene 13); oxidoreductase
261369_at	AT1G53060	legume lectin family protein
254663_at	AT4G18290	KAT2 (K+ ATPase 2); cyclic nucleotide binding / inward rectifier potassium channel
249777_at	AT5G24210	lipase class 3 family protein
256874_at	AT3G26320	CYP71B36 (cytochrome P450, family 71, subfamily B, polypeptide 36); oxygen binding
248888_at	AT5G46240	KAT1 (K+ ATPASE 1); cyclic nucleotide binding / inward rectifier potassium channel
249037_at	AT5G44130	FLA13 (FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 13 PRECURSOR)
266600_at	AT2G46070	ATMPK12 (Arabidopsis thaliana MAP kinase 12); MAP kinase/ kinase
264635_at	AT1G65500	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G65486.1)
248276_at	AT5G53550	YSL3 (YELLOW STRIPE LIKE 3); oligopeptide transporter
263216_s_at	AT1G30720; AT1G30730	AT1G30720, FAD-binding domain-containing protein; AT1G30730, FAD-binding domain-containing protein
259753_at	AT1G71050	heavy-metal-associated domain-containing protein / copper chaperone (CCH)-related
252204_at	AT3G50340	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G67020.1); similar to unknown protein *Oryza sativa* (*japonica* cultivar-group) (GB:BAD09363.1)
255521_at	AT4G02280	SUS3; UDP-glycosyltransferase/ sucrose synthase/ transferase, transferring glycosyl groups
260568_at	AT2G43570	chitinase, putative
255064_at	AT4G08950	phosphate-responsive protein, putative (EXO)
266992_at	AT2G39200	MLO12 (MILDEW RESISTANCE LOCUS O 12); calmodulin binding
263830_at	AT2G40260	myb family transcription factor
260919_at	no_match	no_match
254243_at	AT4G23210	CRK13/HIG1; kinase
264319_at	AT1G04110	SDD1 (STOMATAL DENSITY AND DISTRIBUTION); subtilase
254232_at	AT4G23600	CORI3 (CORONATINE INDUCED 1, JASMONIC ACID RESPONSIVE 2); transaminase
265917_at	AT2G15080	disease resistance family protein
259173_at	AT3G03640	GLUC (Beta-glucosidase homolog); hydrolase, hydrolyzing O-glycosyl compounds
250062_at	AT5G17760	AAA-type ATPase family protein
253414_at	AT4G33050	EDA39 (embryo sac development arrest 39); calmodulin binding
246825_at	AT5G26260	meprin and TRAF homology domain-containing protein / MATH domain-containing protein
247095_at	AT5G66400	RAB18 (RESPONSIVE TO ABA 18)
247794_at	AT5G58670	ATPLC1 (PHOSPHOLIPASE C 1); phospholipase C

ATH1 Array identifier	Gene identifier	Gene description
246860_at	AT5G25840	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G79770.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO45837.1); contains InterPro domain Protein of unknown function DUF1677, plant (InterPro:IPR012876)
267389_at	AT2G44460	glycosyl hydrolase family 1 protein
254889_at	AT4G11650	ATOSM34 (OSMOTIN 34)
246831_at	AT5G26340	MSS1 (SUGAR TRANSPORT PROTEIN 13); carbohydrate transmembrane transporter/ hexose:hydrogen symporter/ high-affinity hydrogen:glucose symporter/ sugar:hydrogen ion symporter
264891_at	AT1G23200	pectinesterase family protein
250152_at	AT5G15120	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G39890.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO14912.1); contains InterPro domain Protein of unknown function DUF1637 (InterPro:IPR012864)
255059_at	AT4G09420	disease resistance protein (TIR-NBS class), putative
258321_at	AT3G22840	ELIP1 (EARLY LIGHT-INDUCABLE PROTEIN); chlorophyll binding
246228_at	AT4G36430	peroxidase, putative
248118_at	AT5G55050	GDSL-motif lipase/hydrolase family protein
245393_at	AT4G16260	glycosyl hydrolase family 17 protein
252117_at	AT3G51430	YLS2 (yellow-leaf-specific gene 2); strictosidine synthase
254952_at	AT4G10955; AT4G10960	AT4G10955, lipase class 3 family protein; AT4G10960, UGE5 (UDP-D-GLUCOSE/UDP-D-GALACTOSE 4-EPIMERASE 5); UDP-glucose 4-epimerase/ protein dimerization
266227_at	AT2G28870	unknown protein
252114_at	AT3G51450	strictosidine synthase family protein
251400_at	AT3G60420	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G60450.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO70569.1); contains InterPro domain Phosphoglycerate mutase (InterPro:IPR013078); contains InterPro domain PRIB5 (InterPro:IPR012398)
248686_at	AT5G48540	33 kDa secretory protein-related
255516_at	AT4G02270	pollen Ole e 1 allergen and extensin family protein
264400_at	AT1G61800	GPT2 (glucose-6-phosphate/phosphate translocator 2); antiporter/ glucose-6-phosphate transmembrane transporter
247957_at	AT5G57050	ABI2 (ABA INSENSITIVE 2); protein serine/threonine phosphatase
265530_at	AT2G06050	OPR3 (OPDA-REDUCTASE 3); 12-oxophytodienoate reductase
258277_at	AT3G26830	PAD3 (PHYTOALEXIN DEFICIENT 3); oxygen binding
257008_at	AT3G14210	ESM1 (EPITHIOSPECIFIER MODIFIER 1); carboxylesterase
259120_at	AT3G02240	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G02242.1)
260549_at	AT2G43535	trypsin inhibitor, putative
246884_at	AT5G26220	ChaC-like family protein

ATH1 Array identifier	Gene identifier	Gene description
257672_at	AT3G20300	extracellular ligand-gated ion channel
260046_at	AT1G73805	calmodulin binding
267279_at	AT2G19460	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G11970.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO65328.1)
246297_at	AT3G51760	unknown protein
262680_at	AT1G75880	family II extracellular lipase 1 (EXL1)
267425_at	AT2G34810	FAD-binding domain-containing protein
254333_at	AT4G22753	SMO1-3 (STEROL 4-ALPHA METHYL OXIDASE); catalytic
266319_s_at	AT3G10280; AT2G46720	AT3G10280, fatty acid elongase 3-ketoacyl-CoA synthase, putative; AT2G46720, HIC (HIGH CARBON DIOXIDE); acyltransferase
264886_at	AT1G61120	terpene synthase/cyclase family protein
254155_at	AT4G24480	serine/threonine protein kinase, putative
260656_at	AT1G19380	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G65650.1); similar to unknown *Vitis pseudoreticulata* (GB:ABC69762.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO61173.1); contains InterPro domain Protein of unknown function DUF1195 (InterPro:IPR010608)
245089_at	AT2G45290	transketolase, putative
247097_at	AT5G66460	(1-4)-beta-mannan endohydrolase, putative
265028_at	AT1G24530	transducin family protein / WD-40 repeat family protein
260633_at	AT1G62400	HT1 (HIGH LEAF TEMPERATURE 1); kinase/ protein serine/threonine/tyrosine kinase
251598_at	AT3G57600	AP2 domain-containing transcription factor, putative
251226_at	AT3G62680	PRP3 (PROLINE-RICH PROTEIN 3); structural constituent of cell wall
259507_at	AT1G43910	AAA-type ATPase family protein
258002_at	AT3G28930	AIG2 (AVRRPT2-INDUCED GENE 2)
264223_s_at	AT3G16030	CES101 (CALLUS EXPRESSION OF RBCS 101); carbohydrate binding / kinase
255230_at	AT4G05390	ATRFNR1 (ROOT FNR 1); oxidoreductase
249911_at	AT5G22740	ATCSLA02 (Cellulose synthase-like A2); transferase, transferring glycosyl groups
266168_at	AT2G38870	protease inhibitor, putative
264872_at	AT1G24260	SEP3 (SEPALLATA3); transcription factor
257939_at	AT3G19930	STP4 (SUGAR TRANSPORTER 4); carbohydrate transmembrane transporter/ sugar:hydrogen ion symporter
260408_at	AT1G69880	ATH8 (thioredoxin H-type 8); thiol-disulfide exchange intermediate
245760_s_at	AT1G66920; AT1G66910	AT1G66920, serine/threonine protein kinase, putative; AT1G66910, protein kinase, putative
255795_at	AT2G33380	RD20 (RESPONSIVE TO DESSICATION 20); calcium ion binding

ATH1 Array identifier	Gene identifier	Gene description
254697_at	AT4G17970	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G46600.1); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G46610.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO63429.1); contains InterPro domain Protein of unknown function UPF0005 (InterPro:IPR006214)
254163_s_at	AT4G24340; AT4G24350	AT4G24340, phosphorylase family protein; AT4G24350, phosphorylase family protein
246302_at	AT3G51860	CAX3 (cation exchanger 3); cation:cation antiporter
248763_at	AT5G47550	cysteine protease inhibitor, putative / cystatin, putative
250801_at	AT5G04960	pectinesterase family protein
254744_at	AT4G13345	MEE55 (maternal effect embryo arrest 55)
255807_at	AT4G10270	wound-responsive family protein
254819_at	AT4G12500	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
251673_at	AT3G57240	BG3 (BETA-1,3-GLUCANASE 3); hydrolase, hydrolyzing O-glycosyl compounds
247800_at	AT5G58570	unknown protein
246368_at	AT1G51890	leucine-rich repeat protein kinase, putative
249814_at	AT5G23840	MD-2-related lipid recognition domain-containing protein / ML domain-containing protein
261135_at	AT1G19610	LCR78/PDF1.4 (Low-molecular-weight cysteine-rich 78)
247266_at	AT5G64570	XYL4 (beta-xylosidase 4); hydrolase, hydrolyzing O-glycosyl compounds
267264_at	AT2G22970	SCPL11; serine carboxypeptidase
249767_at	AT5G24090	acidic endochitinase (CHIB1)
248327_at	AT5G52750	heavy-metal-associated domain-containing protein
249867_at	AT5G23020	IMS2/MAM-L/MAM3 (METHYLTHIOALKYMALATE SYNTHASE-LIKE); 2-isopropylmalate synthase/ methylthioalkylmalate synthase
247109_at	AT5G65870	ATPSK5 (PHYTOSULFOKINE 5 PRECURSOR); growth factor
264782_at	AT1G08810	MYB60 (myb domain protein 60); DNA binding / transcription factor
263876_at	AT2G21880	AtRABG2/AtRab7A (Arabidopsis Rab GTPase homolog G2); GTP binding
263544_at	AT2G21590	APL4 (large subunit of AGP 4); glucose-1-phosphate adenylyltransferase
252265_at	AT3G49620	DIN11 (DARK INDUCIBLE 11); oxidoreductase
257634_s_at	AT3G26170; AT3G26180	AT3G26170, CYP71B19 (cytochrome P450, family 71, subfamily B, polypeptide 19); oxygen binding; AT3G26180, CYP71B20 (cytochrome P450, family 71, subfamily B, polypeptide 20); oxygen binding
254494_at	AT4G20050	QRT3 (QUARTET 3)
256596_at	AT3G28540	AAA-type ATPase family protein
254396_at	AT4G21680	proton-dependent oligopeptide transport (POT) family protein

ATH1 Array identifier	Gene identifier	Gene description
262939_s_at	AT1G79530; AT1G16300	AT1G79530, GAPCP-1; glyceraldehyde-3-phosphate dehydrogenase; AT1G16300, GAPCP-2; glyceraldehyde-3-phosphate dehydrogenase
245946_at	AT5G19580	glyoxal oxidase-related
258063_at	AT3G14620	CYP72A8 (cytochrome P450, family 72, subfamily A, polypeptide 8); oxygen binding
254265_s_at	AT4G23140; AT4G23160	AT4G23140, CRK6 (CYSTEINE-RICH RLK 6); kinase; AT4G23160, protein kinase family protein
246991_at	AT5G67400	peroxidase 73 (PER73) (P73) (PRXR11)
254806_at	AT4G12430; AT4G12432	AT4G12430, trehalose-6-phosphate phosphatase, putative; AT4G12432, CPuORF26 (Conserved peptide upstream open reading frame 26)
261070_at	AT1G07390	protein binding
254231_at	AT4G23810	WRKY53 (WRKY DNA-binding protein 53); DNA binding / protein binding / transcription activator/ transcription factor
258791_at	AT3G04720	PR4 (PATHOGENESIS-RELATED 4)
258498_at	AT3G02480	ABA-responsive protein-related
248330_at	AT5G52810	ornithine cyclodeaminase/mu-crystallin family protein
249167_at	AT5G42860	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G45688.1); similar to H0814G11.12 *Oryza sativa* (*indica* cultivar-group) (GB:CAJ86345.1); similar to CAA30379.1 protein *Oryza sativa* (GB:CAB53482.1)
252888_at	AT4G39210	APL3 (large subunit of AGP 3)
251181_at	AT3G62820	invertase/pectin methylesterase inhibitor family protein
252511_at	AT3G46280	protein kinase-related
261339_at	AT1G35710	leucine-rich repeat transmembrane protein kinase, putative
248062_at	AT5G55450	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
267411_at	AT2G34930	disease resistance family protein
257396_at	AT2G20875	EPF1 (EPIDERMAL PATTERNING FACTOR 1)
260101_at	AT1G73260	trypsin and protease inhibitor family protein / Kunitz family protein
254122_at	AT4G24510	CER2 (ECERIFERUM 2); transferase
247717_at	AT5G59320	LTP3 (LIPID TRANSFER PROTEIN 3); lipid binding
262374_s_at	AT1G72910; AT1G72930	AT1G72910, disease resistance protein (TIR-NBS class), putative; AT1G72930, TIR (TOLL/INTERLEUKIN-1 RECEPTOR-LIKE); transmembrane receptor
263121_at	AT1G78530	protein kinase family protein
246620_at	AT5G36220	CYP81D1 (CYTOCHROME P450 91A1); oxygen binding
254361_at	AT4G22212	Encodes a defensin-like (DEFL) family protein.
250669_at	AT5G06870	PGIP2 (POLYGALACTURONASE INHIBITING PROTEIN 2); protein binding
260116_at	AT1G33960	AIG1 (AVRRPT2-INDUCED GENE 1); GTP binding
252076_at	AT3G51660	macrophage migration inhibitory factor family protein / MIF family protein

ATH1 Array identifier	Gene identifier	Gene description
267622_at	AT2G39690	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G12540.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO69194.1); contains InterPro domain Protein of unknown function DUF547 (InterPro:IPR006869)
249021_at	AT5G44820	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G19970.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO46707.1); contains domain PTHR10483:SF6 (PTHR10483:SF6); contains domain PTHR10483 (PTHR10483)
266993_at	AT2G39210	nodulin family protein
262531_at	AT1G17230	protein binding / protein kinase
254042_at	AT4G25810	XTR6 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 6); hydrolase, acting on glycosyl bonds
266782_at	AT2G29120	ATGLR2.7 (Arabidopsis thaliana glutamate receptor 2.7)
254805_at	AT4G12480	pEARLI 1; lipid binding
248932_at	AT5G46050	ATPTR3/PTR3 (PEPTIDE TRANSPORTER PROTEIN 3); transporter
254534_at	AT4G19680	IRT2 (iron-responsive transporter 2); iron ion transmembrane transporter/ zinc ion transmembrane transporter
257056_at	AT3G15350	glycosyltransferase family 14 protein / core-2/I-branching enzyme family protein
250043_at	AT5G18430	GDSL-motif lipase/hydrolase family protein
264007_at	AT2G21140	ATPRP2 (PROLINE-RICH PROTEIN 2)
253485_at	AT4G31800	WRKY18 (WRKY DNA-binding protein 18); transcription factor
245329_at	AT4G14365	zinc finger (C3HC4-type RING finger) family protein / ankyrin repeat family protein
246366_at	AT1G51850	leucine-rich repeat protein kinase, putative
249474_s_at	AT5G39160; AT5G39190; AT5G39130	AT5G39160, germin-like protein (GLP2a) (GLP5a); AT5G39190, GLP2A (GERMIN-LIKE PROTEIN 2A); manganese ion binding / metal ion binding / nutrient reservoir;[AT5G39130, germin-like protein, putative]
265012_at	AT1G24470	short-chain dehydrogenase/reductase (SDR) family protein
245771_at	no_match	no_match
250778_at	AT5G05500	pollen Ole e 1 allergen and extensin family protein
264525_at	AT1G10060	ATBCAT-1; branched-chain-amino-acid transaminase/ catalytic
267518_at	AT2G30500	kinase interacting family protein
253264_at	AT4G33950	OST1/P44/SNRK2-6/SRK2E (OPEN STOMATA 1, SNF1-RELATED PROTEIN KINASE 2.6); calcium-dependent protein serine/threonine kinase/ kinase/ protein kinase
248684_at	AT5G48485	DIR1 (DEFECTIVE IN INDUCED RESISTANCE 1); lipid binding
266877_at	AT2G44570	ATGH9B12 (ARABIDOPSIS THALIANA GLYCOSYL HYDROLASE 9B12); hydrolase, hydrolyzing O-glycosyl compounds
250832_at	AT5G04950	nicotianamine synthase, putative
253934_at	AT4G26830	hydrolase, hydrolyzing O-glycosyl compounds

ATH1 Array identifier	Gene identifier	Gene description
249052_at	AT5G44420	PDF1.2 (Low-molecular-weight cysteine-rich 77)
258941_at	AT3G09940	ATMDAR3/MDHAR (MONODEHYDROASCORBATE REDUCTASE); monodehydroascorbate reductase (NADH)
258252_at	AT3G15720	glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein
254197_at	AT4G24040	ATTRE1/TRE1 (TREHALASE 1); alpha,alpha-trehalase/ trehalase
248551_at	AT5G50200	WR3 (WOUND-RESPONSIVE 3); nitrate transmembrane transporter
254521_at	AT5G44820	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G19970.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO46707.1); contains domain PTHR10483:SF6 (PTHR10483:SF6); contains domain PTHR10483 (PTHR10483)
247573_at	AT5G61160	AACT1 (ANTHOCYANIN 5-AROMATIC ACYLTRANSFERASE 1); transferase
250690_at	AT5G06530	ABC transporter family protein
245148_at	AT2G45220	pectinesterase family protein
266364_at	AT2G41230	unknown protein
246652_at	AT5G35190	proline-rich extensin-like family protein
252989_at	AT4G38420	SKS9 (SKU5 Similar 9); copper ion binding / oxidoreductase
261037_at	AT1G17420	LOX3 (Lipoxygenase 3); iron ion binding / lipoxygenase/ metal ion binding / oxidoreductase, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen
246550_at	AT5G14920	gibberellin-regulated family protein
266746_s_at	AT2G02930; AT4G02520	AT2G02930, ATGSTF3 (GLUTATHIONE S-TRANSFERASE 16); glutathione transferase; AT4G02520, ATGSTF2 (Arabidopsis thaliana Glutathione S-transferase (class phi) 2); glutathione transferase
245262_at	AT4G16563	aspartyl protease family protein
266866_at	AT2G29940	ATPDR3/PDR3 (PLEIOTROPIC DRUG RESISTANCE 3); ATPase, coupled to transmembrane movement of substances
265823_at	AT2G35760	integral membrane family protein
264840_at	AT1G03440	leucine-rich repeat family protein
252417_at	AT3G47480	calcium-binding EF hand family protein
254093_at	AT4G25110	ATMC2 (METACASPASE 2); caspase
254385_s_at	AT4G21830; AT4G21840	AT4G21830, methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein; AT4G21840, methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein
253960_at	no_match	no_match
263718_at	AT2G13570	CCAAT-box binding transcription factor, putative
254202_at	AT4G24140	hydrolase, alpha/beta fold family protein
254387_at	AT4G21850	methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein

ATH1 Array identifier	Gene identifier	Gene description
257623_at	AT3G26210	CYP71B23 (cytochrome P450, family 71, subfamily B, polypeptide 23); oxygen binding
261653_at	AT1G01900	ATSBT1.1; subtilase

Table 2-S2. Genes showing high expression in *scrm-D*.
List of 407 genes representing gene cluster V from Figure 2-2. ATH1 array, respective gene locus identifiers, and gene descriptions are indicated. Coexpression was determined according to Table 2-S1.

Table 2-S3.

ATH1 Array identifier	Gene identifier	Gene description
248756_at	AT5G47560	ATSDAT/ATTDT (TONOPLAST DICARBOXYLATE TRANSPORTER); malate transmembrane transporter/ sodium:dicarboxylate symporter
253217_at	AT4G34970	actin binding
248064_at	no_match	no_match
257315_at	AT3G30775	ERD5 (EARLY RESPONSIVE TO DEHYDRATION 5); proline dehydrogenase
261567_at	AT1G33055	unknown protein
252367_at	AT3G48360	BT2 (BTB AND TAZ DOMAIN PROTEIN 2); protein binding / transcription factor/ transcription regulator
262381_at	AT1G72900	disease resistance protein (TIR-NBS class), putative
258815_at	AT3G04000	short-chain dehydrogenase/reductase (SDR) family protein
258049_at	AT3G16220	similar to RNA binding / catalytic *Arabidopsis thaliana* (TAIR:AT3G16230.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO45862.1); contains InterPro domain Predicted eukaryotic LigT; (InterPro:IPR009210)
254020_at	AT4G25700	BETA-OHASE 1 (BETA-HYDROXYLASE 1); beta-carotene hydroxylase
260427_at	AT1G72430	auxin-responsive protein-related
251356_at	AT3G61060	ATPP2-A13
262885_at	AT1G64740	TUA1 (ALPHA-1 TUBULIN)
246519_at	AT5G15780	pollen Ole e 1 allergen and extensin family protein
250182_at	AT5G14470	GHMP kinase-related
250972_at	AT5G02840	LCL1 (LHY/CCA1-LIKE 1); DNA binding / transcription factor
264514_at	AT1G09500	cinnamyl-alcohol dehydrogenase family / CAD family
250199_at	AT5G14180	MPL1 (MYZUS PERSICAE-INDUCED LIPASE 1); catalytic
245711_at	AT5G04340	C2H2 (ZINC FINGER OF ARABIDOPSIS THALIANA 6); nucleic acid binding / transcription factor/ zinc ion binding
255764_at	AT1G16720	HCF173 (HIGH CHLOROPHYLL FLUORESCENCE PHENOTYPE 173); binding / catalytic/ oxidoreductase/ transcription repressor
264720_at	AT1G70080	terpene synthase/cyclase family protein
259432_at	AT1G01520	myb family transcription factor
265481_at	AT2G15960	unknown protein
263402_at	AT2G04050	MATE efflux family protein
266658_at	AT2G25735	unknown protein
256970_at	AT3G21090	ABC transporter family protein
247543_at	AT5G61600	ethylene-responsive element-binding family protein
262038_at	AT1G35580	CINV1 (CYTOSOLIC INVERTASE 1); beta-fructofuranosidase
259751_at	AT1G71030	ATMYBL2 (Arabidopsis myb-like 2); DNA binding / transcription factor
245506_at	AT4G15700	glutaredoxin family protein
263593_at	AT2G01860	EMB975 (EMBRYO DEFECTIVE 975)

ATH1 Array identifier	Gene identifier	Gene description
264923_s_at	AT1G65970; AT1G60740	AT1G65970, TPX2 (THIOREDOXIN-DEPENDENT PEROXIDASE 2); antioxidant; AT1G60740, peroxiredoxin type 2, putative
259982_at	AT1G76410	ATL8; protein binding / zinc ion binding
246028_at	AT5G21170	5'-AMP-activated protein kinase beta-2 subunit, putative
248606_at	AT5G49450; AT5G49448	AT5G49450, ATBZIP1 (ARABIDOPSIS THALIANA BASIC LEUCINE-ZIPPER 1); DNA binding / protein heterodimerization/ transcription factor; AT5G49448, CPuORF4 (Conserved peptide upstream open reading frame 4)
264124_at	AT1G79360	ATOCT2 (ARABIDOPSIS THALIANA ORGANIC CATION/CARNITINE TRANSPORTER2); carbohydrate transmembrane transporter/ sugar:hydrogen ion symporter
256751_at	AT3G27170	CLC-B (chloride channel protein B); anion channel/ voltage-gated chloride channel
254787_at	AT4G12690	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G04220.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO41124.1); contains InterPro domain Protein of unknown function DUF868, plant (InterPro:IPR008586)
245264_at	AT4G17245	zinc finger (C3HC4-type RING finger) family protein
263847_at	AT2G36970	UDP-glucoronosyl/UDP-glucosyl transferase family protein
247979_at	AT5G56750	Ndr family protein
258939_at	AT3G10020	similar to Os12g0147200 *Oryza sativa* (*japonica* cultivar-group) (GB:NP_001066153.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO15981.1); similar to Os11g0149200 *Oryza sativa* (*japonica* cultivar-group) (GB:NP_001065754.1)
246510_at	AT5G15410	DND1 (DEFENSE NO DEATH 1); calcium channel/ calmodulin binding / cation channel/ cyclic nucleotide binding / inward rectifier potassium channel
262705_at	AT1G16260	protein kinase family protein
254715_at	AT4G13550	lipase class 3 family protein
249101_at	AT5G43580	serine-type endopeptidase inhibitor
261618_at	AT1G33110	MATE efflux family protein
250666_at	AT5G07100	WRKY26 (WRKY DNA-binding protein 26); transcription factor
254016_at	AT4G26150	CGA1 (CYTOKININ-RESPONSIVE GATA FACTOR 1); transcription factor
260527_at	AT2G47270	transcription factor/ transcription regulator
260856_at	AT1G21910	AP2 domain-containing transcription factor family protein
262049_at	AT1G80180	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G15400.2); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G15400.3); similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G15400.1); similar to hypothetical protein MtrDRAFT_AC148340g12v2 *Medicago truncatula* (GB:ABD28396.1)
251438_s_at	AT3G59930; AT5G33355	AT3G59930, Encodes a defensin-like (DEFL) family protein.; AT5G33355, Encodes a defensin-like (DEFL) family protein.
263118_at	AT1G03090	MCCA (3-methylcrotonyl-CoA carboxylase 1)

ATH1 Array identifier	Gene identifier	Gene description
245504_at	AT4G15660	glutaredoxin family protein
265999_at	AT2G24100	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G30780.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO66612.1)
265817_at	AT2G18050	HIS1-3 (HISTONE H1-3); DNA binding
255742_at	AT1G25560	AP2 domain-containing transcription factor, putative
247954_at	AT5G56870	BGAL4 (beta-galactosidase 4); beta-galactosidase
248282_at	AT5G52900	similar to unnamed protein product *Vitis vinifera* (GB:CAO49548.1)
260567_at	AT2G43820	GT/UGT74F2 (UDP-GLUCOSYLTRANSFERASE 74F2); UDP-glucosyltransferase/ UDP-glycosyltransferase/ transferase, transferring glycosyl groups / transferase, transferring hexosyl groups
248207_at	AT5G53970	aminotransferase, putative
262882_at	AT1G64900	CYP89A2 (CYTOCHROME P450 89A2); oxygen binding
246962_s_at	AT5G24800	ATBZIP9/BZO2H2 (BASIC LEUCINE ZIPPER O2 HOMOLOG 2); DNA binding / protein heterodimerization/ transcription factor
253496_at	AT4G31870	ATGPX7 (GLUTATHIONE PEROXIDASE 7); glutathione peroxidase
253874_at	AT4G27450	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G15450.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO39242.1); contains domain G3DSA:3.60.20.10 (G3DSA:3.60.20.10); contains domain SSF56235 (SSF56235)
246114_at	AT5G20250	DIN10 (DARK INDUCIBLE 10); hydrolase, hydrolyzing O-glycosyl compounds
249125_at	AT5G43450	2-oxoglutarate-dependent dioxygenase, putative
263122_at	AT1G78510	SPS1 (SOLANESYL DIPHOSPHATE SYNTHASE 1); dimethylallyltranstransferase/ trans-octaprenyltranstransferase
250293_s_at	AT5G13370; AT5G13360	AT5G13370, auxin-responsive GH3 family protein; AT5G13360, auxin-responsive GH3 family protein
248311_at	AT5G52570	BETA-OHASE 2 (BETA-CAROTENE HYDROXYLASE 2); beta-carotene hydroxylase
247925_at	AT5G57560	TCH4 (TOUCH 4); hydrolase, acting on glycosyl bonds / xyloglucan:xyloglucosyl transferase
245306_at	AT4G14690	ELIP2 (EARLY LIGHT-INDUCIBLE PROTEIN 2); chlorophyll binding
250217_at	AT5G14120	nodulin family protein
254384_at	AT4G21870	26.5 kDa class P-related heat shock protein (HSP26.5-P)
265387_at	AT2G20670	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G32480.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO69754.1); contains InterPro domain Protein of unknown function DUF506, plant (InterPro:IPR006502)
247488_at	AT5G61820	similar to MtN19-like protein *Pisum sativum* (GB:AAU14999.2); contains InterPro domain Stress up-regulated Nod 19 (InterPro:IPR011692)
248868_at	AT5G46780	VQ motif-containing protein

ATH1 Array identifier	Gene identifier	Gene description
266983_at	AT2G39400	hydrolase, alpha/beta fold family protein
262456_at	AT1G11260	STP1 (SUGAR TRANSPORTER 1); carbohydrate transmembrane transporter/ sugar:hydrogen ion symporter
255617_at	AT4G01330	protein kinase family protein
258395_at	AT3G15500	ATNAC3 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 55); transcription factor
258244_at	AT3G27770	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G62960.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO14740.1)
255160_at	AT4G07820	pathogenesis-related protein, putative
251196_at	AT3G62950	glutaredoxin family protein
260536_at	AT2G43400	ETFQO (ELECTRON-TRANSFER FLAVOPROTEIN:UBIQUINONE OXIDOREDUCTASE); catalytic/ electron acceptor
261395_at	AT1G79700	ovule development protein, putative
253841_at	AT4G27830	glycosyl hydrolase family 1 protein
245317_at	AT4G15610	integral membrane family protein
253281_at	AT4G34138	UGT73B1 (UDP-GLUCOSYL TRANSFERASE 73B1); UDP-glycosyltransferase/ abscisic acid glucosyltransferase
266290_at	AT2G29490	ATGSTU1 (GLUTATHIONE S-TRANSFERASE 19); glutathione transferase
264788_at	AT2G17880	DNAJ heat shock protein, putative
263184_at	AT1G05560	UGT1 (UDP-glucosyl transferase 75B1); UDP-glycosyltransferase/ transferase, transferring glycosyl groups
256965_at	AT3G13450	DIN4 (DARK INDUCIBLE 4); 3-methyl-2-oxobutanoate dehydrogenase (2-methylpropanoyl-transferring)
254432_at	AT4G20830	FAD-binding domain-containing protein
261934_at	AT1G22400	ATUGT85A1/UGT85A1 (UDP-GLUCOSYL TRANSFERASE 85A1); UDP-glycosyltransferase/ glucuronosyltransferase/ transferase, transferring glycosyl groups / transferase, transferring hexosyl groups
265648_at	AT2G27500	glycosyl hydrolase family 17 protein
266892_at	AT2G26080	ATGLDP2 (ARABIDOPSIS THALIANA GLYCINE DECARBOXYLASE P-PROTEIN 2); glycine dehydrogenase (decarboxylating)
259977_at	AT1G76590	zinc-binding family protein
253161_at	AT4G35770	SEN1 (DARK INDUCIBLE 1)
258456_at	AT3G22420	WNK2 (WITH NO K 2); kinase
257670_at	AT3G20340	Expression of the gene is downregulated in the presence of paraquat, an inducer of photoxidative stress.
249942_at	AT5G22300	NIT4 (NITRILASE 4)
267357_at	AT2G40000	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G55840.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO41329.1); contains InterPro domain Hs1pro-1, C-terminal (InterPro:IPR009743); contains InterPro domain Hs1pro-1, N-terminal (InterPro:IPR009869)

ATH1 Array identifier	Gene identifier	Gene description
258935_at	AT3G10120	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G03890.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN62766.1)
259058_at	AT3G03470	CYP89A9 (cytochrome P450, family 87, subfamily A, polypeptide 9); oxygen binding
262844_at	AT1G14890	pectinesterase inhibitor
256324_at	AT1G66760	MATE efflux family protein
266072_at	AT2G18700	ATTPS11 (Arabidopsis thaliana trehalose phosphatase/synthase 11); transferase, transferring glycosyl groups
262346_at	AT1G63980	D111/G-patch domain-containing protein
245668_at	AT1G28330	DRM1 (DORMANCY-ASSOCIATED PROTEIN 1)
258472_at	AT3G06080	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G19160.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO71110.1); contains InterPro domain Protein of unknown function DUF231, plant (InterPro:IPR004253)
245936_at	AT5G19850	hydrolase, alpha/beta fold family protein
262935_at	AT1G79410	ATOCT5 (ARABIDOPSIS THALIANA ORGANIC CATION/CARNITINE TRANSPORTER5); carbohydrate transmembrane transporter/ sugar:hydrogen ion symporter
264580_at	AT1G05340	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G32210.1)
265203_at	AT2G36630	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G25737.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO16008.1); contains InterPro domain Protein of unknown function DUF81; (InterPro:IPR002781)
253125_at	AT4G36040	DNAJ heat shock N-terminal domain-containing protein (J11)
262518_at	AT1G17170	ATGSTU24 (ARABIDOPSIS THALIANA GLUTATHIONE S-TRANSFERASE (CLASS TAU) 24); glutathione transferase
253343_at	AT4G33540	metallo-beta-lactamase family protein
254691_at	AT4G17840	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G35260.1); similar to hypothetical protein 40.t00061 *Brassica oleracea* (GB:ABD65174.1)
249769_at	AT5G24120	SIGE (RNA polymerase sigma subunit E); DNA binding / DNA-directed RNA polymerase/ sigma factor/ transcription factor
247994_at	AT5G56140	KH domain-containing protein
261068_at	AT1G07450	tropinone reductase, putative / tropine dehydrogenase, putative
246464_at	AT5G16980	NADP-dependent oxidoreductase, putative
253373_at	AT4G33150	LKR (SACCHAROPINE DEHYDROGENASE)
246260_at	AT1G31820	amino acid permease family protein
249174_at	AT5G42900	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G33980.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN75496.1)
264289_at	AT1G61890	MATE efflux family protein
252882_at	AT4G39675	unknown protein

ATH1 Array identifier	Gene identifier	Gene description
248193_at	AT5G54080	HGO (HOMOGENTISATE 1,2-DIOXYGENASE); homogentisate 1,2-dioxygenase
250777_at	AT5G05440	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT2G38310.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO48777.1); contains InterPro domain Bet v I allergen; (InterPro:IPR000916); contains InterPro domain Streptomyces cyclase/dehydrase (InterPro:IPR005031)
253872_at	AT4G27410	RD26 (RESPONSIVE TO DESSICATION 26); transcription factor
253046_at	AT4G37370	CYP81D8 (cytochrome P450, family 81, subfamily D, polypeptide 8); oxygen binding
260900_s_at	AT1G21400; AT5G34780	AT1G21400, 2-oxoisovalerate dehydrogenase, putative / 3-methyl-2-oxobutanoate dehydrogenase, putative / branched-chain alpha-keto acid dehydrogenase E1 alpha subunit, putative; AT5G34780, dehydrogenase E1 component family protein
267230_at	AT2G44080	ARL (ARGOS-LIKE)
254638_at	AT4G18740	transcription termination factor
261821_at	AT1G11530	ATCXXS1 (C-TERMINAL CYSTEINE RESIDUE IS CHANGED TO A SERINE 1); thiol-disulfide exchange intermediate
256589_at	AT3G28740	cytochrome P450 family protein
260914_at	AT1G02640	BXL2 (BETA-XYLOSIDASE 2); hydrolase, hydrolyzing O-glycosyl compounds
249862_at	AT5G22920	zinc finger (C3HC4-type RING finger) family protein
267337_at	AT2G39980	transferase family protein
248050_at	AT5G56100	glycine-rich protein / oleosin
264238_at	AT1G54740	similar to structural constituent of ribosome *Arabidopsis thaliana* (TAIR:AT1G22110.1); similar to unknown *Populus trichocarpa* × *Populus deltoides* (GB:ABK96485.1)
261674_at	AT1G18270	ketose-bisphosphate aldolase class-II family protein
245136_at	AT2G45210	auxin-responsive protein-related
255381_at	AT4G03510	RMA1 (Ring finger protein with Membrane Anchor 1); protein binding / ubiquitin-protein ligase/ zinc ion binding
253268_s_at	AT4G34131; AT4G34135	AT4G34131, UGT73B3 (UDP-GLUCOSYL TRANSFERASE 73B3); UDP-glycosyltransferase/ abscisic acid glucosyltransferase/ transferase, transferring hexosyl groups; AT4G34135, UGT73B2; UDP-glucosyltransferase/ UDP-glycosyltransferase/ flavonol 3-O-glucosyltransferase
265511_at	AT2G05540	glycine-rich protein
262505_at	AT1G21680	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G21670.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN73514.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO61906.1); similar to hypothetical protein OsJ_012725 *Oryza sativa* (*japonica* cultivar-group) (GB:EAZ29242.1); contains InterPro domain WD40-like Beta Propeller (InterPro:IPR011659); contains InterPro domain Six-bladed beta-propeller, TolB-like (InterPro:IPR011042)

ATH1 Array identifier	Gene identifier	Gene description
256894_at	AT3G21870	CYCP2;1 (cyclin p2;1); cyclin-dependent protein kinase
267523_at	AT2G30600	BTB/POZ domain-containing protein
261353_at	AT1G79600	ABC1 family protein
257830_at	AT3G26690	ATNUDT13 (ARABIDOPSIS THALIANA NUDIX HYDROLASE HOMOLOG 13); hydrolase
267254_at	AT2G23030	SNRK2-9/SNRK2.9 (SNF1-RELATED PROTEIN KINASE 2.9); kinase
265111_at	AT1G62510	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
260915_at	AT1G02660	lipase class 3 family protein
265494_at	AT2G15680	calmodulin-related protein, putative
252539_at	AT3G45730	unknown protein
245757_at	AT1G35140	PHI-1 (PHOSPHATE-INDUCED 1)
258227_at	AT3G15620	UVR3 (UV REPAIR DEFECTIVE 4)
246411_at	AT1G57770	amine oxidase family
263296_at	AT2G38800	calmodulin-binding protein-related
252057_at	AT3G52480	unknown protein
252081_at	AT3G51910	AT-HSFA7A (Arabidopsis thaliana heat shock transcription factor A7A); DNA binding / transcription factor
246523_at	AT5G15850	COL1 (CONSTANS-LIKE 1); transcription factor/ zinc ion binding
250868_at	AT5G03860	MLS (MALATE SYNTHASE); malate synthase
262118_at	AT1G02850	glycosyl hydrolase family 1 protein
245352_at	AT4G15490	UGT84A3; UDP-glycosyltransferase/ sinapate 1-glucosyltransferase/ transferase, transferring glycosyl groups
267477_at	AT2G02710	PAC motif-containing protein
263403_at	AT2G04040	ATDTX1; antiporter/ multidrug efflux pump/ multidrug transporter/ transporter
255543_at	AT4G01870	tolB protein-related
254496_at	AT4G20070	ATAAH (ARABIDOPSIS THALIANA ALLANTOATE AMIDOHYDROLASE); allantoate deiminase/ metallopeptidase
263019_at	AT1G23870	ATTPS9 (Arabidopsis thaliana trehalose-phosphatase/synthase 9); transferase, transferring glycosyl groups / trehalose-phosphatase
259511_at	AT1G12520	CCS1 (copper chaperone for superoxide dismutase 1); superoxide dismutase copper chaperone
246584_at	AT5G14730	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G01513.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO64332.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN65788.1)
250781_at	AT5G05410	DREB2A (DRE-BINDING PROTEIN 2A); DNA binding / transcription activator/ transcription factor
245193_at	AT1G67810	Fe-S metabolism associated domain-containing protein
266984_at	AT2G39570	ACT domain-containing protein

ATH1 Array identifier	Gene identifier	Gene description
251869_at	AT3G54500	similar to dentin sialophosphoprotein-related *Arabidopsis thaliana* (TAIR:AT5G64170.1); similar to dentin sialophosphoprotein-related *Arabidopsis thaliana* (TAIR:AT5G64170.2); similar to conserved hypothetical protein *Medicago truncatula* (GB:ABD28297.1)
251503_at	AT3G59140	ATMRP14 (Arabidopsis thaliana multidrug resistance-associated protein 14)
259226_at	AT3G07700	ABC1 family protein
261892_at	AT1G80840	WRKY40 (WRKY DNA-binding protein 40); transcription factor
260070_at	AT1G73830	BEE3 (BR ENHANCED EXPRESSION 3); DNA binding / transcription factor
254175_at	AT4G24050	short-chain dehydrogenase/reductase (SDR) family protein
247222_at	AT5G64840	ATGCN5 (Arabidopsis thaliana general control non-repressible 5)
257827_at	AT3G26630	pentatricopeptide (PPR) repeat-containing protein
247374_at	AT5G63190	MA3 domain-containing protein
248392_at	AT5G52050	MATE efflux protein-related
262947_at	AT1G75750	GASA1 (GAST1 PROTEIN HOMOLOG 1)
247704_at	AT5G59510	DVL18/RTFL5 (ROTUNDIFOLIA LIKE 5)
245448_at	AT4G16860	RPP4 (RECOGNITION OF PERONOSPORA PARASITICA 4)
264700_at	AT1G70100	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G24160.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO40971.1)
258383_at	AT3G15440	similar to zinc finger (C3HC4-type RING finger) family protein *Arabidopsis thaliana* (TAIR:AT3G15740.1)
254001_at	AT4G26260	MIOX4 (MYO-INOSITOL OXYGENASE 4)
247696_at	AT5G59780	MYB59 (myb domain protein 59); DNA binding / transcription factor
259681_at	AT1G77760	NIA1 (NITRATE REDUCTASE 1)
259129_at	AT3G02150	PTF1 (PLASTID TRANSCRIPTION FACTOR 1); transcription factor
258100_at	AT3G23550	MATE efflux family protein
246595_at	AT5G14780	FDH (FORMATE DEHYDROGENASE); NAD binding / binding / catalytic/ cofactor binding / oxidoreductase, acting on the CH-OH group of donors, NAD or NADP as acceptor
258507_at	AT3G06500	beta-fructofuranosidase, putative / invertase, putative / saccharase, putative / beta-fructosidase, putative
267266_at	AT2G23150	NRAMP3 (NRAMP metal ion transporter 3); manganese ion transmembrane transporter/ metal ion transmembrane transporter
252570_at	AT3G45300	IVD (ISOVALERYL-COA-DEHYDROGENASE)
245277_at	AT4G15550	IAGLU (INDOLE-3-ACETATE BETA-D-GLUCOSYLTRANSFERASE); UDP-glycosyltransferase/ transferase, transferring glycosyl groups
263067_at	AT2G17550	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G02390.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO14481.1)

ATH1 Array identifier	Gene identifier	Gene description
266269_at	AT2G29480	ATGSTU2 (GLUTATHIONE S-TRANSFERASE 20); glutathione transferase
261957_at	AT1G64660	ATMGL; catalytic/ methionine gamma-lyase
267178_at	AT2G37750	unknown protein
251428_at	AT3G60140	DIN2 (DARK INDUCIBLE 2); hydrolase, hydrolyzing O-glycosyl compounds
254057_at	AT4G25170	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G61490.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO60860.1); contains InterPro domain Uncharacterised conserved protein UCP012943 (InterPro:IPR016606)
249823_s_at	AT5G23350; AT5G23360	AT5G23350, GRAM domain-containing protein / ABA-responsive protein-related; AT5G23360, GRAM domain-containing protein / ABA-responsive protein-related
264524_at	AT1G10070	ATBCAT-2; branched-chain-amino-acid transaminase/ catalytic
250287_at	AT5G13330	RAP2.6L (related to AP2 6L); DNA binding / transcription factor
245528_at	AT4G15530	PPDK (PYRUVATE ORTHOPHOSPHATE DIKINASE); kinase/ pyruvate, phosphate dikinase
256725_at	AT2G34070	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G29050.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO48076.1); contains InterPro domain Protein of unknown function DUF231, plant (InterPro:IPR004253)
245392_at	AT4G15680	glutaredoxin family protein
264752_at	AT1G23010	LPR1 (LOW PHOSPHATE ROOT1); copper ion binding / oxidoreductase, acting on diphenols and related substances as donors, oxygen as acceptor
248622_at	AT5G49360	BXL1 (BETA-XYLOSIDASE 1); hydrolase, hydrolyzing O-glycosyl compounds
253048_at	AT4G37560	formamidase, putative / formamide amidohydrolase, putative
267524_at	AT2G30600	BTB/POZ domain-containing protein
258434_at	AT3G16770	ATEBP/ERF72/RAP2.3 (RELATED TO AP2 3); DNA binding / protein binding / transcription activator/ transcription factor
259015_at	AT3G07350	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G25240.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO38687.1); contains InterPro domain Protein of unknown function DUF506, plant (InterPro:IPR006502)
251235_at	AT3G62860	esterase/lipase/thioesterase family protein
260741_at	AT1G15040	glutamine amidotransferase-related
258033_at	AT3G21250	ATMRP6 (Arabidopsis thaliana multidrug resistance-associated protein 6)
255793_at	AT2G33250	similar to unnamed protein product *Vitis vinifera* (GB:CAO71510.1)

ATH1 Array identifier	Gene identifier	Gene description
264704_at	AT1G70090	GATL9/LGT8 (Galacturonosyltransferase-like 9); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups / transferase, transferring hexosyl groups
246099_at	AT5G20230	ATBCB (ARABIDOPSIS BLUE-COPPER-BINDING PROTEIN); copper ion binding
262916_at	AT1G59700	ATGSTU16 (Arabidopsis thaliana Glutathione S-transferase (class tau) 16); glutathione transferase
262626_at	AT1G06430	FTSH8 (FtsH protease 8); ATP-dependent peptidase/ ATPase/ metallopeptidase/ zinc ion binding
267028_at	AT2G38470	WRKY33 (WRKY DNA-binding protein 33); transcription factor
264777_at	AT1G08630	THA1 (THREONINE ALDOLASE 1); aldehyde-lyase/ threonine aldolase
252250_at	AT3G49790	ATP binding
258350_at	AT3G17510	CIPK1 (CBL-INTERACTING PROTEIN KINASE 1); kinase
257791_at	AT3G27110	peptidase M48 family protein
266313_at	AT2G26980	CIPK3 (CBL-INTERACTING PROTEIN KINASE 3); kinase/ protein kinase
263210_at	AT1G10585	transcription factor
255285_at	AT4G04630	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G21970.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO17943.1); contains InterPro domain Protein of unknown function DUF584 (InterPro:IPR007608)
251929_at	AT3G53920	SIGC (RNA polymerase sigma subunit C); DNA binding / DNA-directed RNA polymerase/ transcription factor
250099_at	AT5G17300	myb family transcription factor
261193_at	AT1G32920	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G32928.1)
255895_at	AT1G18020; AT1G17990	AT1G18020, 12-oxophytodienoate reductase, putative; AT1G17990, 12-oxophytodienoate reductase, putative
260637_at	AT1G62380	ACO2 (ACC OXIDASE 2)
247951_at	AT5G57240	oxysterol-binding family protein
251248_at	AT3G62150	PGP21 (P-GLYCOPROTEIN 21); ATPase, coupled to transmembrane movement of substances
259573_at	AT1G20390	transposable element gene
264042_at	AT2G03760	ST (steroid sulfotransferase); sulfotransferase
258275_at	AT3G15760	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G52565.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO62742.1)
264590_at	AT2G17710	similar to unnamed protein product *Vitis vinifera* (GB:CAO42932.1)
245925_at	AT5G28770	BZO2H3 (ARABIDOPSIS THALIANA BASIC LEUCINE ZIPPER 63); DNA binding / protein heterodimerization/ transcription factor
245349_at	AT4G16690	esterase/lipase/thioesterase family protein
247655_at	AT5G59820	RHL41 (RESPONSIVE TO HIGH LIGHT 41); nucleic acid binding / transcription factor/ zinc ion binding

ATH1 Array identifier	Gene identifier	Gene description
261211_at	AT1G12780	UGE1 (UDP-D-GLUCOSE/UDP-D-GALACTOSE 4-EPIMERASE 1); UDP-glucose 4-epimerase/ protein dimerization
250054_at	AT5G17860	CAX7 (CALCIUM EXCHANGER 7); calcium:sodium antiporter/ cation:cation antiporter
246034_at	AT5G08350	GRAM domain-containing protein / ABA-responsive protein-related
266150_s_at	AT2G12290; AT4G19700	AT2G12290, similar to protein binding / zinc ion binding *Arabidopsis thaliana* (TAIR:AT4G19700.1); similar to unknown *Populus trichocarpa* (GB:ABK94891.1); AT4G19700, protein binding / zinc ion binding
266140_at	AT2G28120	nodulin family protein
248040_at	AT5G55970	zinc finger (C3HC4-type RING finger) family protein
262488_at	AT1G21830	similar to unnamed protein product *Vitis vinifera* (GB:CAO61865.1)
257947_at	AT3G21720	isocitrate lyase, putative
254410_at	AT4G21410	protein kinase family protein
249029_at	AT5G44870	disease resistance protein (TIR-NBS-LRR class), putative
250017_at	AT5G18140	DNAJ heat shock N-terminal domain-containing protein
264527_at	AT1G30760	FAD-binding domain-containing protein
254931_at	AT4G11460	protein kinase family protein
250114_s_at	AT5G16370; AT5G16340	AT5G16370, AMP-binding protein, putative; AT5G16340, AMP-binding protein, putative
260774_at	AT1G78290	serine/threonine protein kinase, putative
264467_at	AT1G10140	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G58420.1); contains InterPro domain Uncharacterised conserved protein UCP031279 (InterPro:IPR016972)
245244_at	AT1G44350	ILL6 (IAA-leucine resistant (ILR)-like gene 6); metallopeptidase
258327_at	AT3G22640	cupin family protein
246072_at	AT5G20240	PI (PISTILLATA); DNA binding / transcription factor
256542_at	AT1G42550	PMI1 (PLASTID MOVEMENT IMPAIRED1)
257860_at	AT3G13062	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G55960.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO41766.1); contains InterPro domain Lipid-binding START (InterPro:IPR002913)
260668_at	AT1G19530	similar to unnamed protein product *Vitis vinifera* (GB:CAO61141.1)
267181_at	AT2G37760	aldo/keto reductase family protein
246158_at	AT5G19855	similar to unnamed protein product *Vitis vinifera* (GB:CAO21907.1)
262113_at	AT1G02820	late embryogenesis abundant 3 family protein / LEA3 family protein

ATH1 Array identifier	Gene identifier	Gene description
245016_at	ATCG00500	Encodes the carboxytransferase beta subunit of the Acetyl-CoA carboxylase (ACCase) complex in plastids. This complex catalyzes the carboxylation of acetyl-CoA to produce malonyl-CoA, the first committed step in fatty acid synthesis.
262373_at	AT1G73120	similar to hypothetical protein *Vitis vinifera* (GB:CAN69175.1)
250881_at	AT5G04080	similar to unknown protein *Oryza sativa* (*japonica* cultivar-group) (GB:BAD87214.1)
254797_at	AT4G13030	similar to unnamed protein product *Vitis vinifera* (GB:CAO18001.1); contains domain SSF52540 (SSF52540)
245543_at	AT4G15260	UDP-glucoronosyl/UDP-glucosyl transferase family protein
255881_at	AT1G67070	DIN9 (DARK INDUCIBLE 9); mannose-6-phosphate isomerase
260983_at	AT1G53560	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G17080.1); similar to unknown *Populus trichocarpa* (GB:ABK95219.1); contains domain PTHR10052 (PTHR10052); contains domain PTHR10052:SF2 (PTHR10052:SF2)
260056_at	AT1G78140	methyltransferase-related
256024_at	AT1G58340	ZF14; transporter
259685_at	AT1G63090	ATPP2-A11 (Phloem protein 2-A11); carbohydrate binding
263475_at	AT2G31945	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G05575.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO22015.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN61524.1); similar to hypothetical protein *Vitis vinifera* (GB:CAN71356.1)
251342_at	AT3G60690	auxin-responsive family protein
263231_at	AT1G05680	UDP-glucoronosyl/UDP-glucosyl transferase family protein
264246_at	AT1G60140	ATTPS10 (TREHALOSE PHOSPHATE SYNTHASE); transferase, transferring glycosyl groups / trehalose-phosphatase
258923_at	AT3G10450	SCPL7; serine carboxypeptidase
252167_at	AT3G50560	short-chain dehydrogenase/reductase (SDR) family protein
247326_at	AT5G64110	peroxidase, putative
250971_at	AT5G02810	PRR7 (PSEUDO-RESPONSE REGULATOR 7); transcription regulator
256548_at	AT3G14770	nodulin MtN3 family protein
264774_at	AT1G22890	unknown protein
250098_at	AT5G17350	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT3G03280.1); similar to unknown *Populus trichocarpa* (GB:ABK95625.1)
256308_s_at	AT1G30420; AT1G30410	AT1G30420, ATMRP12 (Arabidopsis thaliana multidrug resistance-associated protein 12); AT1G30410, ATMRP13 (Arabidopsis thaliana multidrug resistance-associated protein 13)
253829_at	AT4G28040	nodulin MtN21 family protein

ATH1 Array identifier	Gene identifier	Gene description
259937_s_at	AT3G13080; AT1G71330	AT3G13080, ATMRP3 (Arabidopsis thaliana multidrug resistance-associated protein 3); AT1G71330, ATNAP5 (Arabidopsis thaliana non-intrinsic ABC protein 5)
265414_at	AT2G16660	nodulin family protein
248564_at	AT5G49700	DNA-binding protein-related
252482_at	AT3G46670	UDP-glucoronosyl/UDP-glucosyl transferase family protein
248487_at	AT5G51070	ERD1 (EARLY RESPONSIVE TO DEHYDRATION 1); ATP binding / ATPase
259970_at	AT1G76570	chlorophyll A-B binding family protein
258114_at	AT3G14660	CYP72A13 (cytochrome P450, family 72, subfamily A, polypeptide 13); oxygen binding
264383_at	AT2G25080	ATGPX1 (GLUTATHIONE PEROXIDASE 1); glutathione peroxidase
252234_at	AT3G49780	ATPSK4 (PHYTOSULFOKINE 4 PRECURSOR); growth factor
246197_at	AT4G37010	caltractin, putative / centrin, putative
266832_at	AT2G30040	MAPKKK14 (Mitogen-activated protein kinase kinase kinase 14); kinase
263182_at	AT1G05575	unknown protein
253517_at	AT4G31390	ABC1 family protein
247734_at	AT5G59400	similar to hypothetical protein *Vitis vinifera* (GB:CAN64889.1)
265680_at	AT2G32150	haloacid dehalogenase-like hydrolase family protein
262607_at	AT1G13990	similar to unnamed protein product *Vitis vinifera* (GB:CAO68469.1)
253666_at	AT4G30270	MERI5B (MERISTEM-5); hydrolase, acting on glycosyl bonds / xyloglucan:xyloglucosyl transferase
246463_at	AT5G16970	AT-AER (ALKENAL REDUCTASE); 2-alkenal reductase
262635_at	AT1G06570	PDS1 (PHYTOENE DESATURATION 1)
248790_at	AT5G47450	AtTIP2;3 (Arabidopsis thaliana tonoplast intrinsic protein 2;3); water channel
246854_at	AT5G26200	mitochondrial substrate carrier family protein
249467_at	AT5G39610	ANAC092/ATNAC2/ATNAC6 (Arabidopsis NAC domain containing protein 92); protein heterodimerization/ protein homodimerization/ transcription factor
249923_at	AT5G19120	aspartic-type endopeptidase/ pepsin A
264102_at	AT1G79270	ECT8 (evolutionarily conserved C-terminal region 8)
247848_at	AT5G58120	disease resistance protein (TIR-NBS-LRR class), putative
267300_at	AT2G30140	UDP-glucoronosyl/UDP-glucosyl transferase family protein
253061_at	AT4G37610	BT5 (BTB and TAZ domain protein 5); protein binding / transcription regulator
253859_at	AT4G27657	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G27652.1)

ATH1 Array identifier	Gene identifier	Gene description
265200_s_at	AT2G36800; AT2G36790	AT2G36800, DOGT1 (DON-GLUCOSYLTRANSFERASE); UDP-glycosyltransferase/ transferase, transferring glycosyl groups; AT2G36790, UGT73C6 (UDP-GLUCOSYL TRANSFERASE 73C6); UDP-glucosyltransferase/ UDP-glycosyltransferase/ transferase, transferring glycosyl groups
250007_at	AT5G18670	BMY3 (BETA-AMYLASE 9); beta-amylase
263443_at	AT2G28630	beta-ketoacyl-CoA synthase family protein
251642_at	AT3G57520	ATSIP2 (ARABIDOPSIS THALIANA SEED IMBIBITION 2); hydrolase, hydrolyzing O-glycosyl compounds
251603_at	AT3G57760	protein kinase family protein
260662_at	AT1G19540	isoflavone reductase, putative
259856_at	AT1G68440	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G25400.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO42150.1)
262033_at	AT1G37140	MCT1 (mei2 C-Terminal RRM only like 1); RNA binding
262229_at	AT1G68620	hydrolase
267461_at	AT2G33830	dormancy/auxin associated family protein
256302_at	AT1G69526	UbiE/COQ5 methyltransferase family protein
264280_at	AT1G61820	BGLU46; hydrolase, hydrolyzing O-glycosyl compounds
256772_at	AT3G13750	BGAL1 (BETA GALACTOSIDASE 1); beta-galactosidase
265501_at	AT2G15490	UGT73B4; UDP-glycosyltransferase/ transferase, transferring glycosyl groups
260284_at	AT1G80380	phosphoribulokinase/uridine kinase-related
263827_at	AT2G40420	amino acid transporter family protein
265768_at	AT2G48020	sugar transporter, putative
251195_at	AT3G62930	glutaredoxin family protein
262291_at	AT1G70790	C2 domain-containing protein
248703_at	AT5G48430	aspartic-type endopeptidase/ pepsin A
258975_at	AT3G01970	WRKY45 (WRKY DNA-binding protein 45); transcription factor
266656_at	AT2G25900	ATCTH (Arabidopsis thaliana Cys3His zinc finger protein); transcription factor
255893_at	AT1G17960	threonyl-tRNA synthetase, putative / threonine–tRNA ligase, putative
255733_at	AT1G25400	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT1G68440.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO42150.1)
247745_at	AT5G59030	COPT1 (COPPER TRANSPORTER 1); copper ion transmembrane transporter
265197_at	AT2G36750	UGT72C1 (UDP-GLUCOSYL TRANSFERASE 72C1); UDP-glycosyltransferase/ transferase, transferring glycosyl groups
247282_at	AT5G64240	ATMC3 (METACASPASE 3); caspase
248460_at	AT5G50915	basic helix-loop-helix (bHLH) family protein
254707_at	AT4G18010	IP5PII (INOSITOL POLYPHOSPHATE 5-PHOSPHATASE II); inositol-polyphosphate 5-phosphatase

ATH1 Array identifier	Gene identifier	Gene description
260890_at	AT1G29090	peptidase C1A papain family protein
252415_at	AT3G47340	ASN1 (DARK INDUCIBLE 6)
248759_at	AT5G47610	zinc finger (C3HC4-type RING finger) family protein
255926_at	AT1G22190	AP2 domain-containing transcription factor, putative
266799_at	AT2G22860	ATPSK2 (PHYTOSULFOKINE 2 PRECURSOR); growth factor
264016_at	AT2G21220	auxin-responsive protein, putative
259927_at	AT1G75100	JAC1 (J-DOMAIN PROTEIN REQUIRED FOR CHLOROPLAST ACCUMULATION RESPONSE 1); heat shock protein binding
259911_at	AT1G72680	cinnamyl-alcohol dehydrogenase, putative
261663_at	AT1G18330	EPR1 (EARLY-PHYTOCHROME-RESPONSIVE1); DNA binding / transcription factor
264561_at	AT1G55810	uracil phosphoribosyltransferase, putative / UMP pyrophosphorylase, putative / UPRTase, putative
247780_at	AT5G58770	dehydrodolichyl diphosphate synthase, putative / DEDOL-PP synthase, putative
251072_at	AT5G01740	similar to SAG20 (WOUND-INDUCED PROTEIN 12) *Arabidopsis thaliana* (TAIR:AT3G10985.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO48816.1); contains InterPro domain Wound-induced protein, Wun1 (InterPro:IPR009798)
265573_at	AT2G28200	nucleic acid binding / transcription factor/ zinc ion binding
250935_at	AT5G03240	UBQ3 (POLYUBIQUITIN 3); protein binding
262047_at	AT1G80160	lactoylglutathione lyase family protein / glyoxalase I family protein
260287_at	AT1G80440	kelch repeat-containing F-box family protein
252487_at	AT3G46660	UDP-glucoronosyl/UDP-glucosyl transferase family protein
266296_at	AT2G29420	ATGSTU7 (GLUTATHIONE S-TRANSFERASE 25); glutathione transferase
263157_at	AT1G54100	ALDH7B4 (ALDEHYDE DEHYDROGENASE 7B4); 3-chloroallyl aldehyde dehydrogenase
261648_at	AT1G27730	STZ (SALT TOLERANCE ZINC FINGER); nucleic acid binding / transcription factor/ zinc ion binding
266719_at	AT2G46830	CCA1 (CIRCADIAN CLOCK ASSOCIATED 1); transcription factor
249148_at	AT5G43260	chaperone protein dnaJ-related
258402_at	AT3G15450	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT4G27450.1); similar to unknown *Populus trichocarpa* (GB:ABK93866.1); contains domain PTHR11772 (PTHR11772); contains domain G3DSA:3.60.20.10 (G3DSA:3.60.20.10); contains domain SSF56235 (SSF56235)
263632_at	AT2G04795	similar to unknown protein *Arabidopsis thaliana* (TAIR:AT5G35732.1); similar to unnamed protein product *Vitis vinifera* (GB:CAO24128.1)
262473_at	AT1G50250	FTSH1 (FtsH protease 1); ATP-dependent peptidase/ ATPase/ metallopeptidase

ATH1 Array identifier	Gene identifier	Gene description
245505_at	AT4G15690	glutaredoxin family protein
267168_at	AT2G37770	aldo/keto reductase family protein
266693_at	AT2G19800	MIOX2 (MYO-INOSITOL OXYGENASE 2)
247013_at	AT5G67480	BT4 (BTB AND TAZ DOMAIN PROTEIN 4); protein binding / transcription regulator

Table 2-S3. Genes showing high expression in *spch*.
List of 392 genes representing gene cluster II from Figure 2. ATH1 array, respective gene locus identifiers and gene descriptions are indicated. Coexpression was determined according to Table 2-S1.

Table 2-S4. RT-PCR

Gene	Forward Primer	Reverse Primer
ACTIN2	GCCATCCAAGCTGTTCTCTC	GCTCGTAGTCAACAGCAACAA
ARR16	CACCCTCGAGATGAACAGTTCAGGAGGTTC	TGATGACTCCTTCACTTTC
CLE9	CACCGTTCGTGTGATTCCTTC	CCAGAGGGGACGAGTCTTTTG
EPF1	CACCCTCGAGATGAAGTCTCTTCTTCTCCTTG	TTGAATTCTAGGGACAGGGTAGGACTTAT
EPF2	CACCCTCGAGATGACGAAGTTTGTACGCAAGT	ACGGCGGAGATTCAATTCAAG
ERL1	CGCATAACTTGCGGGAATTTG	AGTCCTTGTGCAGCTCCAACC
ERL2	GCTGTGGATAACGAGGCCAAC	CATGGTGGGTCTCTCCAAAGG
JAZ10	ACGAGTCGTCGATGGAGACAG	GGCCGATGTCGGATAGTAAGG
MUTE	TCGAAGAAGGCAAATGAACGAG	GGGTAGTGGCGGCTCCTAAAC
POLAR	TCTTGGTCCCTCCTACGAAGC	GCACGATAGTGCACCCATCTC
TMM	CACTCACAAGCTGTGGATCGTT	GACGGGAATGGACCTGATAACC

Table 2-S4. qRT-PCR

Gene	Forward Primer	Reverse Primer
ACTIN2	GATGAGGCAGGTCCAGGAATC	AACCCCAGCTTTTTAAGCCTTT
ARR16	CCGATTACTGTATGCCTGGAA	TGAGCTCCACTCGCTAAACAT
CLE9	CACCGTTCGTGTGATTCCTTC	CCAGAGGGGACGAGTCTTTTG
EPF1	ATGCCGTCTTGTGATGGTTAG	TCAAGGGACAGGGTAGGACTT
EPF2	TTTGGTCGTTAACTCCATTCG	ATCCGGTAAGCTTGATCCTGT
ERL1	CGCATAACTTGCGGGAATTTG	AGTCCTTGTGCAGCTCCAACC
ERL2	GCTGTGGATAACGAGGCCAAC	CATGGTGGGTCTCTCCAAAGG
JAZ10	ACGAGTCGTCGATGGAGACAG	GGCCGATGTCGGATAGTAAGG
MUTE	TCGAAGAAGGCAAATGAACGAG	GGGTAGTGGCGGCTCCTAAAC
POLAR	TCTTGGTCCCTCCTACGAAGC	GCACGATAGTGCACCCATCTC
TMM	AGCTGAGGCTCAACGATAACA	CCTCAGCTTTCTCCTCATCCT
SCZ	ACGAGAGATTACGACGAAGCA	AACAGGCTTGACATGGTTTTG
CDKB1;2	ATGCAAAGGTATGGCATTCTG	TGTCTTGGGATCCATCAAGAG
CYCA1;1	GAGGCTAAGAAGCGACCTGAT	ACGTCTCCGGTACAAGCCTAT
CYCA2;2	CATCTCCCAAACGAGTCACAT	GTTTTCGATGTTGCTTGGTGT
CYCA2;3	AAAGACCAGCCTTGGAGGATA	ATGACCGCGTCCTTTCTTTAT
CYCB1;2	TGTACCTCACCGTCAACATCA	AAGCAATAAGCAAGGCACTGA
CYCB1;3	GCTAGAAAACGCACAAGTTGC	GCTTTCTTTTGAACCGCTCTT
AT5G10890	CCTTCCGAATTGCAAAGAAGG	TTGCCTCAGTTCTTCGTTTCG

Table 2-S4. Primers used for transcript amplification in RT-PCR and qRT-PCR.
Gene name and sequences for forward and reverse primers are given for each gene amplified. All sequences are in 5′ to 3′ orientation.

Table 2-S5.

Fold change	Gene identifier	Gene description
7.23	At1g33811	GDSL-like Lipase/Acylhydrolase superfamily protein
6.57	At5g06530	ABC-2 type transporter family protein
11.17	265674_at	no match
12.92	At4g18280	glycine-rich cell wall protein-related
60.61	At2g39030	Acyl-CoA N-acyltransferases (NAT) superfamily protein
11.49	At1g10585	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
13.40	At1g24400	AATL2_ATLHT2_LHT2__lysine histidine transporter 2
6.12	At1g78530	Protein kinase superfamily protein
18.86	At1g75880	SGNH hydrolase-type esterase superfamily protein
9.21	At1g19620	unknown protein; Has 44 Blast hits to 24 proteins in 12 species: Bacteria - 8; Metazoa - 22; Plants - 6; Other Eukaryotes - 8 (source: NCBI BLink).
6.29	At3g24140	FMA__basic helix-loop-helix (bHLH) DNA-binding superfamily protein
18.72	At5g18430	GDSL-like Lipase/Acylhydrolase superfamily protein
3.96	At5g07260	START (StAR-related lipid-transfer) lipid-binding domain
5.01	At2g30140	UDP-Glycosyltransferase superfamily protein
5.04	At1g10060	ATBCAT-1_BCAT-1__branched-chain amino acid transaminase 1
10.02	At2g46070	ATMPK12_MAPK12_MPK12__mitogen-activated protein kinase 12
3.78	At1g60630	Leucine-rich repeat protein kinase family protein
5.20	At1g28010	ABCB14_ATABCB14_PGP14__P-glycoprotein 14
15.70	At3g17070 At3g42570	Peroxidase family protein
12.68	At2g39690	Protein of unknown function, DUF547
33.05	At1g19610	LCR78_PDF1.4__Arabidopsis defensin-like protein
4.93	At1g17230	Leucine-rich receptor-like protein kinase family protein
12.68	At4g24040	ATTRE1_TRE1__trehalase 1
3.69	At1g11340	S-locus lectin protein kinase family protein
7.23	At2g16630	Pollen Ole e 1 allergen and extensin family protein
4.81	At1g24470	ATKCR2_KCR2__beta-ketoacyl reductase 2
3.58	At1g75730	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages.
11.74	At4g12430 At4g12432	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
4.81	At3g26230	CYP71B24__cytochrome P450, family 71, subfamily B, polypeptide 24
6.45	At2g28410	unknown protein; Has 18 Blast hits to 18 proteins in 5 species: Plants - 18 (source: NCBI BLink).
10.67	At4g38420	sks9__SKU5 similar 9
8.28	At5g19580	glyoxal oxidase-related protein
2.57	At1g75030	ATLP-3_TLP-3__thaumatin-like protein 3
10.15	At4g14480	Protein kinase superfamily protein
9.37	253960_at	no match
8.90	At3g53720	ATCHX20_CHX20__cation/H+ exchanger 20

Fold change	Gene identifier	Gene description
6.73	At5g25840	Protein of unknown function (DUF1677)
3.79	At4g20050	QRT3__Pectin lyase-like superfamily protein
4.20	At1g62300	ATWRKY6_WRKY6__WRKY family transcription factor
3.29	At1g12320	Protein of unknown function (DUF1442)
4.44	At3g24670	Pectin lyase-like superfamily protein
3.12	At5g43330	Lactate/malate dehydrogenase family protein
12.97	At4g12500	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
11.56	At3g44860 At3g44870	FAMT__farnesoic acid carboxyl-O-methyltransferase
4.19	At3g27400	Pectin lyase-like superfamily protein
19.24	At3g62820	Plant invertase/pectin methylesterase inhibitor superfamily protein
3.86	At2g04040	ATDTX1_TX1__MATE efflux family protein
20.00	At2g24850	TAT_TAT3__tyrosine aminotransferase 3
29.87	At4g12490	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
56.72	At1g65500	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G65486.1); Has 23 Blast hits to 23 proteins in 2 species: Plants - 23 (source: NCBI BLink).
3.86	At1g47410	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages.
7.10	At5g02540	NAD(P)-binding Rossmann-fold superfamily protein
4.69	At3g20300	Protein of unknown function (DUF3537)
6.19	At3g14620	CYP72A8__cytochrome P450, family 72, subfamily A, polypeptide 8
11.12	At4g02330	ATPMEPCRB__Plant invertase/pectin methylesterase inhibitor superfamily
28.09	At3g55500	ATEXP16_ATEXPA16_ATHEXP ALPHA 1.7_EXP16_EXPA16__expansin A16
6.11	At1g62400	HT1__Protein kinase superfamily protein
7.50	At4g25260	Plant invertase/pectin methylesterase inhibitor superfamily protein
10.15	At2g04570	GDSL-like Lipase/Acylhydrolase superfamily protein
7.92	At2g25470	AtRLP21_RLP21__receptor like protein 21
3.48	At2g25000	ATWRKY60_WRKY60__WRKY DNA-binding protein 60
4.86	At4g37370	CYP81D8__cytochrome P450, family 81, subfamily D, polypeptide 8
13.51	At4g17970	ALMT12_ATALMT12__aluminum-activated, malate transporter 12
3.77	At5g58580	ATL63_TL63__TOXICOS EN LEVADURA 63
8.13	At2g43520	ATTI2_TI2__trypsin inhibitor protein 2
2.56	At5g66460	Glycosyl hydrolase superfamily protein
2.12	At4g36250	ALDH3F1__aldehyde dehydrogenase 3F1
3.10	At2g31980	AtCYS2_CYS2__PHYTOCYSTATIN 2

Fold change	Gene identifier	Gene description
2.99	At3g03680	C2 calcium/lipid-binding plant phosphoribosyltransferase family protein
9.49	At1g04110	SDD1__Subtilase family protein
4.94	At2g21080	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: vacuole; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF3537 (InterPro:IPR021924); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF3537) (TAIR:AT3G20300.1); Has 141 Blast hits to 141 proteins in 16 species: Plants - 140; Other Eukaryotes - 1 (source: NCBI BLink).
2.70	At3g47750 At3g47760	ABCA4_ATATH3__ATP binding cassette subfamily A4
6.23	At5g22300	AtNIT4_NIT4__nitrilase 4
2.19	At1g20630	CAT1__catalase 1
3.85	At2g44570	AtGH9B12_GH9B12__glycosyl hydrolase 9B12
8.45	At3g57260	BG2_BGL2_PR-2_PR2__beta-1,3-glucanase 2
8.30	At1g08810	AtMYB60_MYB60__myb domain protein 60
3.43	At5g62080	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
2.08	At3g48580	XTH11__xyloglucan endotransglucosylase/hydrolase 11
3.97	At5g52810	NAD(P)-binding Rossmann-fold superfamily protein
4.42	At1g52910	Protein of unknown function (DUF1218)
6.51	At1g30720 At1g30730	FAD-binding Berberine family protein
3.15	At4g38810	Calcium-binding EF-hand family protein
4.67	At1g77020	DNAJ heat shock N-terminal domain-containing protein
5.27	At2g24520	AHA5_HA5__H(+)-ATPase 5
3.65	At5g39050	HXXXD-type acyl-transferase family protein
7.79	At2g20875	EPF1__epidermal patterning factor 1
3.96	At4g23920	ATUGE2_UGE2__UDP-D-glucose/UDP-D-galactose 4-epimerase 2
4.41	At2g29940	ATPDR3_PDR3__pleiotropic drug resistance 3
8.12	At1g21400 At5g34780	Thiamin diphosphate-binding fold (THDP-binding) superfamily protein
11.29	At2g43510	ATTI1_TI1__trypsin inhibitor protein 1
7.22	At2g46720 At3g10280	HIC_KCS13__3-ketoacyl-CoA synthase 13
3.40	At5g62850	AtVEX1__Nodulin MtN3 family protein
7.23	At2g28870	unknown protein; Has 34 Blast hits to 34 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 34; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).
24.08	At1g21250	PRO25_WAK1__cell wall-associated kinase
3.82	At5g66440	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G34560.1); Has 1807 Blast hits to 1807 proteins in 277 species: Metazoa - 736; Fungi - 347; Plants - 385; Other Eukaryotes - 339 (source: NCBI BLink).

Fold change	Gene identifier	Gene description
3.33	At2g26560	PLA IIA_PLA2A_PLP2_PLP2__phospholipase A 2A
58.48	At2g43620	Chitinase family protein
5.86	At1g26600	CLE9__CLAVATA3/ESR-RELATED 9
2.38	At1g04990	Zinc finger C-x8-C-x5-C-x3-H type family protein
3.74	At1g66910 At1g66920	Protein kinase superfamily protein
4.05	At1g22690	Gibberellin-regulated family protein
6.85	At1g63180	UGE3__UDP-D-glucose/UDP-D-galactose 4-epimerase 3
4.69	At1g64660	ATMGL_MGL__methionine gamma-lyase
2.93	At2g37760	NAD(P)-linked oxidoreductase superfamily protein
6.52	At3g60420	Phosphoglycerate mutase family protein
2.26	At3g58540	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 4 anthesis, F mature embryo stage, petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G06190.2); Has 6 Blast hits to 6 proteins in 1 species: Plants - 6 (source: NCBI BLink).
2.75	At2g42280	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
9.15	At1g68500	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G25422.1); Has 16 Blast hits to 16 proteins in 6 species: Plants - 16 (source: NCBI BLink).
3.97	At3g25620	ABC-2 type transporter family protein
6.39	At1g68600	Aluminium activated malate transporter family protein
4.28	At5g48485	DIR1__Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
4.90	At3g51660	Tautomerase/MIF superfamily protein
3.42	At5g46230	Protein of unknown function, DUF538
6.99	At1g02850	BGLU11__beta glucosidase 11
4.24	At4g18050	PGP9__P-glycoprotein 9
4.31	At3g47480	Calcium-binding EF-hand family protein
3.77	At3g13450	DIN4__Transketolase family protein
3.49	At1g78780	pathogenesis-related family protein
7.85	At5g13220	JAS1_JAZ10_TIFY9__jasmonate-zim-domain protein 10
4.36	At5g55930	ATOPT1_OPT1__oligopeptide transporter 1
10.81	At2g37770	NAD(P)-linked oxidoreductase superfamily protein
4.58	At4g16690	ATMES16_MES16__methyl esterase 16
2.37	At1g75170 At5g04781	Sec14p-like phosphatidylinositol transfer family protein
3.06	At4g28390	AAC3_ATAAC3__ADP/ATP carrier 3
4.68	At3g51760	Protein of unknown function (DUF688)
5.44	At5g47330	alpha/beta-Hydrolases superfamily protein
10.58	At3g18250	Putative membrane lipoprotein

Fold change	Gene identifier	Gene description
2.09	At4g18910	ATNLM2_NIP1;2_NLM2__NOD26-like intrinsic protein 1;2
4.18	At3g30720	QQS__qua-quine starch
6.26	At4g36670	Major facilitator superfamily protein
4.20	At5g43450	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
5.16	At2g21590	APL4__Glucose-1-phosphate adenylyltransferase family protein
3.48	At5g67020	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G50340.1); Has 1807 Blast hits to 1807 proteins in 277 species: Metazoa - 736; Fungi - 347; Plants - 385; Other Eukaryotes - 339 (source: NCBI BLink).
4.83	At2g38870	Serine protease inhibitor, potato inhibitor I-type family protein
2.00	At3g18040	MPK9__MAP kinase 9
9.50	At1g21540	AMP-dependent synthetase and ligase family protein
7.18	At1g04800	glycine-rich protein
2.39	At4g37060 At4g37070	AtPLAIVB_PLA IVB_PLP5__PATATIN-like protein 5
2.74	At4g13210	Pectin lyase-like superfamily protein
3.70	At1g72680	ATCAD1_CAD1__cinnamyl-alcohol dehydrogenase
2.02	At3g43720	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
4.49	At5g49520	ATWRKY48_WRKY48__WRKY DNA-binding protein 48
3.44	At3g28960	Transmembrane amino acid transporter family protein
4.03	At5g06870	ATPGIP2_PGIP2__polygalacturonase inhibiting protein 2
17.01	At2g17740	Cysteine/Histidine-rich C1 domain family protein
21.04	At3g49620	DIN11__2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
37.27	At4g22470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
3.41	At5g17860	CAX7__calcium exchanger 7
4.94	At5g55450	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
3.16	At3g16150	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
2.25	At2g46420	Plant protein 1589 of unknown function
5.11	At4g12470	AZI1__azelaic acid induced 1
2.95	At4g02280	ATSUS3_SUS3__sucrose synthase 3
4.02	At5g24160	SQE6__squalene monoxygenase 6
3.45	At2g13610	ABC-2 type transporter family protein
22.79	At5g46050	ATPTR3_PTR3__peptide transporter 3
2.59	At4g24510	CER2_VC-2_VC2__HXXXD-type acyl-transferase family protein
7.71	At5g26340	ATSTP13_MSS1_STP13__Major facilitator superfamily protein
2.33	At3g19850	Phototropic-responsive NPH3 family protein
2.12	At1g71770	PAB5__poly(A)-binding protein 5
2.05	At2g34070	TBL37__TRICHOME BIREFRINGENCE-LIKE 37
2.90	At3g19930	ATSTP4_STP4__sugar transporter 4

Fold change	Gene identifier	Gene description
2.09	At1g25450	CER60_KCS5__3-ketoacyl-CoA synthase 5
2.72	At4g26830	O-Glycosyl hydrolases family 17 protein
5.99	At1g79770	Protein of unknown function (DUF1677)
4.44	At3g51860	ATCAX3_ATHCX1_CAX1-LIKE_CAX3__cation exchanger 3
3.29	At3g17130	Plant invertase/pectin methylesterase inhibitor superfamily protein
6.11	At3g06120	MUTE__basic helix-loop-helix (bHLH) DNA-binding superfamily protein
5.62	At4g15680	Thioredoxin superfamily protein
2.55	At3g28910	ATMYB30_MYB30__myb domain protein 30
2.91	At2g21140	ATPRP2_PRP2__proline-rich protein 2
2.05	At2g21640	Encodes a protein of unknown function that is a marker for oxidative stress response.
7.81	At2g38300	myb-like HTH transcriptional regulator family protein
9.07	At2g32190 At2g32210	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G32210.1); Has 74 Blast hits to 74 proteins in 7 species: Plants - 74 (source: NCBI BLink).
10.33	At2g26440	Plant invertase/pectin methylesterase inhibitor superfamily
7.01	At2g39210	Major facilitator superfamily protein
0.46	At4g15630	Uncharacterised protein family (UPF0497)
0.43	At3g16370	GDSL-like Lipase/Acylhydrolase superfamily protein
0.44	At1g70560	SAV3_TAA1_WEI8__tryptophan aminotransferase of Arabidopsis 1
0.27	At5g17700	MATE efflux family protein
0.41	At3g04290	ATLTL1_LTL1__Li-tolerant lipase 1
0.28	At1g55260	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
0.39	At3g62270	HCO3- transporter family
0.35	At1g01120	KCS1__3-ketoacyl-CoA synthase 1
0.12	At2g28780	unknown protein; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: inflorescence meristem, root, flower; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF939, bacterial (InterPro:IPR010343); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G09450.1); Has 671 Blast hits to 667 proteins in 305 species: Bacteria - 588; Fungi - 2; Plants - 66; Other Eukaryotes - 15 (source: NCBI BLink).
0.18	At5g03120	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; Has 14 Blast hits to 14 proteins in 4 species: Plants - 14 (source: NCBI BLink).
0.28	At1g32450	NRT1.5__nitrate transporter 1.5
0.07	At1g29510	SAUR68__SAUR-like auxin-responsive protein family
0.22	At2g30010	TBL45__TRICHOME BIREFRINGENCE-LIKE 45

Fold change	Gene identifier	Gene description
0.32	At5g06270	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G11600.1); Has 1807 Blast hits to 1807 proteins in 277 species: Metazoa - 736; Fungi - 347; Plants - 385; Other Eukaryotes - 339 (source: NCBI BLink).
0.09	At5g36910	THI2.2__thionin 2.2
0.41	At3g26450	Polyketide cyclase/dehydrase and lipid transport superfamily protein
0.28	At4g10770	ATOPT7_OPT7__oligopeptide transporter 7
0.29	At1g65450	HXXXD-type acyl-transferase family protein
0.11	At5g65730	XTH6__xyloglucan endotransglucosylase/hydrolase 6
0.05	At1g66100	Plant thionin
0.24	At2g43010	PIF4_SRL2__phytochrome interacting factor 4
0.48	At1g75500	WAT1__Walls Are Thin 1
0.13	At1g29440	SAUR-like auxin-responsive protein family
0.40	At3g46780	PTAC16__plastid transcriptionally active 16
0.32	At1g04240	IAA3_SHY2__AUX/IAA transcriptional regulator family protein
0.31	At3g13980	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G54200.1); Has 1485 Blast hits to 418 proteins in 98 species: Bacteria - 6; Metazoa - 246; Fungi - 61; Plants - 107; Viruses - 6; Other Eukaryotes - 1059 (source: NCBI BLink).
0.26	At3g24290 At3g24300	AMT1;5__ammonium transporter 1;5
0.38	At1g24020	MLP423__MLP-like protein 423
0.33	At4g02850	phenazine biosynthesis PhzC/PhzF family protein
0.27	At1g14280	PKS2__phytochrome kinase substrate 2
0.40	At1g52290	Protein kinase superfamily protein
0.35	At4g17470	alpha/beta-Hydrolases superfamily protein
0.28	At2g15090	KCS8__3-ketoacyl-CoA synthase 8
0.48	At1g71695	Peroxidase superfamily protein
0.26	At3g46270	receptor protein kinase-related
0.49	At2g38120	AUX1_MAP1_PIR1_WAV5__Transmembrane amino acid transporter family protein
0.27	At5g02890	HXXXD-type acyl-transferase family protein
0.42	At1g52190	Major facilitator superfamily protein
0.40	At5g25610	ATRD22_RD22__BURP domain-containing protein
0.07	At1g29430 At5g27780	SAUR-like auxin-responsive protein family
0.35	At1g77490	TAPX__thylakoidal ascorbate peroxidase
0.29	At2g15050	LTP_LTP7__lipid transfer protein

Fold change	Gene identifier	Gene description
0.12	At1g13650	BEST Arabidopsis thaliana protein match is: 18S pre-ribosomal assembly protein gar2-related (TAIR:AT2G03810.4); Has 3688 Blast hits to 1629 proteins in 255 species: Archae - 22; Bacteria - 222; Metazoa - 684; Fungi - 292; Plants - 62; Viruses - 14; Other Eukaryotes - 2392 (source: NCBI BLink).
0.20	At5g44680	DNA glycosylase superfamily protein
0.45	At4g14550	IAA14_SLR__indole-3-acetic acid inducible 14
0.32	At5g19220	ADG2_APL1__ADP glucose pyrophosphorylase large subunit 1
0.31	At5g55620	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G09950.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Other Eukaryotes - 2996 (source: NCBI BLink).
0.41	At2g37640	ATEXP3_ATEXPA3_ATHEXP ALPHA 1.9_EXP3__Barwin-like endoglucanases superfamily protein
0.37	At4g37800	XTH7__xyloglucan endotransglucosylase/hydrolase 7
0.21	At5g01790	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 13 growth stages; Has 121 Blast hits to 121 proteins in 12 species: Plants - 121 (source: NCBI BLink).
0.24	At5g01240	LAX1__like AUXIN RESISTANT 1
0.42	At4g34980	SLP2__subtilisin-like serine protease 2
0.25	At1g70830	MLP28__MLP-like protein 28
0.32	At5g14970	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G14910.1); Has 579 Blast hits to 397 proteins in 95 species: Bacteria - 294; Plants - 86; Other Eukaryotes - 199 (source: NCBI BLink).
0.18	At5g45070	AtPP2-A8_PP2-A8__phloem protein 2-A8
0.19	At1g31690	Copper amine oxidase family protein
0.49	At5g23860	TUB8__tubulin beta 8
0.36	At1g51840	protein kinase-related
0.44	At3g52360	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to karrikin; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G35850.1); Has 34 Blast hits to 34 proteins in 10 species: Plants - 34 (source: NCBI BLink).
0.11	At1g29500	SAUR-like auxin-responsive protein family
0.21	At3g21950	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
0.40	At5g44400	FAD-binding Berberine family protein
0.20	At2g44230	Plant protein of unknown function (DUF946)
0.28	At4g00950	MEE47__Protein of unknown function (DUF688)

Fold change	Gene identifier	Gene description
0.30	At5g03760	ATCSLA09_ATCSLA9_CSLA09_CSLA9_RAT4__Nucleotide-diphospho-sugar transferases superfamily protein
0.35	At3g23050	AXR2_IAA7__indole-3-acetic acid 7
0.11	At1g06100	Fatty acid desaturase family protein
0.14	At4g04840	ATMSRB6_MSRB6__methionine sulfoxide reductase B6
0.03	At2g33850	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G28400.1); Has 3053 Blast hits to 2119 proteins in 133 species: Archae - 6; Bacteria - 52; Metazoa - 135; Fungi - 96; Plants - 73; Viruses - 2; Other Eukaryotes - 2689 (source: NCBI BLink).
0.09	At1g31690	Copper amine oxidase family protein
0.10	At5g33370	GDSL-like Lipase/Acylhydrolase superfamily protein

Table 2-S5. Genes significantly deregulated in the stomata-enriched background, *scrm-D*.
List of 248 genes with statistically significant deregulation in the *scrm-D* background compared to wild type. Fold-change, gene locus identifiers, and gene descriptions are given. Gene descriptions and significance were determined as described in Table 2-S1.

Table 2-S6.

Fold change	Gene identifier	Gene description
33.51	At1g19530	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: N-terminal protein myristoylation, anaerobic respiration; LOCATED IN: cellular_component unknown; EXPRESSED IN: leaf apex, inflorescence meristem, hypocotyl, root, flower; EXPRESSED DURING: petal differentiation and expansion stage; Has 47 Blast hits to 47 proteins in 13 species:Plants - 47 (source: NCBI BLink).
14.08	At4g31870	ATGPX7_GPX7__glutathione peroxidase 7
12.71	At4g16690	ATMES16_MES16__methyl esterase 16
11.42	At1g60740 At1g65970	Thioredoxin superfamily protein
7.89	At3g45300	ATIVD_IVD_IVDH__isovaleryl-CoA-dehydrogenase
7.50	At3g15620	UVR3__DNA photolyase family protein
5.82	At5g24800	ATBZIP9_BZIP9_BZO2H2__basic leucine zipper 9
4.86	At1g66760	MATE efflux family protein
4.60	At3g21090	ABC-2 type transporter family protein
4.40	At4g34138	UGT73B1__UDP-glucosyl transferase 73B1
4.26	At5g57240	ORP4C__OSBP(oxysterol binding protein)-related protein 4C
4.09	At1g58340	ZF14__MATE efflux family protein
3.81	At2g24100	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G30780.1); Has 101 Blast hits to 101 proteins in 16 species: Plants - 95; Other Eukaryotes - 6 (source: NCBI BLink).
3.53	At3g21870	CYCP2;1__cyclin p2;1
3.49	At4g15550	IAGLU__indole-3-acetate beta-D-glucosyltransferase
3.34	At1g62510	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
3.27	At2g34070	TBL37__TRICHOME BIREFRINGENCE-LIKE 37
3.23	At1g28010	ABCB14_ATABCB14_PGP14__P-glycoprotein 14
2.98	At5g03860	MLS__malate synthase
2.95	At2g33250	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G04310.1); Has 41 Blast hits to 41 proteins in 12 species: Plants - 41 (source: NCBI BLink).
2.84	At4g37150	ATMES9_MES9__methyl esterase 9
2.83	At1g64740	TUA1__alpha-1 tubulin
2.67	At4g09830	Uncharacterised conserved protein UCP009193
2.65	At1g66540	Cytochrome P450 superfamily protein
2.49	At3g56140	Protein of unknown function (DUF399 and DUF3411)
2.46	At1g72770	HAB1__homology to ABI1
2.41	At1g66330	senescence-associated family protein
2.28	At1g19680	RING/U-box superfamily protein
2.27	At1g78140	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein

Fold change	Gene identifier	Gene description
2.26	At1g53090	SPA4__SPA1-related 4
2.16	At2g31380	STH__salt tolerance homologue
2.13	At5g53160	PYL8_RCAR3__regulatory components of ABA receptor 3
2.10	At1g35670	ATCDPK2_ATCPK11_CDPK2_CPK11__calcium-dependent protein kinase 2
2.09	At5g20960	AAO1_AO1_AOalpha_AT-AO1_ATAO_AtAO1__aldehyde oxidase 1
2.06	At3g21670	Major facilitator superfamily protein
2.05	At5g48120	ARM repeat superfamily protein
2.04	At1g56280	ATDI19_DI19__drought-induced 19
2.03	At2g14060	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
2.02	At1g62960	ACS10__ACC synthase 10
0.47	At2g31390	pfkB-like carbohydrate kinase family protein
0.47	At4g00080	UNE11__Plant invertase/pectin methylesterase inhibitor superfamily protein
0.46	At2g44450	BGLU15__beta glucosidase 15
0.46	At4g31500	ATR4_CYP83B1_RED1_RNT1_SUR2__cytochrome P450, family 83, subfamily B, polypeptide 1
0.45	At1g64405	unknown protein; Has 13 Blast hits to 13 proteins in 7 species: Plants - 13 (source: NCBI BLink).
0.45	At3g22240	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G22235.2); Has 177 Blast hits to 177 proteins in 14 species: Plants - 177 (source: NCBI BLink).
0.44	At2g28250	NCRK__Protein kinase superfamily protein
0.44	At1g56320	BEST Arabidopsis thaliana protein match is: Glycine-rich protein family (TAIR:AT5G49350.2); Has 60 Blast hits to 60 proteins in 12 species: Plants - 60 (source: NCBI BLink).
0.43	At5g50375	CPI1__cyclopropyl isomerase
0.42	At1g06640	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
0.41	At1g78000	SEL1_SULTR1;2__sulfate transporter 1;2
0.41	At5g19890	Peroxidase superfamily protein
0.40	At3g14940	ATPPC3_PPC3__phosphoenolpyruvate carboxylase 3
0.39	At1g20450	ERD10_LTI29_LTI45__Dehydrin family protein
0.39	At2g14900	Gibberellin-regulated family protein
0.37	At3g07320	O-Glycosyl hydrolases family 17 protein
0.37	At2g23540	GDSL-like Lipase/Acylhydrolase superfamily protein
0.37	254756_at	no match
0.36	At4g29270	HAD superfamily, subfamily IIIB acid phosphatase
0.35	At3g23190	HR-like lesion-inducing protein-related
0.32	At2g21100	Disease resistance-responsive (dirigent-like protein) family protein
0.31	At3g54040	PAR1 protein
0.29	At3g06035	Glycoprotein membrane precursor GPI-anchored
0.28	At1g72230	Cupredoxin superfamily protein
0.25	At3g54260	TBL36__TRICHOME BIREFRINGENCE-LIKE 36

Fold change	Gene identifier	Gene description
0.25	At4g33720	CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 protein) superfamily protein
0.22	At5g37690	SGNH hydrolase-type esterase superfamily protein
0.20	At3g46270	receptor protein kinase-related
0.19	At2g39700	ATEXP4_ATEXPA4_ATHEXP ALPHA 1.6_EXPA4__expansin A4
0.18	At4g08770	Peroxidase superfamily protein
0.16	At3g13650	Disease resistance-responsive (dirigent-like protein) family protein
0.15	At5g05960	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
0.11	At3g02885	GASA5__GAST1 protein homolog 5
0.10	At5g59320	LTP3__lipid transfer protein 3
0.06	At4g23600	CORI3_JR2__Tyrosine transaminase family protein

Table 2-S6. Genes significantly deregulated in the pavement cell-enriched background, *spch*.
List of 74 genes with statistically significant deregulation in the *spch* background compared to wild type. Fold-change, gene locus identifiers, and gene descriptions are given. Gene descriptions and significance were determined as described in Table 2-S1.

Table 2-S7.

Gene Identifier	Description	Fold increase
At5g07260	Homeobox related	75-fold
At2g19980	TPX-1 related protein	41-fold
At1g26600	CLE9	30-fold
At2g41190	Amino acid transporter	25-fold
At4g28420	Aminotransferase	24-fold
At1g33930	AIG-responsive protein related	23-fold
At1g52910	Unknown	22-fold
At4g31805	POLAR, unknown function	21-fold
At5g62210	Embryo-specific protein related	20-fold *
At1g12845	Unknown	20-fold
At5g07280	EMS1	19-fold
At5g13220	JAZ10	15-fold
At2g40670	ARR16 Arabidopsis Response Regulator	14-fold
At3g17640	LRR-family protein	9-fold
At4g29740	CKX4 cytokinin oxidase	6-fold

* indicates fold increase was in comparison to *spch* seedlings.

Table 2-S7. Genes used for production of GFP reporter constructs.
All genes were identified as upregulated in the meristemoid-enriched *scrm-D mute* genotype as determined by comparison to wild type seedlings (Table 2-S1), except where indicated. Gene locus identifiers, gene description, and fold change are listed.

Gene identifier	Gene Description	T-DNA stock
At1g12845	unknown protein	FLAG_472G02
At1g26600	CLE9 (CLAVATA3/ESR-RELATED 9); receptor binding	SALK_018122
At1g33930	avirulence-responsive family protein / avirulence induced gene (AIG1) family protein	SALK_060790C
At1g52910	unknown protein	SALK_013728
At2g19980	allergen V5/Tpx-1-related family protein	FLAG_094E06
At2g40670	ARR16 (response regulator 16)	SAIL_398_C12
At2g41190	amino acid transporter family protein	SALK_150035
At3g17640	leucine-rich repeat family protein	SAIL_652_C04
At4g29740	CKX4 (CYTOKININ OXIDASE 4)	SALK_146112
At4g31805	POLAR, unknown function	SALK_122914, SALK_122924, SALK_142820
At5g13220	JAZ10	SAIL_92_D08
At5g62210	embryo-specific protein-related	SALK_038225

Table 2-S8. T-DNA insertion lines examined in this study.

Available T-DNA insertion lines for each gene were identified using SIGnAL T-DNA Express Tool (http://signal.salk.edu/cgi-bin/tdnaexpress) and TAIR (www.arabidopsis.org). Gene locus identifiers and gene descriptions are listed.

ATH1 array identifier	Gene identifier	Gene Description
247471_at	At5g62230	ERL1__ERECTA-like 1
259083_at	At3g04810	ATNEK2_NEK2__NIMA-related kinase 2
256389_at	At3g06220	AP2/B3-like transcriptional factor family protein
250642_at	At5g07180	ERL2__ERECTA-like 2
266733_at	At2g03280	O-fucosyltransferase family protein
256780_at	At3g13640	ATRLI1_RLI1__RNAse l inhibitor protein 1
247643_at	At5g60450	ARF4__auxin response factor 4
262081_at	At1g59540	ZCF125__P-loop containing nucleoside triphosphate hydrolases superfamily protein
257929_at	At3g16980	NRPB9A_NRPD9A_NRPE9A__RNA polymerases M/15 Kd subunit
248710_at	At5g48480	Lactoylglutathione lyase / glyoxalase I family protein
262802_at	At1g20930	CDKB2;2__cyclin-dependent kinase B2;2
259407_at	At1g13320	PP2AA3__protein phosphatase 2A subunit A3
259998_at	At1g68120	ATBPC3_BPC3__basic pentacysteine 3
262003_at	At1g64460	Protein kinase superfamily protein
263239_at	At2g16570	ASE1_ATASE_ATASE1__GLN phosphoribosyl pyrophosphate amidotransferase 1
265373_at	At2g06510	ATRPA1A_ATRPA70A_RPA1A_RPA70A__replication protein A 1A
257020_at	At3g19590	BUB3.1__Transducin/WD40 repeat-like superfamily protein
265087_at	At1g03760	Prefoldin chaperone subunit family protein
251470_at	At3g59570	Ypt/Rab-GAP domain of gyp1p superfamily protein

Table 2-S9. Genes common to transient and permanent stem-cell populations.

List of 19 genes identified as upregulated in *scrm-D mute* plants (transient stem-cell population) and cells with active *CLV3*, *FIL*, or *WUS* promoters (permanent stem-cell populations). For all stem-cell populations, significance of upregulation was determined from comparison to wild type seedlings. ATH1 array, gene locus identifiers, and gene descriptions are given for each. Gene descriptions were derived from BAR (http://142.150.214.117/welcome.htm).

Chapter 3.

Arabidopsis homeodomain-leucine zipper IV proteins promote stomatal development and their ectopic expression induces stomata beyond the epidermis

This chapter was previously published as:
Kylee M. Peterson, Christine Shyu, Christian A. Burr, Robin J. Horst, Masahiro M. Kanaoka, Minami Omae, Yutaka Sato, and Keiko U. Torii (2013) *Development* 140, 1924-1935.

Introduction

Epidermis, the outermost cell layer of land plants, serves as an interface between plants and the surrounding environment. The shoot epidermis, which is derived from the L1 layer of the shoot apex, gives rise to specialized cell types -- pavement cells, stomatal guard cells, and trichomes -- to optimize the balance between protection and gas exchange. The cuticulated pavement cells form a tightly sealed barrier protecting plants from desiccation, UV damage, and pathogen entry, while the stomata act as valves for efficient gas exchange and transpiration (Dong and Bergmann, 2010; Javelle *et al.*, 2011b; Pillitteri and Torii, 2012). Trichomes are appendages for herbivore protection, and, in some plants, accumulate and secrete defensive chemicals (Hulskamp, 2004; Serna and Martin, 2006).

The epidermal layer is specified during embryogenesis. In *Arabidopsis*, two closely related genes, *A. thaliana MERISTEM LAYER1* (*AtML1*) and *PROTODERMAL FACTOR2* (*PDF2*), redundantly specify epidermal identity, and

their double loss of function results in an embryo lethally devoid of a protoderm (Abe *et al.,* 2003). AtML1 and PDF2 belong to a family of plant-specific homeodomain leucine zipper class IV (HD-ZIP IV) transcription factors (Mukherjee *et al.,* 2009), many showing preferential expression in the epidermis (Javelle *et al.*, 2011a; Nakamura *et al.,* 2006). The *Arabidopsis* genome contains sixteen *HD-ZIP IV* (*AtHD-ZIP IV*) genes, of which at least nine exhibit expression in leaf epidermal cells (Nakamura *et al*. 2006; Figure 3-S1). It is not known whether *AtML1* and *PDF2* are sufficient to specify epidermal fate. Some other HD-ZIP IV members are associated with differentiation of shoot epidermal cell types. For instance, *Arabidopsis GLABRA2* (*GL2*) promotes trichome differentiation (Rerie *et al.,* 1994), and *HDG11* and *HDG12* regulate trichome branching (Nakamura *et al.,* 2006). Microarray studies of maize plants overexpressing *HD-ZIP IV* gene *OUTER CELL LAYER1* (*OCL1*) identified target genes in lipid metabolism and cuticle biosynthesis, suggesting a role in pavement cell differentiation (Javelle *et al.,* 2010). Likewise, *Arabidopsis HDG1* overexpression upregulates epicuticular wax biosynthesis genes (Wu *et al.,* 2011). Thus far, it is not known if any HD-ZIP IV members promote differentiation of the remaining epidermal cell type, stomata.

Stomatal development occurs through a series of cell-state transitions. In *Arabidopsis*, a subpopulation of protodermal cells adopts the identity of a meristemoid mother cell (MMC) and initiates an asymmetric entry division that creates a meristemoid, a transiently amplifying stomatal lineage cell. The meristemoid reiterates asymmetric amplifying divisions but eventually differentiates into a guard mother cell (GMC), which divides symmetrically to form paired guard cells (GCs) constituting a stoma (Dong *et al.,* 2010; Pillitteri *et*

al., 2012).

The stomatal cell-state transitions are specified by combinatorial activities of five basic-helix-loop-helix (bHLH) proteins, SPEECHLESS (SPCH), MUTE, FAMA, SCREAM (SCRM, also known as ICE1), and SCRM2 (Kanaoka *et al.*, 2008; MacAlister *et al.*, 2007; Ohashi-Ito *et al.*, 2006; Pillitteri *et al.*, 2007). These bHLH proteins are necessary and sufficient for specifying each stomatal precursor state. Among them, *MUTE* and *SCRM* are capable of converting the aerial shoot epidermis into stomata when ectopically overexpressed or stabilized (Kanaoka *et al.*, 2008; Pillitteri *et al.*, 2008). In contrast, ectopic overexpression of *SPCH* and *FAMA* confer excessive asymmetric entry divisions and formation of guard-cell-like cells, respectively (MacAlister *et al.*, 2007; Ohashi-Ito *et al.*, 2006; Pillitteri *et al.*, 2007). Importantly, the ability of *MUTE* and *SCRM* to trigger stomatal differentiation is limited to the epidermis. Factors that enable stomatal differentiation beyond epidermal identity remain unknown.

Here we report a meristemoid-enriched AtHD-Zip IV protein, HDG2, as a key component linking epidermal identity to stomatal differentiation. *HDG2* ectopic expression confers striking, ectopic differentiation of stomata in internal mesophyll tissues and occasional multiple epidermal layers. Conversely, the loss-of-function *hdg2* mutant exhibits delayed meristemoid-to-GMC transition. Our work highlights the similarity and difference between *HDG2* and its close relative *AtML1* in specification of protodermal identity and their ability to override epidermal constraints of stomatal differentiation when ectopically expressed.

Materials and Methods

Plant materials and growth conditions

Arabidopsis thaliana Columbia (Col) accession was used as wild type. Mutants used in the study were in the Col background. The following mutants and reporter lines were reported previously: *TMMpro::GUS-GFP* (Nadeau and Sack, 2002), *TMMpro::GUS* (Shpak *et al.*, 2005), *EPF2pro::erGFP* (Hara *et al.*, 2009), *AtML1pro::GUS* (Uchida *et al.*, 2012), *SPCHpro::GUS* and *E994* (Pillitteri *et al.*, 2007), *GL2pro::GUS* (Masucci *et al.*, 1996), and *At5g17710pro::nuc-3xVENUS* (Roeder *et al.*, 2012). T-DNA insertion lines *hdg2-2* (SALK_127828C), *hdg2-3* (SALK_138646C), *hdg2-4* (SALK_120064), *atml1-3* (SALK_128172), *atml1-4* (SALK_033408), and *pdf2-2* (SALK_109425C) were obtained from the *Arabidopsis* Biological Resource Center (ABRC), and loss or disruption of their transcripts were confirmed (Figure 3-S2). Seedlings and plants were grown as described previously (Pillitteri *et al.*, 2011).

Microarray validation

Preparation and analysis of the meristemoid cell state microarray were described in Pillitteri et al. (2011). Original, DNase I-treated RNA samples for all microarray genotypes were converted to cDNA using the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA). Quantitative RT-PCR was performed as described previously (Pillitteri *et al.*, 2011) and normalized using the ΔΔCq method (Livak *et al.*, 2001). RT-PCR was performed as described previously (Lee *et al.*, 2012). See Table 3-1 for a list of primer sequences.

Molecular cloning and generation of transgenic plants

The following constructs were generated for transgenic studies using Gateway

cloning (Nakagawa *et al.,* 2008; Nakagawa *et al.,* 2007): pLJP249 (*35S::HDG2*), pKMP114 (*HDG2pro::HDG2-GFP*), pKMP139 (*HDG2pro::nls-3xGFP*). Plasmid pER8 (Zuo *et al.,* 2000) was used for estradiol-inducible constructs: pKMP145 (inducible *HDG2*), pKMP141 (inducible *HDG2-SRDX*), pKMP151 (inducible *AtML1*). Generation and selection of transgenic plants were performed as described by Pillitteri et al. (2007). For all transgenic *Arabidopsis* lines, more than 20 T1 plants per construct were isolated and characterized for phenotypes. Among them, 5-6 lines were selected based on single insertion status inferred by the segregation of resistance genes and stability of the phenotype in subsequent generations. Estradiol induction of *HDG2*, *HDG2-SRDX*, and *AtML1* were as previously reported (Lee *et al.,* 2012), and their induction was confirmed by RT-PCR using transgene-specific primers (see Figures 3-S3 and 3-S4). See Table 3-1 for a list of primer sequences used for molecular cloning and Table 3-2 for a list of plasmid constructs generated in this study.

Microscopy

Confocal laser scanning microscopy (CLSM) images were taken with either the Zeiss LSM700 (Thornwood, NY, USA) or the Leica SP5 (Solms, Germany). Cell peripheries were visualized with propidium iodide (PI: Molecular Probes, Carlsbad, CA), and GFP and PI signals were detected as described previously (Pillitteri *et al.,* 2011). mPS-PI staining was performed as previously described (Truernit *et al.,* 2008), with an ethanol treatment step of 2-3 minutes. Z-stack movies were produced using Quicktime 7 Pro (Apple, Cupertino, CA).

***In situ* hybridization**

Wild-type seedlings were grown on sterile plates for 10 days. *esHDG2* seedlings

were grown for 5 days and transplanted to induction medium with 10 μM estradiol for 5 more days. The seedlings were subjected to fixation, paraffin embedding, and in situ mRNA hybridization according to Kouchi and Hata (1993) with few modifications, as follows.

The seedlings were harvested and fixed in FAA (50% ethanol, 10% formaldehyde, 5% acetic acid) for 1 hour under vacuum at room temperature and for one more hour at 4°C without vacuum. Samples were dehydrated through a graded ethanol series then the solution was replaced with Histo-Clear (Cosmo Bio Inc., Japan). The samples were embedded in paraffin (Paraplast, Sigma-Aldrich) and cut to make 8 μm sections.

Partial *HDG2* cDNA fragments corresponding to exons 1 and 4 were amplified from pKMP144 using gene-specific primers and cloned into pCR4-TOPO vector (Invitrogen). Partial *FDH* cDNA fragment was amplified from first strand cDNA from 10-day-old Col seedlings using gene-specific primers and cloned into pCR4-TOPO vector. For primers and plasmid information, see Tables 3-1 and 3-2. These plasmids were used as templates for PCR amplification of *in situ* probes using M13 Forward (-20) and M13 Reverse primers. For probe synthesis, in vitro transcription and digoxigenin (DIG)-UTP labeling were done using MAXIscript T3 or T7 kit (Ambion, Life Technologies) according to manufacturer's instructions. RNA polymerases used for probe synthesis were as follows: *HDG2_exon1* (pMK370): antisense T3, sense T7; *HDG2_exon4* (pMK371): antisense T3, sense T7; *FDH_exon3* (pMK372): antisense T7, sense T3.

The sections were de-paraffinized in Histo-Clear, then rehydrated through a graded ethanol series. They were treated with proteinase K (final 5 μg/ml) in

ProK buffer (100 mM Tris-Cl [pH 7.5], 50 mM EDTA) at 37° C for 30 min and then washed twice with distilled water for 5 min each. They were subsequently treated with post-fixation solution (4% [w/v] paraformaldehyde in 0.1M Na-P buffer [pH 7.2]) for 10 min at room temperature and then washed twice with distilled water for 5 min each. Samples were treated with TEA solution (0.5% acetic anhydride in 0.1M triethanolamine) for 10 min at room temperature with stirring, washed twice in 2x SSPE for 5 min each, and then dehydrated through a graded ethanol series. Sections were dried in vacuum at room temperature for 1 hour, and subsequently incubated in a humid box at 45°C for 18 hours with hybridization buffer (200 μl per slide) containing the probes (0.1 μg/slide). The hybridization buffer consisted of 50% deionized formamide, 0.3 M NaC1, 10% dextran sulfate, 10 mM Tris-HCl, pH 7.5, 1mM EDTA, 60mM DTT, 1x Denhardt's, 1 μg/μl *Escherichia coli* tRNA, and 0.1 μg/μl polyA. For *HDG2* detection, both exon1 and exon4 probes were mixed and used.

After hybridization, successive washing steps were performed as follows: twice in 4x SSC (for 7 min and 8 min, at 45°C), treated with RNase A (50 μg/ml) in RNase buffer (0.5 M NaC1, 10 mM Tris-HCl [pH 7.5], 5 mM EDTA) at 37°C for 30 min, 3 times in RNase buffer at 37°C for 5 min each, and twice in 0.5x SSC at 45°C for 20 min each. The slides were soaked twice with Buffer1 (150 mM NaC1, 100 mM Tris-HCl, [pH7.5]) at room temperature for 5 min each and then incubated with 0.5% blocking reagent (Boehringer) in Buffer1 for 30 min at room temperature. After removing the blocking reagent, the slides were briefly rinsed with Buffer1, followed by incubation with the diluted anti-digoxigenin-alkaline phosphatase (Anti-DIG-AP) conjugate (1:1000) in Buffer1 containing 0.1% BSA and 0.25% Tween-20 for 60 min at room temperature. The slides were

subsequently washed three times with Buffer1 for 10 min each with agitation and with Buffer3 (100 mM NaC1, 50 mM MgCl2, 100mM Tris-HC1 [pH 9.5]) for 5 min, and then covered with Buffer3 containing 4.5 μl/ml nitroblue tetrazolium salt (NBT) solution and 3.5 μl/ml 5-bromo-4-chloro-3-indolyl phosphate toluidinium salt (BCIP) solution. After incubation at 37°C for 19 hours in the dark, the color reaction was stopped by immersing the slides in TE (pH 7.5). The sections were dehydrated through a graded ethanol series and then mounted in Eukitt (O. Kindler).

Tissue sectioning

Plastic embedding and sectioning using Technovit 7100 (Heraeus Kulzer, Wehrheim, Germany) and toluidine blue O (TBO) staining were performed as described previously (Shpak *et al.,* 2003). Histochemical staining for GUS activity was performed as described previously (Sessions *et al.,* 1999). Images were taken with the Olympus FV1000 (Center Valley, PA, USA). To visualize GFP expression inside leaves, fresh tissue was embedded in 3% low-melt agarose. The blocks were cut into 150-μm sections using Vibratome Series 1000 (Ted Pella. Inc., Redding, California), stained with propidium iodide, and imaged with the Zeiss LSM700 as above.

Quantitative analysis of stomatal phenotype

Stomatal index (SI = number of stomata/total number of stomata + non-stomatal epidermal cells x100), meristemoid index (MI =number of meristemoids/total number of stomata + non-stomatal epidermal cells x100), and stomatal-lineage index (SLI=number of stomata + meristemoids/total number of stomata + non-stomatal epidermal cells x100) were calculated using TBO-stained 10-day-old

cotyledons and fully expanded rosette leaves 4 from 6-week-old plants as previously reported (Guseman *et al.*, 2010).

Yeast one-hybrid and HDG2 autoactivation assays

The following plasmids were constructed and used in yeast one-hybrid assays: pCAB120 (Gal4AD-HDG2), pCAB121 (HAHR1-A box), pCAB122 (HAHR1-T box), pCAB123 (HAHR1-m box), pCAB124 (L1-C box), pCAB125 (L1-T box), pCAB126 (L1-m box), pCAB130 (Gal4AD-AtML1) and pCAB132 (Gal4AD-SCRM). Gal4AD-PDF2 was reported previously (Wu *et al.*, 2011). SCRM was used as a negative control. DNA binding was tested using the Matchmaker One-Hybrid system (Clontech Laboratories) and the yeast expression vector pDEST22 (Invitrogen). Target binding sequences were created by annealing oligonucleotide pairs containing three tandemly repeated target-binding sequences (see Table 3-3). Each target binding sequence was integrated into yeast strain YM4271 per manufacturer's protocols. DNA binding ability was assayed by testing β-galactosidase activity using 2-Nitrophenyl-β-D-galactopyranoside (ONPG; Sigma). All protocols were run according to manufacturer specifications. Yeast cell density and colorimetric change due to ONPG were measured using the Victor3 V Plate Reader (Perkin Elmer) at 620 nm and 405 nm, respectively.

The following constructs were generated for HDG2 transactivation assays using the Matchmaker Gold system (Clontech Laboratories): pCAB127 (Gal4DB-HDG2) and pCAB128 (Gal4DB-HDG2-SRDX). Gal4DB-SCRM and Gal4DB-MUTE (Kanaoka *et al.*, 2008) were used as positive and negative controls of activation, respectively. These constructs were transformed into yeast strain AH109 according to manufacturer's protocols. Activation was measured by

testing for ONPG (Sigma) as a substrate. See Tables 3-1 and 3-2 for details of plasmid construction and oligo DNA sequences used.

Dual luciferase transactivation assay *in planta*

The following constructs were generated and used in transactivation assays in planta. pCS001 (pGREEN-0800-LUC), pCS003 (pGREEN-0800-LUC-*TMMpro*), and pCS004 (pGREEN-0800-LUC-*MUTEpro*). *Agrobacterium* carrying both effector and reporter constructs were infiltrated into 4-5 week old *N. benthamiana* as described in Hellens et al., (2005). Five to seven days after infiltration, firefly luciferase (LUC) and *Renilla* luciferase (REN) were assayed using dual luciferase reagents (Promega) as previously described (Li et al., 2010). LUC and REN activity were measured using a Victor3 V Plate Reader. At least three biological replicates were measured for each sample. See Tables 3-1 and 3-2 for details about plasmid construction and oligo DNA sequences used.

Statistical analysis of promoter motifs

Enrichment of specific promoter motifs in epidermally expressed genes regulating stomatal development (*EPF1*, *EPF2*, *ERECTA*, *ERL1*, *ERL2*, *TMM*, *SPCH*, *MUTE*, *FAMA*, *SCRM/ICE1*, *SCRM2*, *HDG2*, *AtML1*, *POLAR*, *BASL*, *MYB88*, *FLP*, and *SDD1*) was tested using the ATHENA database (O'Connor *et al.*, 2005). Up to 3 kb upstream of each gene was considered, with a cutoff at adjacent genes.

Molecular phylogenetic analyses

Full-length amino acid sequences of 16 AtHD-ZIP IV members, together with five AtHD-ZIP III members as outgroup, were aligned in MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/). The model of evolution for the

data set was determined by ProtTest version 2.4 (Abascal *et al.*, 2005). The model selected under the Akaike Information Criterion (Akaike, 1974) was JTT+I+G+F.

Bayesian analyses were conducted in MrBayes version 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) via the CIPRES Science Gateway version 3.1 (Miller *et al.*, 2010) under the default priors. Three independent Markov Chain Monte Carlo (Yang and Rannala, 1997) analyses of one million generations were run under the default settings. Convergence was determined when the average standard deviation of split frequencies remained less than 0.01. The first 40% of trees was discarded before convergence. The remaining trees from each run were pooled to construct a 50% majority rule consensus tree and to obtain posterior probabilities. Graphical molecular phylogenetic tree was drawn using FigTree version 1.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Results

HDG2 is a meristemoid-expressed nuclear protein

We previously performed a comparative transcriptomic analysis of seedlings specifically enriched in particular stomatal cell state: *spch,* in which epidermis is devoid of any stomatal-lineage cells (thus pavement cells only), *scrm-D mute*, which produces an epidermis solely composed of meristemoids and sister cells (known as stomatal lineage ground cells, SLGCs), and *scrm-D*, which produces a stomata-only epidermis (Pillitteri *et al.*, 2011). Through that analysis we identified *HDG2,* a member of the HD-ZIP IV family, as a meristemoid-enriched gene. *HDG2* shows strikingly high expression in the meristemoid-enriched population (2069.7±200.8, mean±s.e.m.) with over 18-fold, 20-fold, and

10-fold enrichment over wild type (114.6±4.76), *spch* (102.2±18.3) and *scrm-D* (200.5±9.02), respectively (Figure 3-1A). Quantitative RT-PCR analysis confirmed specific and high *HDG2* expression in *scrm-D mute*, with over 31-fold, 62-fold, and 18-fold expression compared to wild type, *spch,* and *scrm-D*, respectively (Figure 3-1A).

We next examined the cell-type expression patterns of HDG2 in developing cotyledon and leaf epidermis. *HDG2* transcriptional reporter (*HDG2pro::nls-3xGFP*) showed strong GFP signals in meristemoids and SLGCs (Fig 1B). This is consistent with a previous report on *HDG2pro::GUS* (Nakamura *et al.,* 2006). Furthermore, the HDG2 translational reporter (*HDG2pro::HDG2-GFP*) accumulated in the nuclei of meristemoids and SLGCs (Figure 3-1C). The nuclear localization of HDG2-GFP accords with its predicted structure as an HD-ZIP IV transcription factor.

Our molecular phylogenetic analysis of 16 AtHD-ZIP IV family members using Bayesian algorithms shows, similar to the neighbor-joining analysis by Nakamura et al. (2006), that *HDG2* is most closely related to *AtML1* and *PDF2*, two well-studied L1-layer (protoderm)-specific genes (Abe *et al.,* 2003; Lu *et al.,* 1996; Figure 3-S1). To exclude the possibility that high expression of *HDG2* in the meristemoid-enriched population may simply be reflecting a protodermal character in the actively dividing *scrm-D mute* epidermis, we further surveyed the expression of *AtML1* and *PDF2* in stomatal cell-state mutants. Unlike *HDG2*, both *AtML1* and *PDF2* showed moderately high expression in all genetic backgrounds (Figure 3-1D). The expression levels of *AtML1* and *PDF2* were only 1.5-fold and 1.4-fold higher in *scrm-D mute* compared to wild type, which are significantly less than the >18-fold increase in *HDG2* (Figure 3-1D). Combined,

these results highlight the difference in *HDG2* expression from that of *AtML1* and *PDF2* and establish HDG2 as a stomatal-lineage-expressed nuclear-localized protein.

HDG2 overexpression confers differentiation of ectopic stomata in internal mesophyll tissues

To understand the function of *HDG2*, we first ectopically overexpressed *HDG2*. *CaMV35Spro::HDG2* plants were infertile (data not shown). For this reason we subsequently used an induction system (Zuo *et al.,* 2000), which specifically induced the transgenes after estradiol application (Figures 3-S3, 3-S4). Both *CaMV35Spro::HDG2* and inducible *HDG2* seedlings develop narrow, dark green leaves, suggesting that *HDG2* ectopic overexpression (*HDG2-OX*) affects leaf growth (Figure 3-S3 and data not shown).

In situ hybridization analysis of total *HDG2* transcripts showed strong expression in small epidermal cells (Figure 3-2A), confirming the meristemoid-enriched expression seen in *scrm-D mute* seedlings as well as in *HDG2* GFP reporters (Figure 3-1). In contrast, induced *HDG2-OX* resulted in strong signals through internal cotyledon and leaf tissues (Figure 3-2C). Control sense probes showed no signal., indicating that the probes are highly specific (Figure 3-2B,D).

Strikingly, histological sectioning revealed that *HDG2-OX* confers differentiation of stomata in internal tissues of cotyledons and true leaves. These internal stomata are correctly shaped: round and mirror-symmetric with paired GCs (Figure 3-2E-H). Furthermore, the internal stomata express mature GC marker E994 (Figure 3-2I), indicating that they possess molecular characteristics of stomata. The internal stomata phenotype was consistently observed in multiple T1 *CaMV35Spro::HDG2* seedlings and in all independent inducible

HDG2-OX lines analyzed in subsequent T2-T3 generations (Figures 3-2, 3-S3, 3-S4).

To unravel whether *HDG2* can activate earlier steps of stomatal differentiation, we next examined whether *HDG2-OX* triggers ectopic expression of earlier stomatal cell-lineage markers. As shown in Figure 3-2J and K, a sub-population of *HDG2-OX* mesophyll cells express *TMMpro::GUS-GFP*, which is normally expressed in early stomatal-lineage cells, such as meristemoids and SLGCs (Nadeau and Sack, 2002). The TMM-negative internal cells are round and have numerous chloroplasts (Figure 3-2J,K) indicating that they likely retain mesophyll cell identity. We further detected the expression of *SPCHpro::GUS*, the earliest marker of stomatal cell lineages, in the mesophyll tissue (Figure 3-2L). Based on these observations, we conclude that *HDG2* is sufficient to induce ectopic stomatal identities.

Ectopic *HDG2* overexpression can confer multiple epidermal layers and ectopic expression of epidermal markers

In some instance, we observed formation of multiple epidermal layers in *HDG2-OX* cotyledons and leaves (Figure 3-2F-H). This prompted us to test if *HDG2* overexpression confers ectopic epidermal fate. Indeed, the L1-layer marker *AtML1pro::GUS* was occasionally detected in the internal tissues (Figure 3-3A,B). We next examined the expression of the L1-layer-specific gene *FIDDLEHEAD* (*FDH*: Pruitt *et al.*, 2000) using *in situ* hybridization (Figure 3-3C-F). As expected, strong *FDH* signals are detected in the epidermis of both wild-type and *HDG2*-OX seedlings (Figure 3-3 C,E); however, it was inconclusive whether *FDH* is ectopically expressed in the internal tissues of *HDG2-OX* seedlings (Figure 3-3E). We therefore used an endoreduplicating epidermal cell

marker (*At5g17710pro::nuc-3xVENUS*; Roeder et al., 2012) to address this question (Figure 3-3G-I). The marker was originally identified for sepal epidermal 'giant' cells, but it also exhibits expression in large pavement cells in cotyledons and rosette leaves (Figure 3-3G, and data not shown). The *At5g17710pro::nuc-3xVENUS* signal was detected in occasional subepidermal nuclei in *HDG2-OX* seedlings (Figure 3-3I). On the contrary, we did not detect clear ectopic expression of trichome marker *GL2pro::GUS* or trichome differentiation in internal tissues in *HDG2-OX* (data not shown). These results indicate that, in addition to stomatal differentiation, ectopic *HDG2* could induce ectopic epidermal identities.

Optical sectioning of intact cotyledons reveals architecture of ectopic internal stomata

To conclusively demonstrate that internal stomata were not artifacts of tissue sectioning, we took advantage of the mPS-PI staining protocol (Truernit *et al.*, 2008). This technique enabled us to optically section through whole undamaged cotyledons to document the exact locations of internal stomata (Figure 3-4; Supplementary Movies 3-1, 3-2, 3-3). As expected, optical scanning of wild-type cotyledons showed characteristic layers of adaxial epidermis, palisade mesophyll, midvein, spongy mesophyll, and abaxial epidermis (Figure 3-4A,B top i-v and Supplementary Movie 3-S1). In contrast, optical scanning of *HDG2-OX* cotyledons revealed internal stomata in both palisade and spongy mesophyll layers (Figure 3-4B middle ii-iv, and Supplementary Movie 3-S2). The palisade mesophyll layer was thicker and tightly packed with smaller mesophyll cells (Supplementary Movie 3-S2).

Because *HDG2-OX* induces ectopic *AtML1* expression in mesophyll

tissues (Figure 3-3B) and because *HDG2* and *AtML1* are closely related, we tested whether they confer similar ectopic effects. Since there is no previous report on *AtML1* overexpression studies, we generated an estradiol-inducible *AtML1* overexpressor (*AtML1-OX*) and analyzed the phenotype of six independent transgenic lines (Figures 3-4B, 3-S3, 3-S4). Like *HDG2-OX*, *AtML1-OX* seedlings from all lines developed cotyledons with striking, internal stomatal in mesophyll tissues (Figure 3-4B bottom ii-iv; Supplementary Movie 3-S3). These results unambiguously demonstrate that ectopic *HDG2* overexpression results in the differentiation of stomata in internal tissues and suggest that *HDG2* and *AtML1* can act interchangeably for ectopic stomatal differentiation.

Loss of function in *HDG2* results in delayed progression of stomatal development, enhanced by additional loss of *AtML1*

Our stomatal cell-state-specific transcriptome analysis highlighted expression differences between *HDG2* and *AtML1*. *HDG2* is highly specific to the meristemoid-enriched population (*scrm-D mute*) while *AtML1* is expressed throughout the epidermal mutant genotypes (*spch, scrm-D mute*, and *scrm-D*) (Figure 3-1D). Nevertheless, both *HDG2-OX* and *AtML1-OX* triggered ectopic, internal stomatal differentiation (Figures 3-3, 3-4). This raises the question of whether the endogenous *HDG2* plays a specific role in stomatal differentiation. To address this question, we investigated the *hdg2* loss-of-function phenotype. A previous report on the systematic characterization of the *Arabidopsis* HD-ZIP IV family members did not reveal any phenotype associated with the *hdg2* T-DNA knockout mutation (Nakamura *et al.*, 2006). We therefore predicted that *hdg2* phenotype, if any, would be less obvious.

Careful examination of developing (10-day-old) *hdg2* cotyledons revealed

increased numbers of meristemoids, which are not commonly seen in wild type at this stage (Figure 3-5). For quantitative analysis of stomatal phenotypes, we used three T-DNA insertion alleles of *HDG2*: *hdg2-2*, *hdg2-3*, and *hdg2-4* (Figure 3-S2). Extensive RT-PCR analysis detected aberrant, fused *T-DNA-HDG2* transcripts of partial size, indicating that none of these alleles produce functional *HDG2* transcripts (Figure 3-S2D). Quantitative analysis showed that all three *hdg2* alleles exhibit statistically significant reduction in stomatal index (SI) (Figure 3-5A). In contrast, number of meristemoids per epidermal cell, hereafter defined as meristemoid index (MI), was significantly higher in *hdg2-2, 2-3,* and *2-4* than in wild type (Figure 3-5A). Since total numbers of stomata and stomatal precursors, defined here as stomatal lineage index (SLI), remained unchanged from wild type (Figure 3-5A), *hdg2* seedlings are delayed in progression of stomatal differentiation rather than producing more meristemoids. All three *hdg2* alleles showed statistically indistinguishable phenotypes (Tukey's HSD test, non-significant), strongly supporting that the phenotype is caused by loss of function in *HDG2*.

We next examined the possible involvement of *AtML1* and *PDF2* in stomatal development. As shown in Figure 3-5B-D, G, and H; SI, MI, and SLI of *atml1* and *pdf2* cotyledon abaxial epidermis were not significantly different from those that of wild type. Interestingly, introduction of *atml1* mutation to *hdg2* dramatically enhanced MI and reduced SI. On the contrary, *pdf2* had no such effects (Figure 3-5B-D).

To understand the fate of delayed stomatal precursors we observed fully mature, senescing cotyledon epidermis (30 dpg). In *hdg2,* some arrested stomatal precursors surrounded by excessive SLGCs were observed (Figure 3-5K).

Stomatal precursors were never observed in wild type cotyledons at this age (not shown). Rarely yet consistently (approximately 1-2 per cotyledon), some stomatal complexes only resulted in singular GCs (Figure 3-5 L-N).

Like cotyledons, rosette leaf epidermis of *hdg2* exhibits statistically significant increase in MI with concomitant decrease in SI (Figure 3-S5 A-B). In *hdg2* rosette leaf abaxial epidermis, meristemoids occasionally failed to differentiate into stomata (Figure 3-S5D). The phenotype was exaggerated by the additional *atml1* mutation, and excessive numbers of SLGCs were seen in the *hdg2 atml1* rosette leaf epidermis, implying an extended asymmetric cell division state due to delayed differentiation (Figure 3-S5D).

While the cotyledon epidermis is formed during embryogenesis, the rosette leaf epidermis is derived from the L1 layer of the shoot apical meristem. Our results suggest that *HDG2* promotes progression of stomatal differentiation in both types of organ, and that *AtML1*, but not *PDF2*, shares a redundant role in controlling the timing of meristemoid differentiation.

HDG2 is a transcriptional activator that binds to L1 and HAHR1 boxes

To gain insight into the mode of action of HDG2, we first tested whether HDG2 activates transcription, and if so, whether forceful conversion of HDG2 into a transcriptional repressor mimics its loss-of-function effects (Figure 3-6). HDG2 shares a sequence similarity with the known activation domain of Maize HD-ZIP IV protein OCL1 (Depege-Fargeix *et al.*, 2011) (Figure 3-6A,B). Indeed, HDG2 fused to Gal4 DNA-binding domain strongly transactivated the reporter in yeast at an equivalent level to known transcriptional activator SCRM/ICE1 (Figure 3-6C), indicating that HDG2 is a transcriptional activator.

C-terminal addition of the 12-amino-acid repressor domain SRDX is

known to convert plant TFs to dominant repressors (Hiratsu *et al.,* 2003). The addition of the SRDX sequence significantly decreased HDG2 reporter transactivation in yeast (Figure 3-6C). To test if HDG2-SRDX mimics loss-of-function of *HDG2* in vivo, we generated transgenic *Arabidopsis* carrying inducible *HDG2-SRDX* (Figure 3-6D; Figure 3-S3). The *HDG2-SRDX* induction conferred delayed progression of meristemoids and patches of SLGCs (Figure 3-6D). Expression of stomatal-lineage markers *TMMpro::GUS* and *EPF2::erGFP* supports the stomatal-lineage origins of these aberrant cells (Figure 3-6D).

HD-ZIP IV-family proteins, including AtML1 and PDF2, show in vitro binding to the pseudopalindromic sequences known as the L1 box (5'-TAAATG(C/T)A-3') and HAHR1 box (5'-CATT(A/T)AATG-3') (Abe *et al.,* 2003; Abe *et al.,* 2001; Tron *et al.,* 2001). Previous genome-wide study of AtHD-ZIP proteins did not include DNA-binding assays of HDG2 (Nakamura *et al.,* 2006). We therefore performed yeast one-hybrid analysis to test if HDG2 binds to these DNA sequences (Figure 3-6E,F). The reporter assay shows that HDG2 binds to both L1-A (5'-TAAATGCA-3') and L1-T (5'-TAAATGTA-3') as strongly as AtML1 (Figure 3-6E). HDG2, AtML1, and PDF2 all induced strong LacZ reporter activity when fused with the HAHR1-T (5'-CATTTAATG-3') and HAHR1-A (5'-CATTAAATG-3') boxes, with AtML1 inducing the highest HAHR1-reporter activity (Figure 3-6F). These three HD-ZIP IV proteins did not induce reporter activity when key DNA residues of the L1 and HAHR1 boxes were mutated (L1-m and HAHR1-m; Figure 3-6E,F), confirming the sequence specificity of binding. Similarly, a control bHLH protein, SCRM/ICE1, which has been shown to bind in vitro to an E-box element (CANNTG) (Chinnusamy *et al.,* 2003), did not induce any reporter activity (Figure 3-6E,F). These results suggest that HDG2

acts as a transcriptional activator that binds to L1 and HAHR1 boxes. The similar in vitro binding of HDG2 and AtML1 to these cis-regulatory elements further emphasizes the possibility that they share common downstream targets.

HDG2 transactivates promoters of stomatal development genes *in planta*

To address the molecular connection of HDG2, an L1/HAHR1-binding transcriptional activator, with known regulators of stomatal development, we first performed a bioinformatic analysis and surveyed cis-regulatory elements within the promoters of genes regulating stomatal development. Among the eighteen genes surveyed, a significant ($p=0.003$) enrichment of L1-box motifs in their promoter regions was revealed (Table 3-4). Indeed, the p-value for significance of enrichment for the L1 box was the lowest among the 47 motifs analyzed (Table 3-4).

To demonstrate that HDG2 activates transcription of genes regulating stomatal development, we generated reporter luciferase constructs containing native promoters of *TMM* and *MUTE*, which possess one and two L1-boxes, respectively, within their proximal regions (Figure 3-7A). These reporter and effector constructs were co-expressed in *N. benthamiana*. High luciferase activity was observed in a combination of *HDG2* and *TMMpro::LUC*, and even higher activity with *MUTEpro::LUC*, consistent with the number of L1 boxes present (Figure 3-7B). Therefore, HDG2 is capable of directly transactivating *TMM* and *MUTE in planta*, supporting the possibility that HDG2 promotes stomatal differentiation by binding to the promoters and activating the expression of genes regulating stomatal development.

Discussion

Our work reveals the unique role of the AtHD-ZIP IV protein HDG2 as a transcriptional activator promoting stomatal differentiation. HDG2 is necessary for progression of stomatal cell lineages and, most strikingly, sufficient to drive stomatal differentiation in internal mesophyll tissues. The internal stomata in *HDG2-OX* are normal in appearance, and thus different from *FAMA-OX*, which confers abnormally-shaped mesophyll cells expressing mature GC markers (Ohashi-Ito *et al.*, 2006). Our data support the idea that *HDG2-OX* upregulates whole sets of gene regulatory cascades leading to stomatal differentiation. Since *HDG2* highly accumulates in actively proliferating stomatal cell lineages, the stomatal bHLH proteins may promote HDG2, which in turn promotes stomatal bHLHs for reinforcing stomatal differentiation. The transactivation of TMM and MUTE promoters *in planta* by HDG2 (Figure 3-7) strongly supports this idea.

In addition, *HDG2-OX* confers multiple epidermal layers, indicating that it also specifies protodermal identity. In vitro HD binding site selection experiments have revealed 5'-GCATTATTTTGC-3' as consensus sequence for four HD-ZIP IVs: HDG7, HDG9, AtML1, and PDF2 (Nakamura *et al.*, 2006). The well-described L1 and HAHR1 boxes overlap with this consensus sequence (Abe *et al.*, 2001; Tron *et al.*, 2001). While this experiment has not been performed for HDG2, it is most likely that HDG2 and AtML1, when overexpressed, bind to common in vivo target sequences for the following reasons. First, both HDG2 and AtML1 bind to L1 and HAHR1 boxes to similar extents in yeast (Figure 3-6E,F). Second, four amino-acid residues predicted to serve as DNA-sequence-specific contact surface (F47, Q50, N51, and T54) within the HD domain are absolutely conserved in HDG2 and AtML1 (Gehring *et al.*, 1994; Nakamura *et al.*,

2006). Third, *AtML1-OX* confers ectopic internal stomatal differentiation, phenocopying *HDG2-OX* (Figure 3-4B). Future genome-wide identification and comparison of their target genes will address the extent of unique vs. shared components of gene regulatory networks specified by HDG2 and AtML1.

Auto-regulation of transcription factors is crucial for robust specification of cell fate. *AtML1* auto-activates by directly binding to the L1 box in its own promoter to reinforce protodermal identity (Abe *et al.,* 2003; Takada and Jürgens, 2007). That *HDG2-OX* triggers ectopic *AtML1* expression (Figure 3-3B) is consistent with the presence of such an auto-activation loop, in this case initially triggered by the ectopic *HDG2*. The *HDG2* promoter also possesses L1 boxes (Nakamura *et al.,* 2006), and its auto-activation may contribute to robust progression of stomatal development. Specification of the protoderm (L1 layer) in *Arabidopsis* involves cell-cell signaling mediated by receptor-like kinases ARABIDOPSIS CRINKLY4 (ACR4) and ABNORMAL LEAF SHAPE2 (ALE2), a putative cysteine peptidase DEFECTIVE KERNEL1 (AtDEK1) and a putative subtilase ALE1 (Becraft *et al.,* 1996; Cao *et al.,* 2005; Johnson *et al.,* 2005; Tanaka *et al.,* 2001; Tanaka *et al.,* 2007). The signaling pathways appear to act upstream of AtML1 and PDF2, which in turn may promote expression of AtDEK1 and ALE2 as part of a positive-feedback loop (Javelle *et al.,* 2011b). While these signaling components are required for normal epidermal differentiation, none of them has been shown to drive supernumerary epidermal differentiation. Our work in addition establishes AtML1 as a master regulatory transcription factor both necessary and sufficient for protodermal identity. It is surprising that *AtML1* overexpression phenotype had not been reported since its identification in 1996 (Lu *et al.,* 1996). *PDF2-OX* confers a later-flowering phenotype, likely mimicking

the effects of a distantly related *AtHD-ZIP IV*, *FWA* (Abe *et al.*, 2003); thus HDG2 may share more common targets with AtML1 than PDF2.

While our ectopic overexpression study revealed the equivalence of *HDG2* and *AtML1*, expression patterns and loss-of-function phenotypes distinguish their endogenous functions and provide evidence for *HDG2*'s major role in stomatal development. We further uncovered an unexpected redundant role for *AtML1* in promoting stomatal differentiation, which was not shared by *PDF2*. Thus, these HD-ZIP IVs have unique yet partially overlapping functions in specifying multiple aspects of epidermal development. Bioinformatic co-expression network ATTED II (Obayashi *et al.*, 2009) places *HDG2* at the 'hub' connecting stomatal and protodermal regulators: *HDG2* clusters together with *SPCH*, *TMM*, and *SCRM*/*ICE1*, with a direct connection to *SCRM*/*ICE1* (Figure 3-S6). At the same time, *HDG2* is directly connected to *AtML1* and clusters together with *PDF2*, *HDG1*, and *ERECTA*, a receptor kinase which inhibits stomatal-lineage initiation (Shpak *et al.*, 2005, Figure 3-S6).

The meristemoid-to-GMC transition represents a key step from proliferation to differentiation within the stomatal lineages, and is specified by *MUTE* and *SCRM*. Loss of *HDG2* may delay the onset of *MUTE* or *SCRM* expression or deregulate their functions in a small fraction of meristemoids. The observation that *MUTE* can be directly activated by HDG2 (Figure 3-7) supports this hypothesis. *scrm* loss-of-function mutant produces aberrant stomatal complexes at low penetrance (Kanaoka et al. 2008), a phenotype shared by *hdg2* (Figure 3-5L-N). The very low penetrance of *hdg2* stomatal phenotype may be due to functional redundancy with the remaining members of *AtHD-ZIP IVs*, or alternatively, that HDG2 indirectly influences stomatal development through

promoting cell differentiation.

Proper stomatal differentiation requires an intricate balance of cell proliferation and differentiation through coordinated activities of cell-cycle regulators and transcription factors. Among plant cell-cycle regulators, A-type cyclin CycA2;3 and its partner CDKB1;1 are required for GMC divisions, and their losses of function lead to singular GC phenotype (Boudolf *et al.*, 2004; Vanneste *et al.*, 2011). Myb transcription factors FOUR LIPS (FLP) and Myb88 directly bind to the promoters of *CycA2;3* and *CDKB1;1* and repress their expression (Vanneste *et al.*, 2011; Xie *et al.*, 2010). Due to the very low-penetrant *hdg2* stomatal phenotype, it is difficult to draw clear connections between *HDG2* and these regulators. Nevertheless it is tempting to speculate on a potential role of *HDG2* in cell proliferation and differentiation. Interestingly, in addition to meristemoids, *HDG2-GFP* is highly expressed in developing trichomes and a subset of pavement cells, where active endoreduplication occurs (Figure 3-S7; Hulskamp, 2004). *hdg2* trichomes do not have growth/morphological phenotypes, but exhibit altered cuticle textures (Marks *et al.*, 2009). Further studies, such as cell-type-specific manipulation of *HDG2* activity and identification of its targets, may provide a direct link among HDG2-mediated epidermal identity, cell proliferation, and stomatal differentiation.

Supplemental Movie Legends

Movie 3-1. Complete optical sectioning of a wild-type 10-day-old cotyledon. Z stacks were made from mPS-PI stained confocal optical sections in 1 μm intervals.

Movie 3-2. Complete optical sectioning of a 10-day-old cotyledon from *HDG2*

induced ectopic overexpression. Z stacks were made from mPS-PI stained confocal optical sections in 1 μm intervals.

Movie 3-3. Complete optical sectioning of a 10-day-old cotyledon from *AtML1* induced ectopic overexpression. Z stacks were made from mPS-PI stained confocal optical sections in 1 μm intervals.

Acknowledgements

We thank Lynn Pillitteri, Hongya Gu, Nam-Hai Chua, Tsuyoshi Nakagawa, John Schiefelbein, Xing-Wang Deng, Adrienne Roeder, and Naoyuki Uchida for kindly providing us plasmids and reporter lines; Valerie Soza for molecular phylogenetic analysis; Ya-Chen Lin for plant care; Janneke Hille Ris Lambers for assistance in statistical analysis; Wim Lewis for assistance with cell counting; Shinobu Takada for communicating results; and Lynn Pillitteri and Jin Suk Lee for comments. The work was initially supported by the University of Washington Royalty Research Fund and subsequently by the JST-PRESTO award and NSF (MCB-0855659) to K.U.T., JSPS KAKENHI grant number 23012021, 24113510 and 4570045 to M.M.K, and 23012014 to Y.S. K.M.P. was an NSF Graduate Research Fellow and K.U.T. is an HHMI-GBMF investigator.

Author Contributions

K.M.P. and K.U.T. conceived and designed the research; K.M.P., C.S., C.A.B, R.J.H, and K.U.T. performed the research except that M.M.K., M.O., and Y.S. did *in situ* hybridization; K.M.P., C.S., C.A.B, M.M.K, and K.U.T. generated research materials and reagents; K.M.P., R.J.H. and K.U.T. analyzed the data; K.U.T. wrote the manuscript.

References

Abascal., F., Zardoya, R. and Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21, 2104–2105.

Abe, M., Katsumata, H., Komeda, Y. and Takahashi, T. (2003) Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in *Arabidopsis*. *Development* 130, 635-643.

Abe, M., Takahashi, T. and Komeda, Y. (2001) Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *Plant J* 26, 487-494.

Akaike, H. (1974) A new look at the statistical model identification. I*EEE Transactions on Automatic Control AC* 19, 716–723.

Becraft, P. W., Stinard, P. S. and McCarthy, D. R. (1996) *CRINKLY4*, a TNFR-like receptor kinase invloved in maize epidermal differentiation. *Science* 273, 1406-1409.

Boudolf, V., Barroco, R., Engler Jde, A., Verkest, A., Beeckman, T., Naudts, M., Inze, D. and De Veylder, L. (2004) B1-type cyclin-dependent kinases are essential for the formation of stomatal complexes in *Arabidopsis thaliana*. *Plant Cell* 16, 945-955.

Cao, X., Li, K., Suh, S. G., Guo, T. and Becraft, P. W. (2005) Molecular analysis of the CRINKLY4 gene family in *Arabidopsis thaliana*. *Planta* 220, 645-657.

Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B. H., Hong, X., Agarwal., M. and Zhu, J. K. (2003) ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. *Genes Dev* 17, 1043-1054.

Depege-Fargeix, N., Javelle, M., Chambrier, P., Frangne, N., Gerentes, D., Perez, P., Rogowsky, P. M. and Vernoud, V. (2011) Functional characterization of the HD-ZIP IV transcription factor OCL1 from maize. *J Exp Bot* 62, 293-305.

Dong, J. and Bergmann, D. C. (2010) Stomatal patterning and development. *Curr Top Dev Biol* 91, 267-297.

Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994) Homeodomain-DNA recognition. *Cell* 78, 211-223.

Guseman, J. M., Lee, J. S., Bogenschutz, N. L., Peterson, K. M., Virata, R. E., Xie, B., Kanaoka, M. M., Hong, Z. and Torii, K. U. (2010) Dysregulation of cell-to-cell connectivity and stomatal patterning by loss-of-function mutation in *Arabidopsis* CHORUS (GLUCAN SYNTHASE-LIKE 8). *Development* 137, 1731-1741.

Hara, K., Yokoo, T., Kajita, R., Onishi, T., Yahata, S., Peterson, K. M., Torii, K. U. and Kakimoto, T. (2009) Epidermal cell density is auto-regulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR2 in *Arabidopsis* leaves. *Plant Cell Physiol* 50, 1019-1031.

Hiratsu, K., Matsui, K., Koyama, T. and Ohme-Takagi, M. (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *Plant J* 34, 733-739.

Hellens, R.P., Allan, A.C., Friel, E.N., Bolitho, K., Grafton, K., Templeton, M.D., Karunairetnam, S., Gleave, A.P., and Laing, W.A. (2005) Transient

expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* 1, 13.
Huelsenbeck, J.P. and Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
Hulskamp, M. (2004) Plant trichomes: a model for cell differentiation. *Nat Rev Mol Cell Biol* 5, 471-480.
Ingram, G. C., Boisnard-Lorig, C., Dumas, C. and Rogowsky, P. M. (2000) Expression patterns of genes encoding HD-ZipIV homeo domain proteins define specific domains in maize embryos and meristems. *Plant J* 22, 401-414.
Javelle, M., Klein-Cosson, C., Vernoud, V., Boltz, V., Maher, C., Timmermans, M., Depege-Fargeix, N. and Rogowsky, P. M. (2011a) Genome-wide characterization of the HD-ZIP IV transcription factor family in maize: preferential expression in the epidermis. *Plant Physiol* 157, 790-803.
Javelle, M., Vernoud, V., Depege-Fargeix, N., Arnould, C., Oursel, D., Domergue, F., Sarda, X. and Rogowsky, P. M. (2010) Overexpression of the epidermis-specific homeodomain-leucine zipper IV transcription factor Outer Cell Layer1 in maize identifies target genes involved in lipid metabolism and cuticle biosynthesis. *Plant Physiol* 154, 273-286.
Javelle, M., Vernoud, V., Rogowsky, P. M. and Ingram, G. C. (2011b) Epidermis: the formation and functions of a fundamental plant tissue. *New Phytol* 189, 17-39.
Johnson, K. L., Degnan, K. A., Ross Walker, J. and Ingram, G. C. (2005) AtDEK1 is essential for specification of embryonic epidermal cell fate. *Plant J* 44, 114-127.
Kanaoka, M. M., Pillitteri, L. J., Fujii, H., Yoshida, Y., Bogenschutz, N. L., Takabayashi, J., Zhu, J. K. and Torii, K. U. (2008) SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* 20, 1775-1785.
Kouchi, H. and Hata, S. (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol Gen Genet* 238, 106-119.
Lee, J. S., Kuroha, T., Hnilova, M., Khatayevich, D., Kanaoka, M. M., McAbee, J. M., Sarikaya, M., Tamerler, C. and Torii, K. U. (2012) Direct interaction of ligand-receptor pairs specifying stomatal patterning. *Genes Dev* 26, 126-136.
Li, J., Li, G., Gao, S., Martinez, C., He, G., Zhou, Z., Huang, X., Lee, J.-H., Zhang, H., Shen, Y., Wang, H., and Deng, X.-W. (2010) *Arabidopsis* transcription factor ELONGATED HYPOCOTYL5 plays a role in the feedback regulation of phytochrome A signaling. *Plant Cell* 22, 3634–3649.
Livak, K. J. and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-408.
Lu, P., Porat, R., Nadeau, J. A. and O'Neill, S. D. (1996) Identification of a meristem L1 layer-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox genes. *Plant Cell* 8, 2155-2168.
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.
Marks, M. D., Wenger, J. P., Gilding, E., Jilk, R. and Dixon, R. A. (2009) Transcriptome analysis of *Arabidopsis* wild-type and *gl3-sst sim* trichomes identifies four additional genes required for trichome development. *Mol Plant* 2, 803-822.
Masucci, J.D., Rerie, W.G., Foreman, D.R., Zhang, M., Galway, M.E., Marks, M.D., Schiefelbein, J.W. (1996) The homeobox gene *GLABRA2* is required for position-dependent cell differentiation in the root epidermis of *Arabidopsis thaliana. Development* 122, 1253-1260.
Miller, M.A., Pfeiffer, W. and Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop* (GCE), 2010. New Orleans, LA. pp. 1–8.
Mukherjee, K., Brocchieri, L. and Burglin, T. R. (2009) A comprehensive classification and evolutionary analysis of plant homeobox genes. *Mol Biol Evol* 26, 2775-2794.
Nadeau, J. A. and Sack, F. D. (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.
Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y., *et al* (2007) Improved Gateway binary vectors: high-performance vectors for creation of fusion constructs in transgenic analysis of plants. *Biosci Biotechnol Biochem* 71, 2095-2100.
Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K. T. and Takahashi, T. (2006) Characterization of the class IV homeodomain-Leucine Zipper gene family in *Arabidopsis*. *Plant Physiol* 141, 1363-1375.
O'Connor, T. R., Dyreson, C. and Wyrick, J. J. (2005) Athena: a resource for rapid visualization and systematic analysis of *Arabidopsis* promoter sequences. *Bioinformatics* 21, 4411-4413.
Obayashi, T., Hayashi, S., Saeki, M., Ohta, H. and Kinoshita, K. (2009) ATTED-II provides coexpressed gene networks for *Arabidopsis*. *Nucleic Acids Res* 37, D987-991.
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., Bogenschutz, N. L. and Torii, K. U. (2008) The bHLH protein, MUTE, controls differentiation of stomata and the hydathode pore in *Arabidopsis*. *Plant Cell Physiol* 49, 934-943.
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.
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. and Torii, K. U. (2012) Mechanisms of Stomatal Development.

Annu Rev Plant Biol 63, 591-614.
Pruitt, R.E., Vielle-Calzada, J.-P., Ploense, S.E., Grossniklaus, U. and Lolle, S.J. (2000) *FIDDLEHEAD*, a gene required to suppress epidermal cell interactions in *Arabidopsis*, encodes a putative lipid biosynthetic enzyme. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1311-1316.
Rerie, W. G., Feldmann, K. A. and Marks, M. D. (1994) The *GLABRA2* gene encodes a homeo domain protein required for normal trichome development in *Arabidopsis*. *Genes Dev* 8, 1388-1399.
Roeder, A.H.K., Cunha, A., Ohno, C.K., and Meyerowitz, E.M. (2012) Cell cycle regulates cell types in *Arabidopsis* sepal. *Development* 139, 4416-4427.
Ronquist, F. and Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
Serna, L. and Martin, C. (2006) Trichomes: different regulatory networks lead to convergent structures. *Trends Plant Sci* 11, 274-280.
Sessions, A., Weigel, D. and Yanofsky, M. F. (1999) The *Arabidopsis thaliana MERISTEM LAYER 1* promoter specifies epidermal expression in meristems and young primordia. *Plant J* 20, 259-263.
Shpak, E. D., Lakeman, M. B. and Torii, K. U. (2003) Dominant-negative receptor uncovers redundancy in the *Arabidopsis* ERECTA leucine-rich repeat receptor-like kinase signaling pathway that regulates organ shape. *Plant Cell* 15, 1095-1110.
Shpak, E. D., McAbee, J. M., Pillitteri, L. J. and Torii, K. U. (2005) Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science* 309, 290-293.
Takada, S. and Jurgens, G. (2007) Transcriptional regulation of epidermal cell fate in the *Arabidopsis* embryo. *Development* 134, 1141-1150.
Tanaka, H., Onouchi, H., Kondo, M., Hara-Nishimura, I., Nishimura, M., Machida, C. and Machida, Y. (2001) A subtilisin-like serine protease is required for epidermal surface formation in *Arabidopsis* embryos and juvenile plants. *Development* 128, 4681-4689.
Tanaka, H., Watanabe, M., Sasabe, M., Hiroe, T., Tanaka, T., Tsukaya, H., Ikezaki, M., Machida, C. and Machida, Y. (2007) Novel receptor-like kinase ALE2 controls shoot development by specifying epidermis in *Arabidopsis*. *Development* 134, 1643-1652.
Tron, A. E., Bertoncini, C. W., Palena, C. M., Chan, R. L. and Gonzalez, D. H. (2001) Combinatorial interactions of two amino acids with a single base pair define target site specificity in plant dimeric homeodomain proteins. *Nucleic Acids Res* 29, 4866-4872.
Truernit, E., Bauby, H., Dubreucq, B., Grandjean, O., Runions, J., Barthelemy, J. and Palauqui, J. C. (2008) High-resolution whole-mount imaging of three-dimensional tissue organization and gene expression enables the study of Phloem development and structure in *Arabidopsis*. *Plant Cell* 20, 1494-1503.
Uchida, N., Lee, J. S., Horst, R. J., Lai, H.-H., Kajita, R., Kakimoto, T., Tasaka, M. and Torii, K. U. (2012) Regulation of inflorescence architecture by intertissue layer ligand-receptor communication between *Proc Natl Acad Sci U S A* 109, 6337-6342.
Vanneste, S., Coppens, F., Lee, E., Donner, T. J., Xie, Z., Van Isterdael, G., Dhondt, S., De Winter, F., De Rybel, B., Vuylsteke, M., *et al* (2011) Developmental regulation of CYCA2s contributes to tissue-specific

proliferation in *Arabidopsis*. *EMBO J* 30, 3430-3441.
Wu, R., Li, S., He, S., Wassmann, F., Yu, C., Qin, G., Schreiber, L., Qu, L. J. and Gu, H. (2011) CFL1, a WW domain protein, regulates cuticle development by modulating the function of HDG1, a class IV homeodomain transcription factor, in rice and *Arabidopsis*. *Plant Cell* 23, 3392-3411.
Xie, Z., Lee, E., Lucas, J. R., Morohashi, K., Li, D., Murray, J. A., Sack, F. D. and Grotewold, E. (2010) Regulation of cell proliferation in the stomatal lineage by the *Arabidopsis* MYB FOUR LIPS via direct targeting of core cell cycle genes. *Plant Cell* 22, 2306-2321.
Yang, Z. H. and Rannala, B. (1997) Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. *Mol Biol Evol* 14, 717–724.
Zuo, J., Niu, Q. W. and Chua, N. H. (2000) Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J* 24, 265-273.

Figure 3-1.

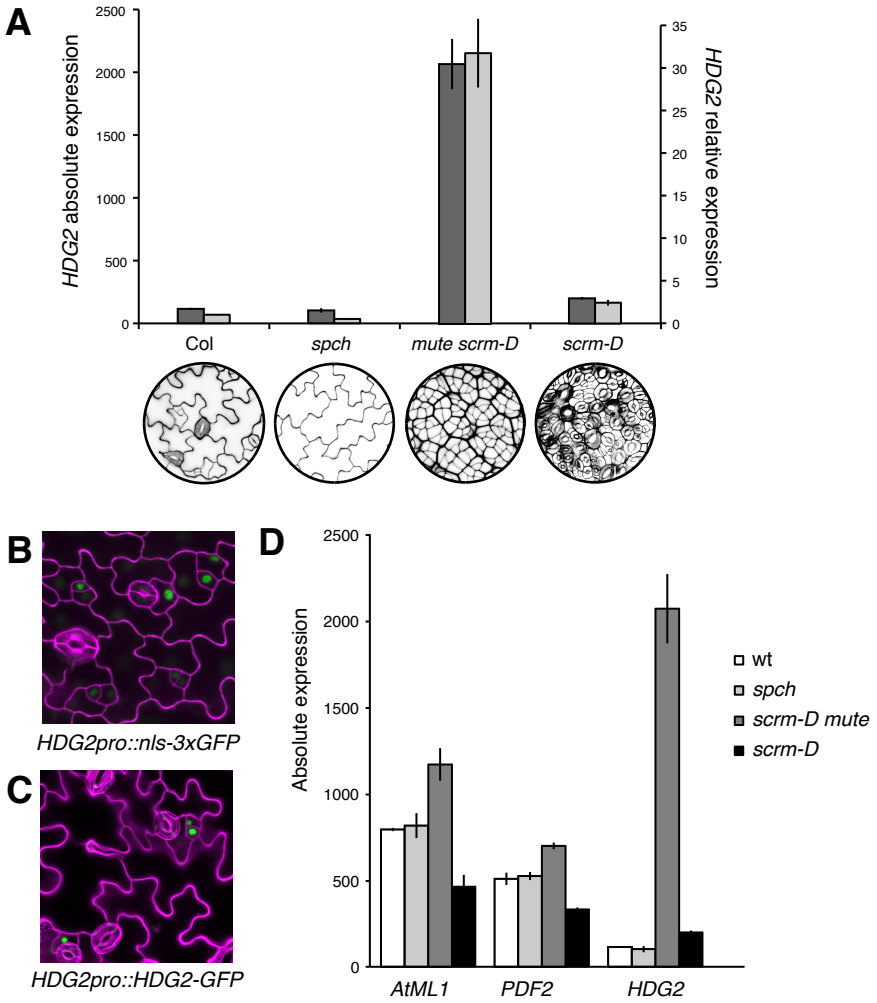


Figure 3-1. HDG2 is highly enriched in meristemoid population of stomatal cell lineages.
(A) *HDG2* absolute and relative expression levels among wild type and stomatal mutants enriched in specific epidermal cell populations. Absolute expressions are from ATH1 microarray data, and relative expressions are from qRT-PCR analysis. Bars, mean value of triplicates; error bars, s.e.m. Col, wt; *spch*, pavement-cell-only; *mute scrm-D*, overwhelmingly enriched in meristemoids; *scrm-D*, stomata-only epidermis. Below each graph are confocal images of cotyledons from corresponding genotypes. (B) Stomatal-lineage accumulation of *HDG2* transcriptional reporter (*HDG2pro::nls-3xGFP*) in seedling epidermis. (C) Stomatal-lineage accumulation of *HDG2* translational reporter (*HDG2pro::HDG2-GFP*) in 10-day-old abaxial cotyledon epidermis. Scale bars, 20 μm. (D) Expression levels of *AtML1* and *PDF2* compared to *HDG2* among wild type and stomatal mutants.

Figure 3-2.

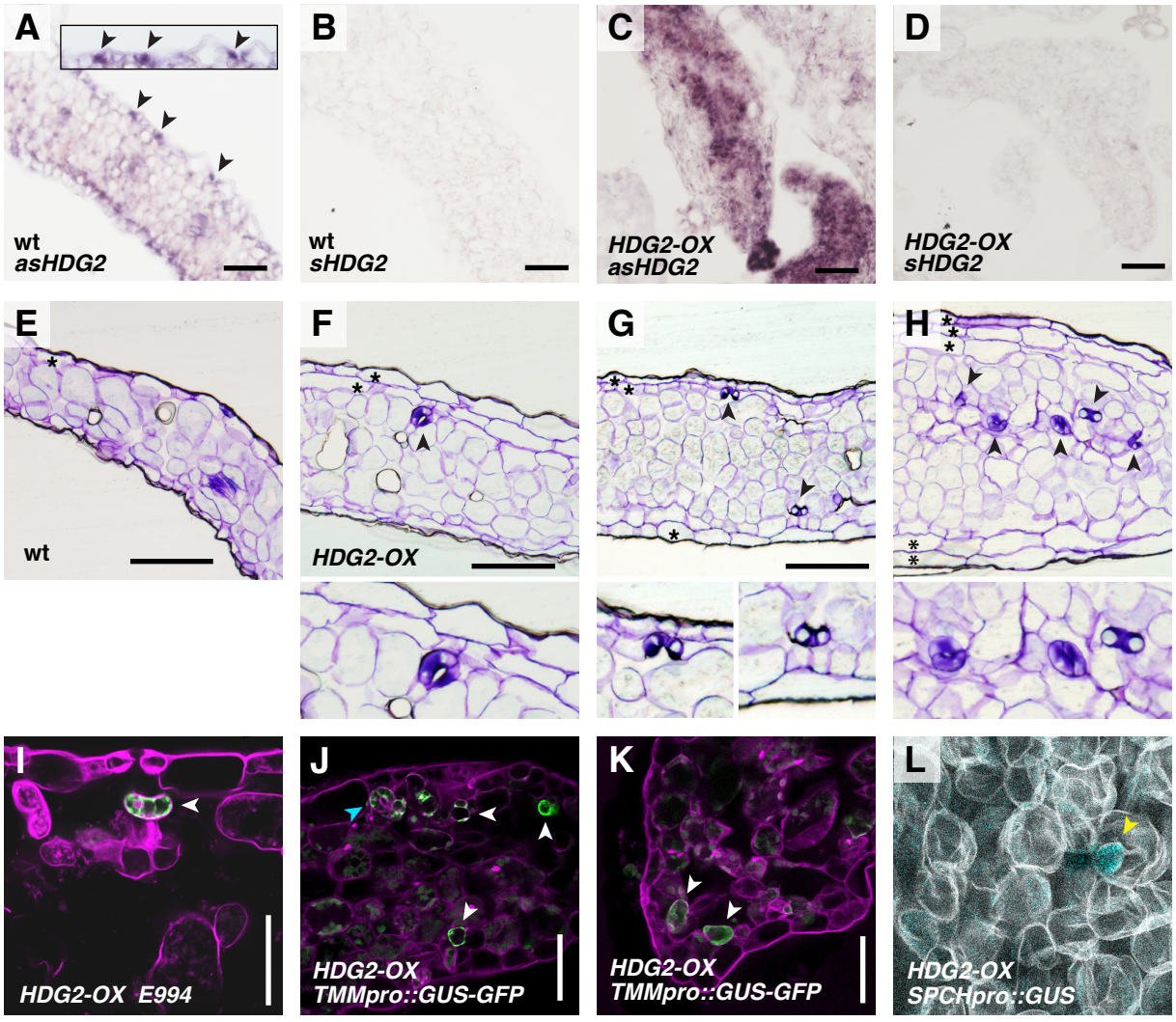


Figure 3-2. Ectopic overexpression of *HDG2* confers ectopic stomata within mesophyll tissues and formation of multiple epidermal layers.
(A-D) In situ hybridization analysis of *HDG2* expression in 10-day-old wild type (A, B) and HDG2-OX (C, D) seedlings treated with *HDG2* antisense (A, C) or sense (B, D) probes. In wild type, endogenous *HDG2* shows stomatal-lineage expression (A: arrowheads). Inset, enlarged image. Scale bars, 40 μm. (E-H) Histological cross sections of plastic-embedded 10-day-old cotyledons from wild type (E) and transgenic plants overexpressing *HDG2* (*HDG2-OX*), which are *CaMV35S::HDG2* (F-H). In *HDG2-OX*, ectopic differentiation of stomata in internal tissues (arrowheads) is evident. Bottom insets, enlarged images. Multiple epidermal layers (asterisks) occasionally formed. Images are taken under the same magnification. Scale bar, 100 μm. (I-K) Live sections of *HDG2-OX* cotyledons expressing mature guard cell marker E994 (I) and stomatal-lineage marker *TMMpro::GUS-GFP* (J, K). GFP is detected in ectopic internal stomata (J; arrowhead) and internal stomatal-lineage cells (J, K arrowheads). A mature ectopic internal stoma is also evident in panel J (cyan arrowhead). Scale bars, 20 μm. (L) Optical section of mPS-PI stained mesophyll layer from *HDG2-OX* cotyledon expressing *SPCHpro::GUS* (arrowhead).

Figure 3-3.

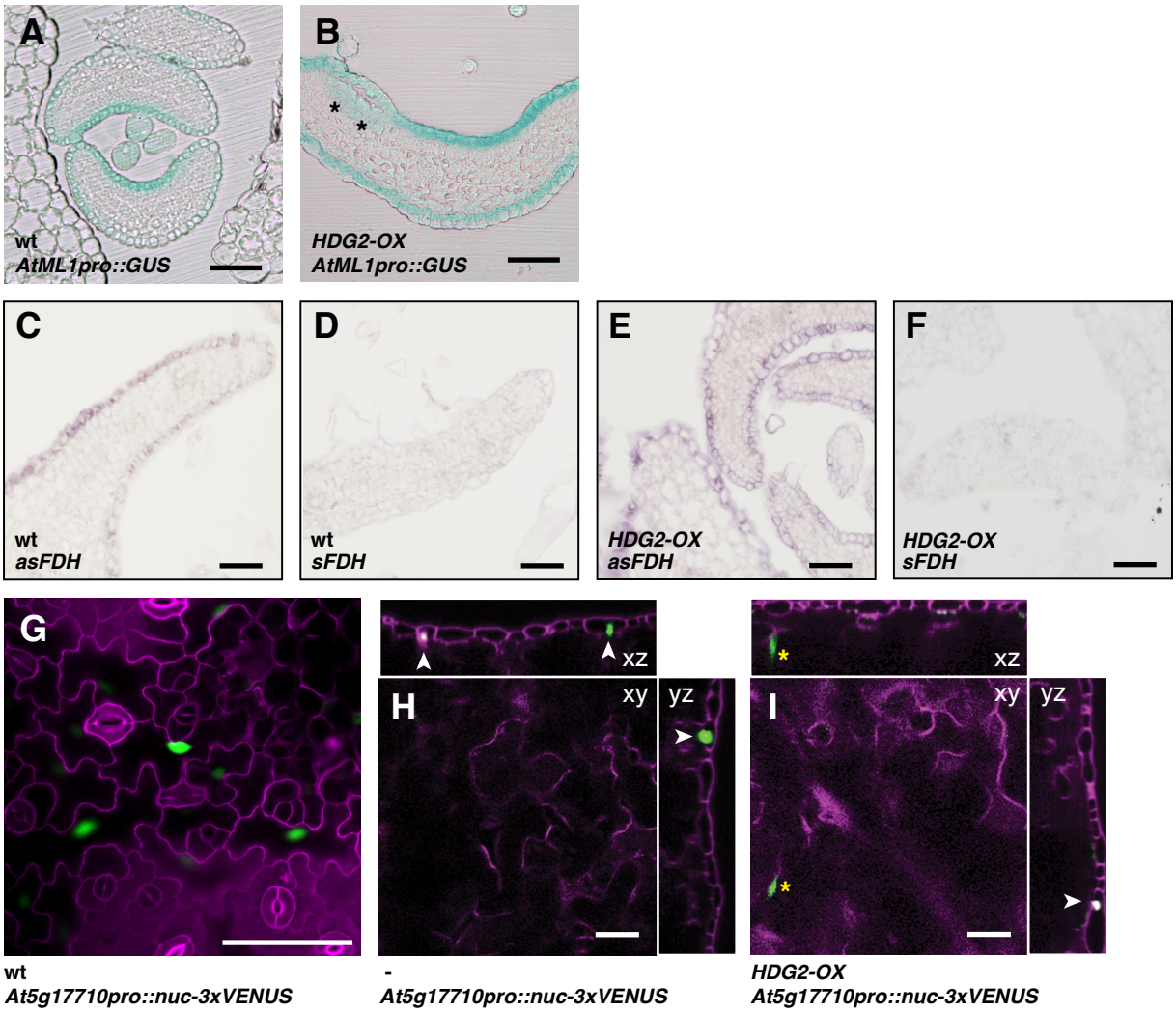


Figure 3-3. Ectopic overexpression of *HDG2* induces some epidermal reporters.
(A, B) Histological cross sections of emerging rosette leaves from 10-day-old wild type (wt: A) and *HDG2-OX* seedlings expressing *AtML1pro::GUS* (B). Ectopic GUS expression (asterisks) can be seen in (B). Scale bars, 100 μm. (C-F) In situ hybridization analysis of *FDH* expression in 10-day-old wild type (C, D) and *HDG2-OX* (E, F) seedlings treated with *FDH* antisense (C, E) or sense (D, F) probes. Scale bars, 40 μm. (G) *At5g17710pro::nuc-3xVENUS* in wild-type rosette leaf abaxial epidermis showing specific signals in large pavement cells. Scale bar, 50 μm. (H, I) Confocal optical sections (xy, xz, and yz axis) of 10-day-old *HDG2-OX* cotyledon expressing *At5g17710pro::nuc-3xVenus*, without (H) or with (I) *HDG2* induction. Asterisk, GFP in subepidermal mesophyll nucleus. Arrowheads, GFP epidermal nuclei (co-stained with propidium iodide). Scale bars, 50 μm.

Figure 3-4.

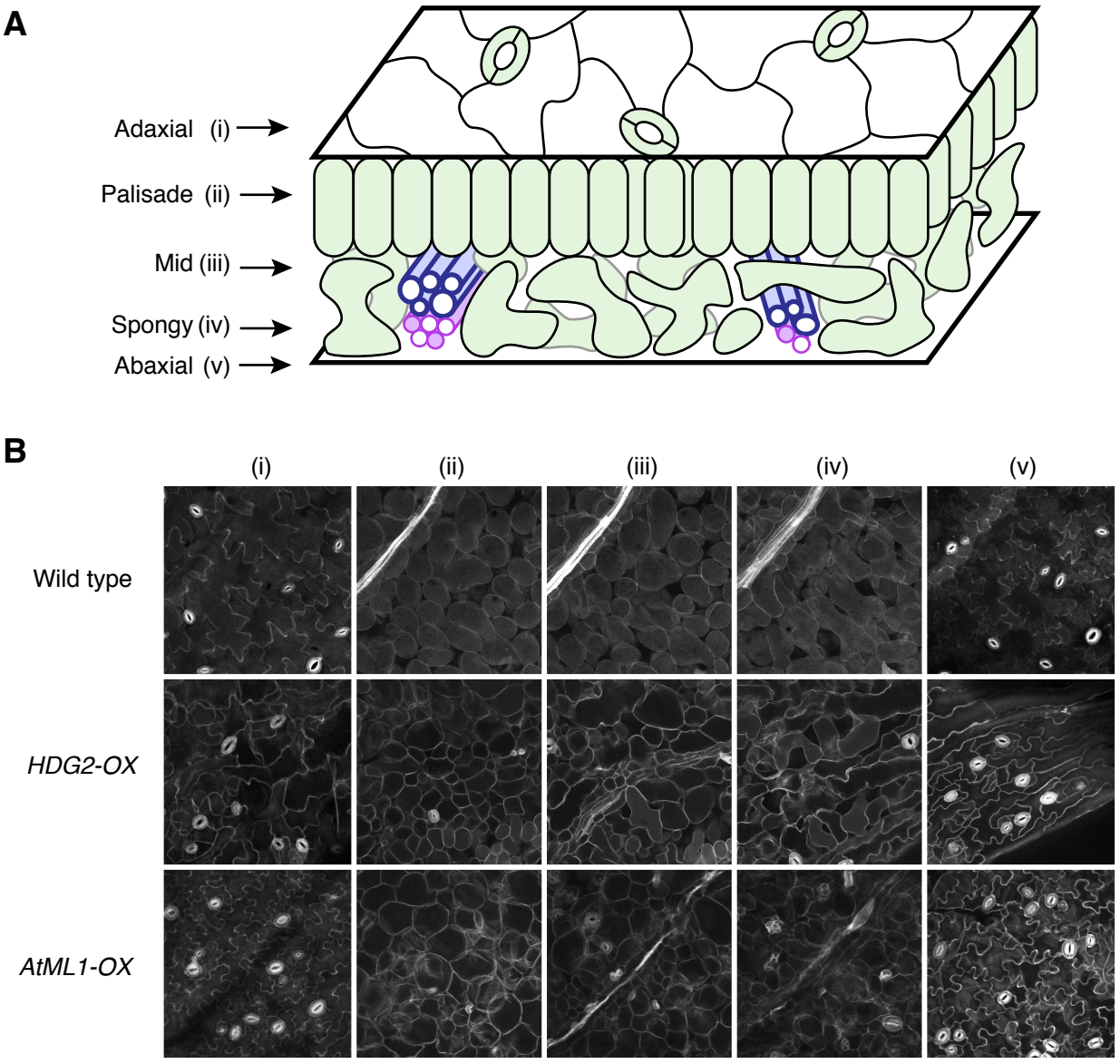


Figure 3-4. Optical serial sectioning through intact cotyledons by mPS-PI staining reveals ectopic internal stomata.
(A) Cartoon of a cotyledon indicating locations of each tissue: (i) adaxial epidermis; (ii) palisade mesophyll; (iii) mid-vein region; (iv) spongy mesophyll; (v) abaxial epidermis. Note that stomata are found in both adaxial and abaxial epidermis in *Arabidopsis*. (B) mPS-PI stained optical serial sections of locations (i-v) in wild-type (top); induced ectopic *HDG2* overexpression (middle: *HDG2-OX*) and induced ectopic *AtML1* overexpression (bottom: *AtML1-OX*). Images were taken under the same magnification. Scale bar, 20 μm. See Supplementary Movies 3-S1, 3-S2, and 3-S3 for Z-stack three-dimensional sections.

Figure 3-5.

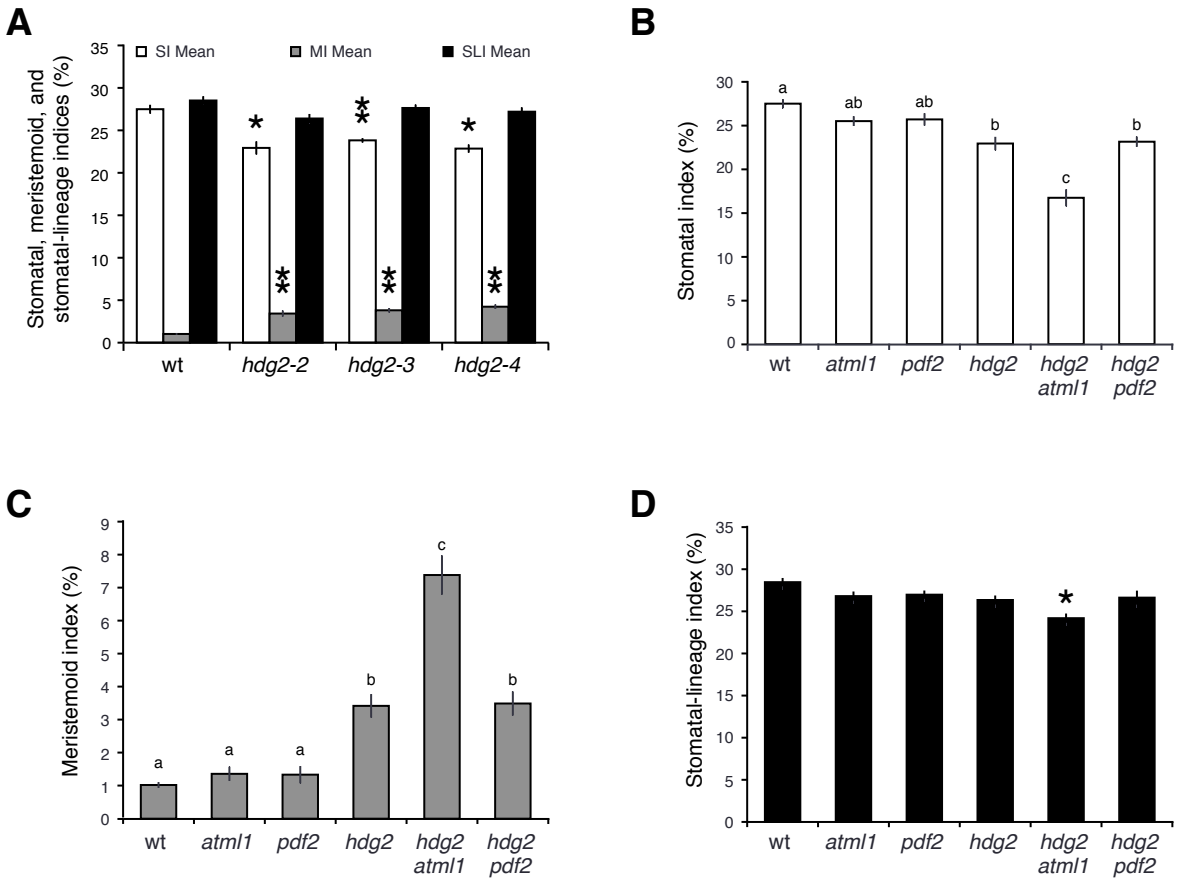


Figure 3-5. Stomatal development defects in *hdg2* mutants and higher-order mutants in closely related HD-ZIP IV genes.
(A) Stomatal index (SI; white bars), meristemoid index (MI; grey bars), and stomatal-lineage index (SLI; black bars) of 10-day-old abaxial cotyledons from wild type (wt) and three independent *hdg2* T-DNA insertion alleles, *hdg2-2, hdg2-3*, and *hdg2-4*. SLI is defined here as the sum of SI and MI. See Figure 3-S1 for RT-PCR analysis. Bars, means (n=8); error bars, s.e.m. * $p<0.0005$; ** $p<0.0001$ (two-tailed Student's t-test of each *hdg2* allele against wild-type). Tukey's HSD did not reveal statistical difference of SI, MI, and SLI among three hdg2 alleles. (B) SI of 10-day-old abaxial cotyledons from wt, *atml1, pdf2, hdg2, hdg2 atml1*, and *hdg2 pdf2*. *hdg2-2* and *atml1-3* alleles were used for the analysis. Bars, means (n=8); Error bars, s.e.m. Total numbers of stomata counted: 680 (wt), 639 (*atml1*), 673 (*pdf2*), 630 (*hdg2*), 550 (*hdg2 atml1*), 620 (*hdg2 pdf2*). (C) MI of the six genotypes described above. Bars, means; Error bars, s.e.m. Total numbers of meristemoids counted: 25 (wt), 34 (*atml1*), 35 (*pdf2*), 93 (*hdg2*), 243 (*hdg2 atml1*), 94 (*hdg2 pdf2*). (D) SLI of the six genotypes described above. Bars, means; Error bars, s.e.m. Total numbers of cells counted: 2516 (wt), 2542 (*atml1*), 2661 (*pdf2*), 2786 (*hdg2*), 3331 (*hdg2 atml1*), 2661 (*hdg2 pdf2*). For (B, C), genotypes with non-significant phenotypes were grouped together with a letter (Tukey's HSD test after One- way ANOVA). For (D), only one genotype was significantly different from others (Tukey's HSD test; $p<0.01$).
[Continued on next page.]

Figure 3-5.

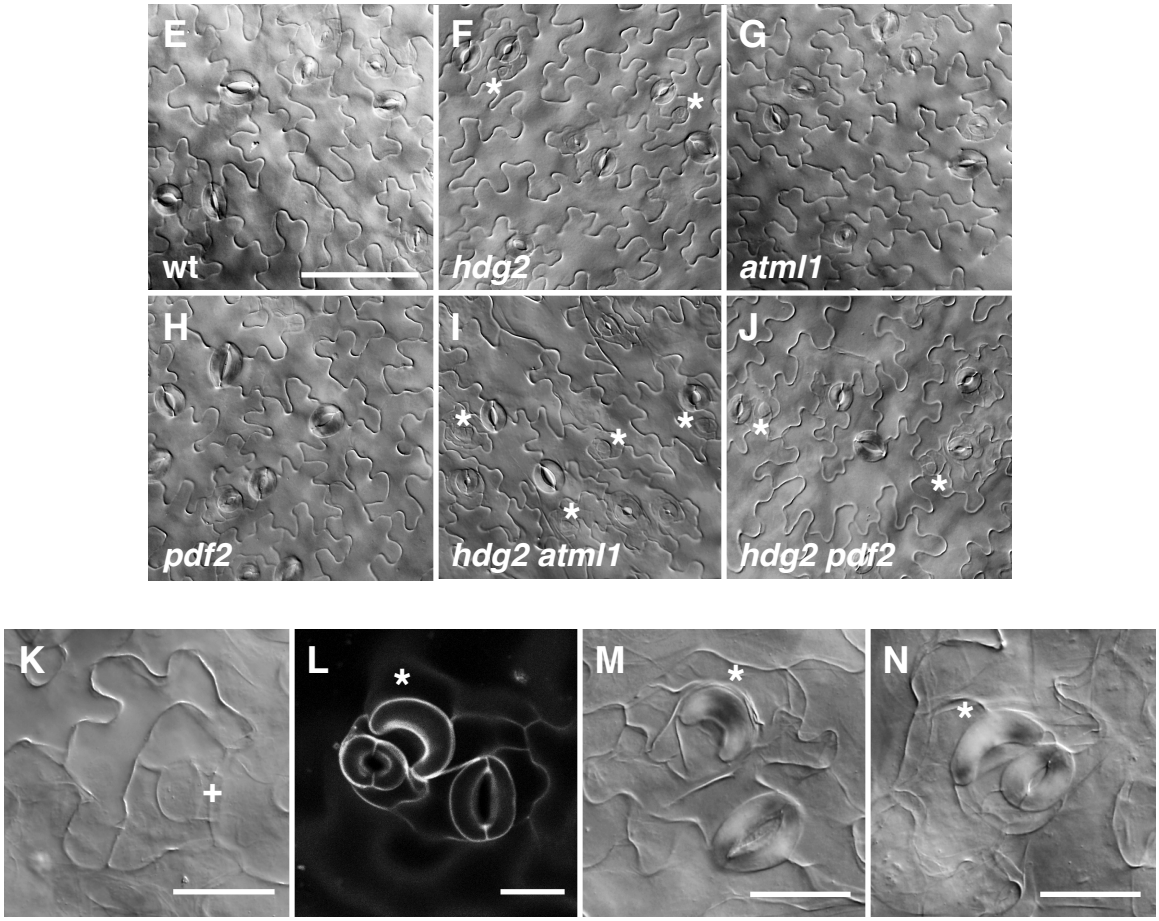


Figure 3-5. Stomatal development defects in *hdg2* mutants and higher-order mutants in closely related HD-ZIP IV genes.

[Continued from previous page.]

(E-J) Representative DIC images of 10-day-old abaxial cotyledons from wild type (wt; E), *hdg2* (F), *atml1* (G), *pdf2* (H), *hdg2 atml1* (I), and *hdg2 pdf2* (J). Asterisks, meristemoids. Images were taken under the same magnification. Scale bar, 100 μm. (K-N) Aberrant stomatal complexes found in *hdg2*. (K) Arrested stomatal precursor after extensive asymmetric amplifying divisions from 30-day-old cotyledon. (L) Confocal image of a stomatal complex with a single GC from 10-day-old abaxial cotyledon. (M, N) Singular GCs from 30-day-old cotyledon.

Figure 3-6.

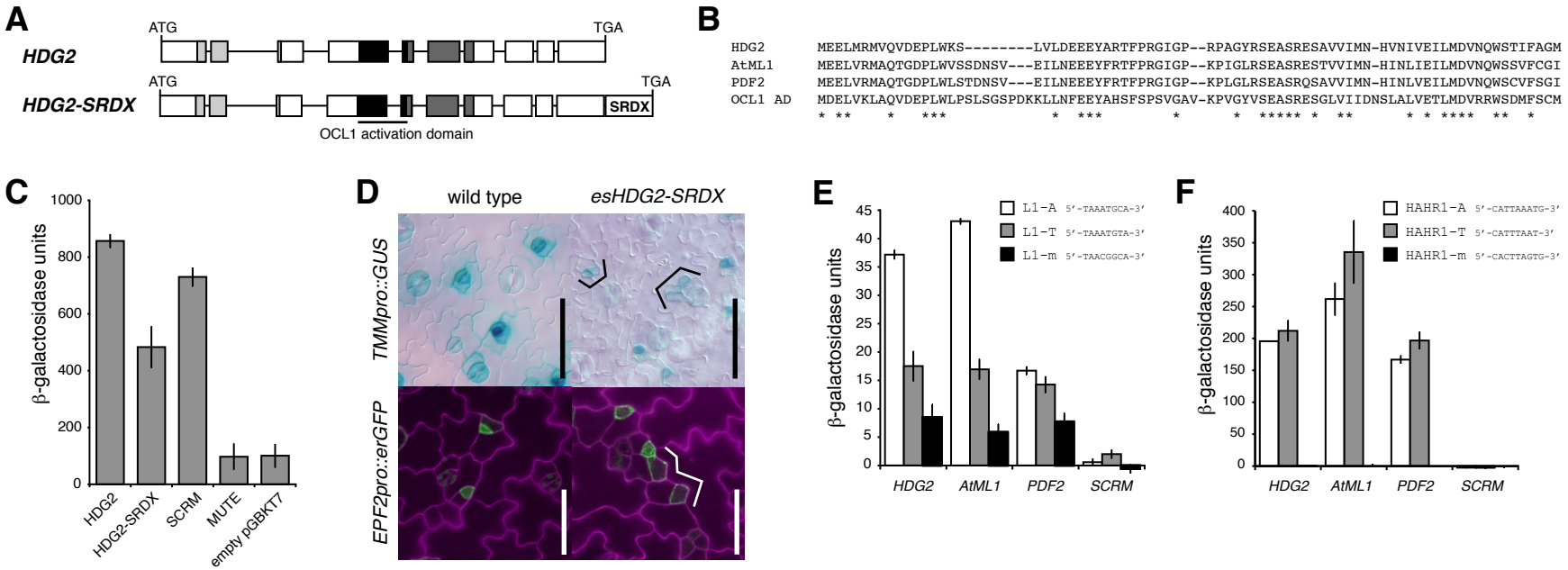


Figure 3-6. HDG2 is a transcriptional activator that can bind to both L1 and HAHR1 boxes.
(A) Schematic diagram of *HDG2* gene, from translational initiation codon to termination codon. Boxes, exons; lines, introns; light gray boxes, homeodomain; dark gray boxes, START domain, which also contains a transcriptional activation domain of maize OCL1 (black boxes). *HDG2-SRDX* has 12 additional amino acids (LDLDLELRLGFA) for forced gene repression. (B) Sequence similarity of *HDG2* with *AtML1*, *PDF2*, and *OCL1* activation domain region. (C) Transcriptional autoactivation of reporter LacZ gene in yeast by HDG2. Addition of SRDX reduces autoactivation. SCRM and MUTE, a known transcriptional activator and repressor, respectively (Kanaoka et al, 2008), were used as positive and negative controls. (D) Induced overexpression of *HDG2-SRDX* confers an epidermal phenotype with delayed stomatal precursors, resulting in expanded, SLGC-like cells (brackets). Shown are 10-day-old cotyledon abaxial epidermis of wild type (left) and *esHDG2-SRDX* (right) expressing *TMMpro::GUS* (top) and *EPF2pro::erGFP* (bottom). Scale bars, 50 μm (top) and 50 μm (bottom). (E) Yeast one-hybrid analysis for activation of LacZ reporter possessing 5′-upstream L1-A-box (white bars), L1-T-box (gray bars), and L1-m-box (black bars) by HDG2, AtML1, or PDF2. SCRM, a known E-box binder (Chinnusamy et al, 2003), was used as a negative control. Bars represent mean value of triplicates. Error bars, s.e.m. (F) Yeast one-hybrid analysis for activation of LacZ reporter possessing 5′-upstream HAHR1-A-box (white bars), HAHR1-T-box (gray bars), and HAHR1-m-box (black bars), respectively, by HDG2, AtML1, or PDF2. SCRM was used as a negative control. Bars represent mean value of triplicates. Error bars, s.e.m.

Figure 3-7.

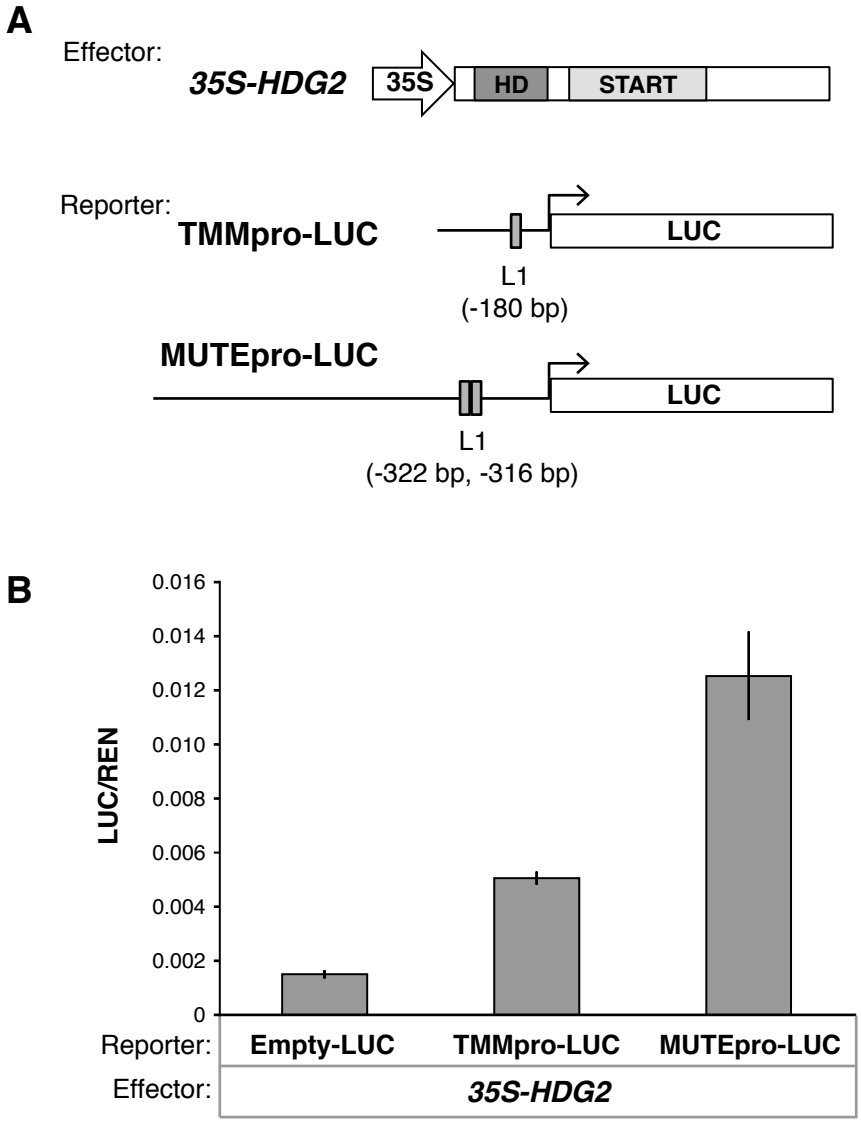


Figure 3-7. HDG2 transactivates *TMM* and *MUTE* promoters in planta.
(A) Schematic diagrams of effector and reporter constructs used. The exact locations of L1 boxes are indicated. (B) Relative luciferase activity in *N. benthamiana* transiently transformed with *35S::HDG2* and the indicated reporters. Luciferase activities were measured seven days after infiltration. Relative luciferase activities are shown by firefly luciferase (LUC) activities normalized to renilla luciferase (REN) activities. Bars, mean; Error bars, s.e.m. (n=4).

Figure 3-S1.

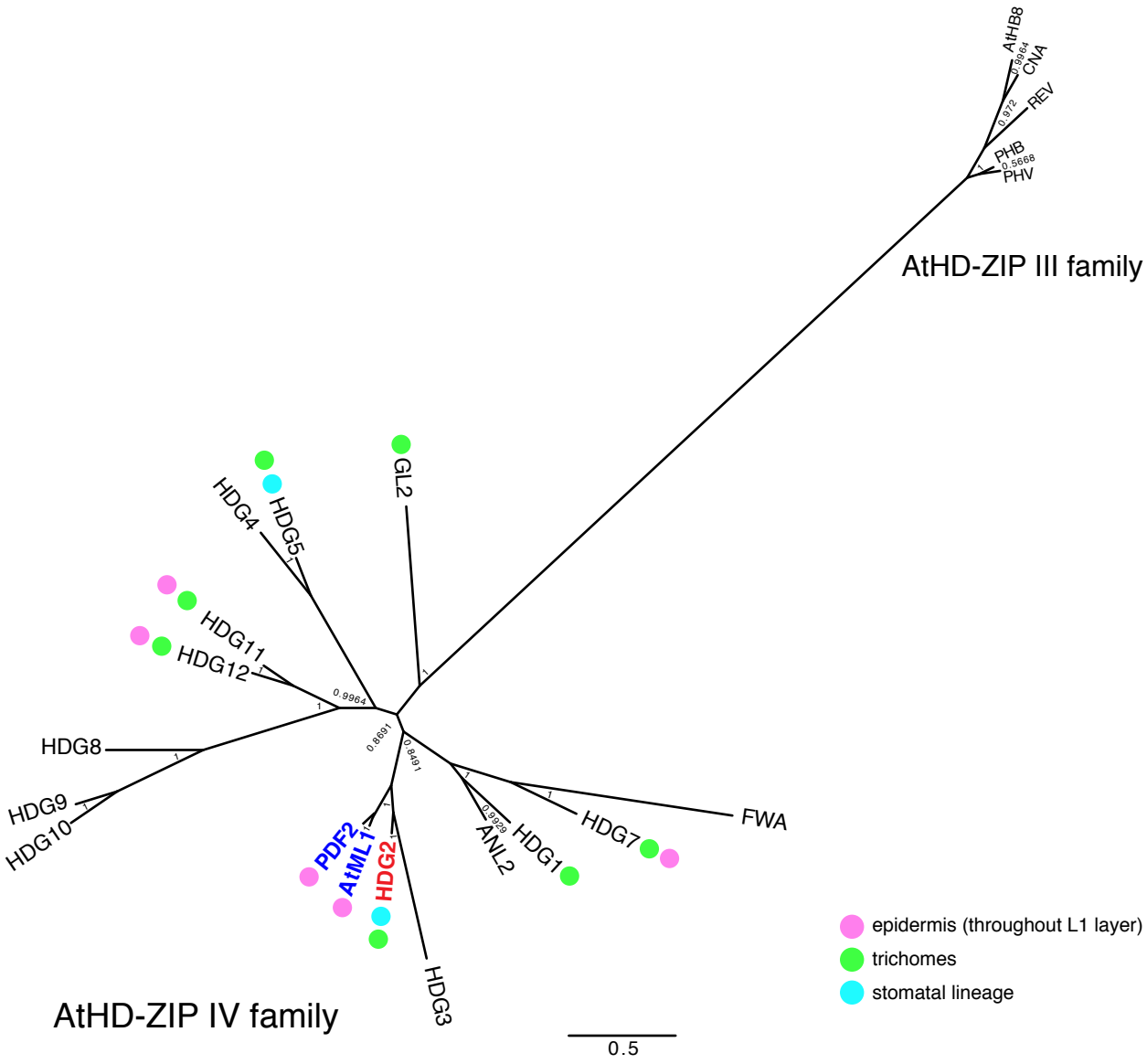


Figure 3-S1. Molecular phylogenetic analysis of AtHD-ZIP IV family.
A phylogenetic tree was constructed using Bayesian analysis with Markov Chain Monte Carlo algorithm for one million generations to obtain a majority-rule consensus tree (see Methods). Full-length amino-acid sequences of sixteen AtHD-ZIP IV family and five AtHD-ZIP III (outgroup) genes were analyzed. Posterior probabilities of key nodes are shown at the respective nodes. The scale bar indicates the number of amino acid changes per branch length. *HDG2* is highlighted in red; *AtML1* and *PDF2* are highlighted in blue to show their close phylogenetic relationship with *HDG2*. Colored dots indicate known expression/promoter activities within the leaf epidermis with the following cell types; pink, uniformly in the protoderm (L1 layer); green, in trichomes; cyan, in stomatal-lineage cells (Rerie et al. 1994; Lu et al. 1996; Session et al. 1999; Abe et al. 2003; Nakamura et al. 2006).

Figure 3-S2.

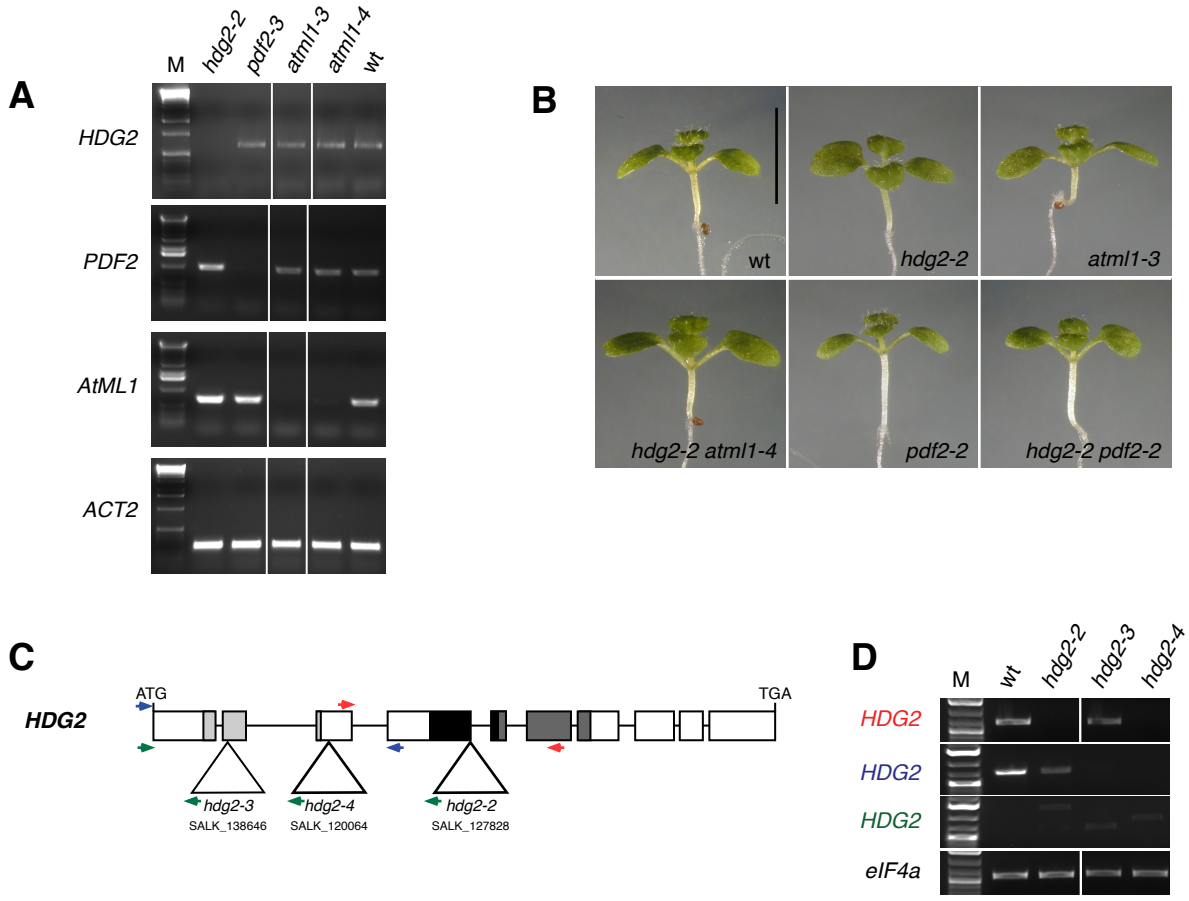


Figure 3-S2. T-DNA insertion knockout alleles of *HDG2*, *PDF2*, and *AtML1* show no obvious growth phenotypes.
(A) RT-PCR analysis of *HDG2*, *PDF2*, and *AtML1* from 10-day-old seedlings of wild type (wt) and T-DNA insertion alleles, *hdg2-2* (SALK_127828C), *pdf2-2* (SALK_109425C), *atml1-3* (SALK_128172), and *atml1-4* (SALK_033408). Corresponding transcripts are undetectable in these T-DNA insertion alleles. *ACT2* transcripts serve as a control. (B) 10-day-old seedlings of the corresponding genotypes. Images were taken under the same magnification. Scale bar, 5 mm. (C) Schematic diagram indicating the T-DNA insertion sites for three mutant alleles within the *HDG2* gene. This diagram includes *HDG2* genomic region from translational initiation codon to termination codon. Light gray boxes, homeodomain; dark gray boxes, START domain, which also contains a transcriptional activation domain of maize OCL1 (black boxes). Color-coded arrows correspond to locations of primer pairs used for RT-PCR analysis (see D). (D) RT-PCR analysis of three *hdg2* T-DNA insertion alleles reveals expression of aberrant transcripts of partial size, which are not derived from genomic DNA contamination based on their size. Each color code corresponds to the following primer pairs used for analysis: green, primers HDG2_1_XhoI and LBa1; blue, primers HDG2_1_XhoI and HDG2_1750.rc; red, HDG2_1168 and HDG2_2359.rc. See panel (C) for locations of primers. For primer sequences, see Table 3-1.

Figure 3-S3.

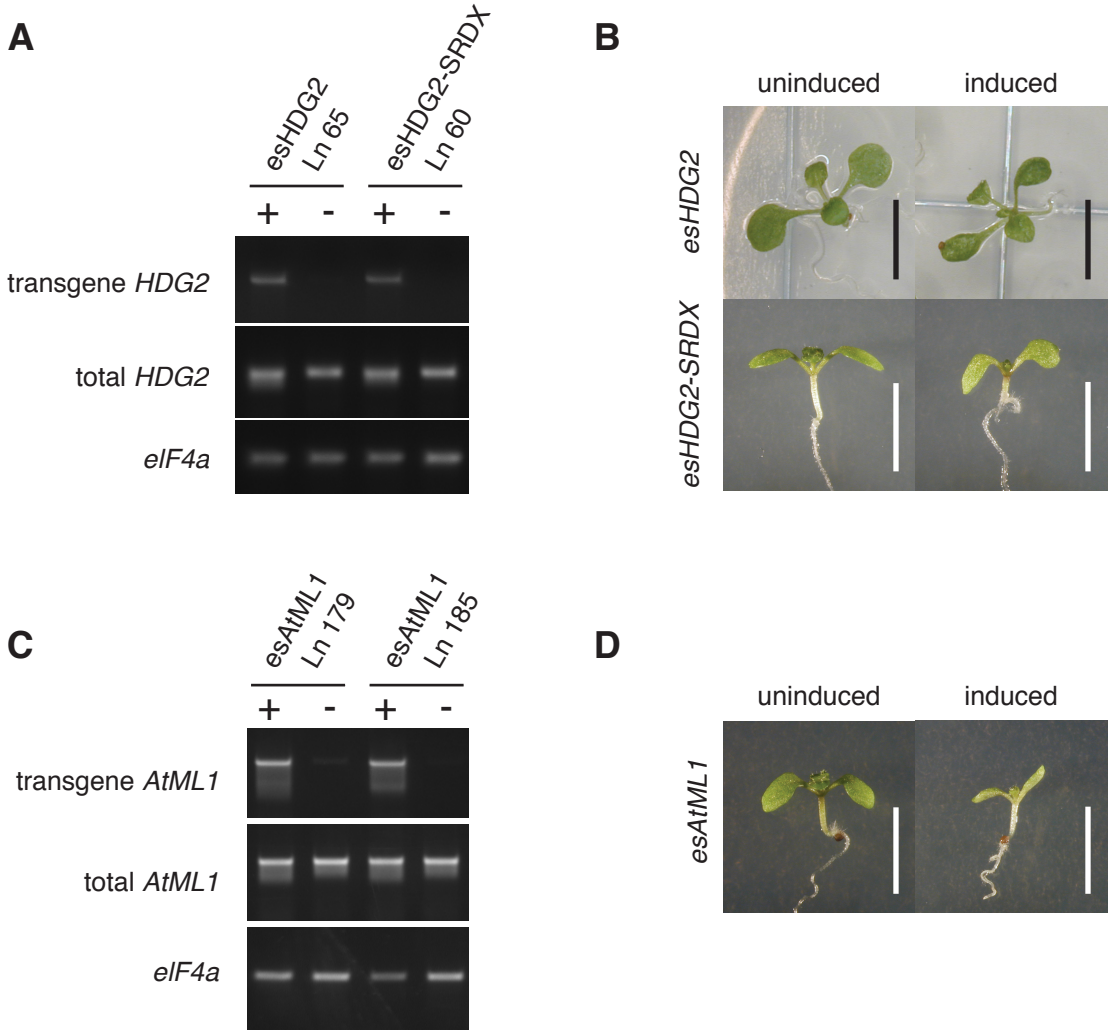


Figure 3-S3. Inducible ectopic overexpressors of *HDG2*, *HDG2-SRDX*, and *AtML1*.
(A) RT-PCR analysis of transgenic Arabidopsis lines carrying inducible *HDG2* and *HDG2-SRDX* with (+) or without (-) estradiol induction. eIF4a transcripts serve as a control. (B) Seedlings of the corresponding transgenic lines, *HDG2* (top; 2 weeks old) and *HDG2-SRDX* (10 days old) with or without induction. Scale bars, 5 mm. (C) RT-PCR analysis of transgenic Arabidopsis lines carrying inducible *AtML1* with (+) or without (-) estradiol induction. eIF4a transcripts serve as a control. (D) Seedlings of a corresponding transgenic line (10 days old) with or without induction. Scale bars, 5 mm.

Figure 3-S4.

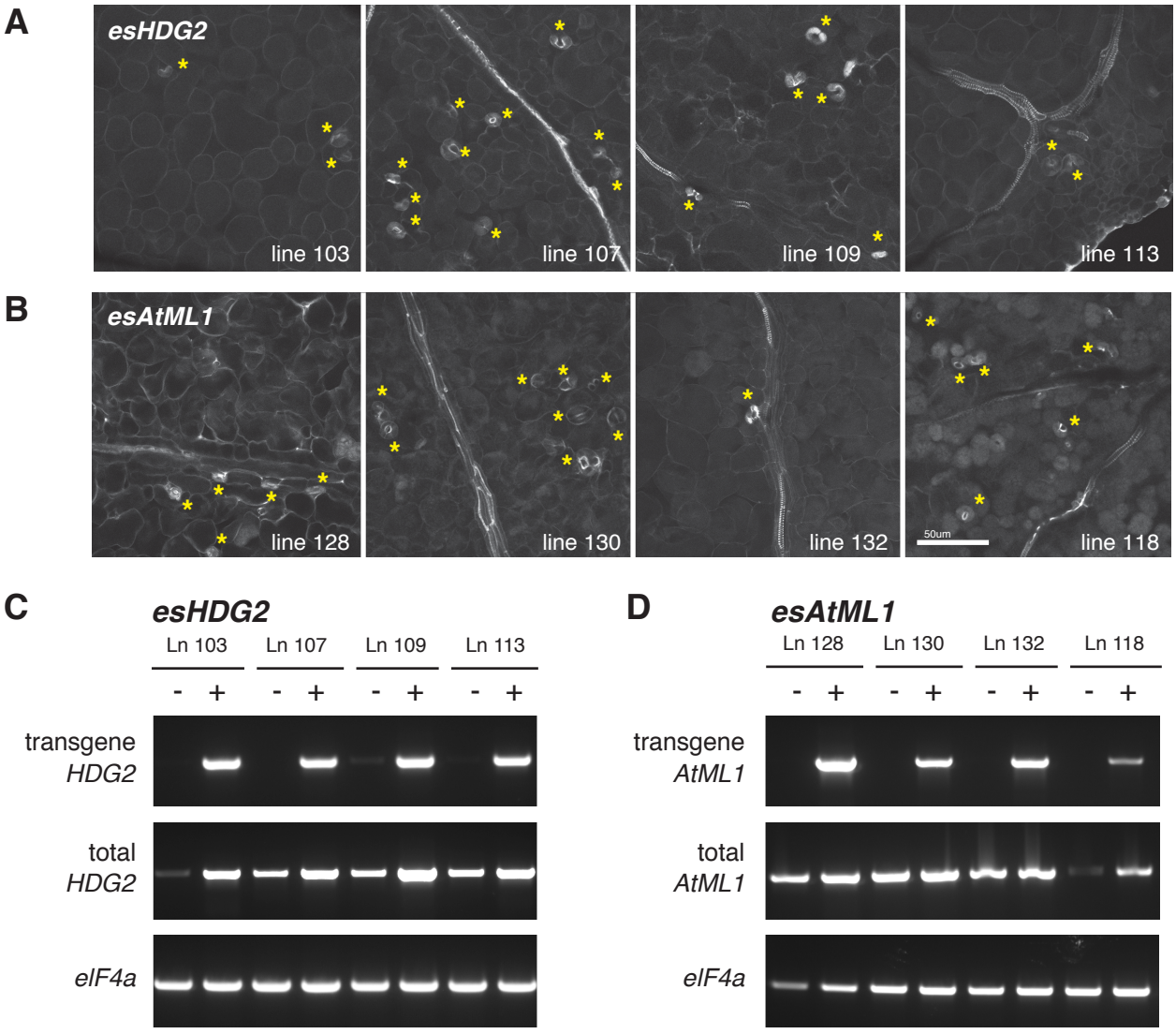


Figure 3-S4. Induced ectopic overexpression of *HDG2* and *AtML1* in four additional transgenic lines for each construct, each showing ectopic stomatal differentiation in mesophyll layers.
(A, B) mPS-PI stained optical sections of mesophyll layer of 12-day-old cotyledons from four independent transgenic lines harboring inducible *HDG2* transgene (A) and four independent transgenic lines harboring inducible *AtML1* transgene (B). Ectopic differentiation of internal stomata (asterisks) is evident. Images were taken under the same magnification. Scale bar, 50 μm. (C, D) RT-PCR analysis of additional transgenic Arabidopsis lines carrying inducible *HDG2* or *AtML1* with (+) or without (-) estradiol induction. eIF4 transcripts serve as a control. The line numbers correspond to the lines examined for phenotypes (A, B).

Figure 3-S5.

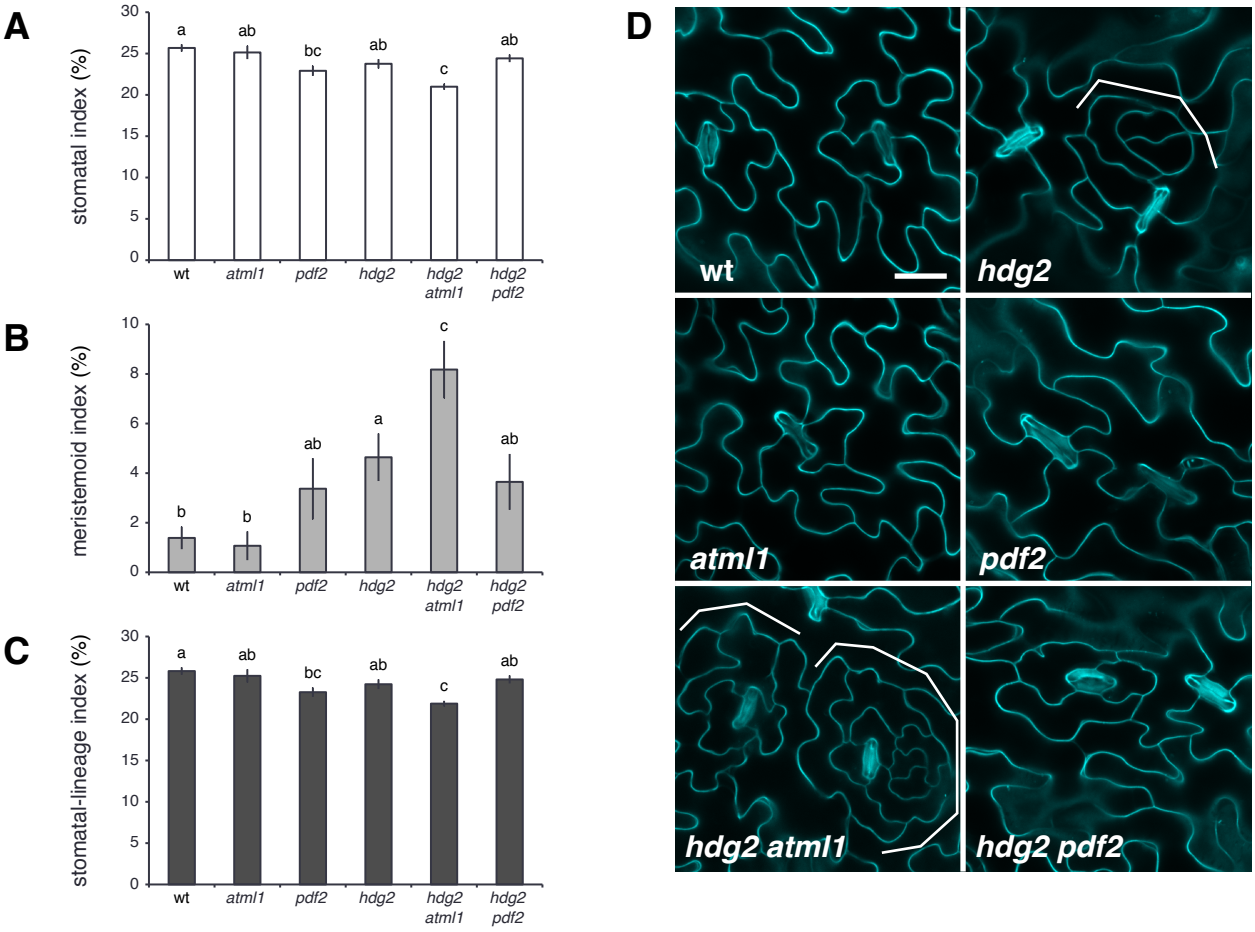


Figure 3-S5. Stomatal development defects in rosette leaves of *hdg2* mutants and higher-order mutants in closely related HD-ZIP IV genes.
(A) SI of 6-week-old abaxial rosette leaf 4 from wt, *atml1*, *pdf2*, *hdg2*, *hdg2 atml1*, and *hdg2 pdf2*. *hdg2-2* and *atml1-3* alleles were used for the analysis. Bars, means (n=16-21); error bars, s.e.m. Total numbers of stomata counted: 1875 (wt), 1224 (*atml1*), 1424 (*pdf2*), 1931 (*hdg2*), 1608 (*hdg2 atml1*), 1749 (*hdg2 pdf2*). (B) MI of the six genotypes described above. Bars, means; Error bars, s.e.m. Total numbers of meristemoids counted: 12 (wt), 5 (*atml1*), 22 (*pdf2*), 40 (*hdg2*), 58 (*hdg2 atml1*), 31 (*hdg2 pdf2*). (C) SLI of the six genotypes described above. Bars, means; Error bars, s.e.m. Total numbers of cells counted: 7272 (wt), 4795 (*atml1*), 6229 (*pdf2*), 8010 (*hdg2*), 7585 (*hdg2 atml1*), 7123 (*hdg2 pdf2*). For (B, C), genotypes with non-significant phenotypes were grouped together with a letter (Tukey's HSD test after One-way ANOVA). For (C), only one genotype was significantly different from others (Tukey's HSD test; $p<0.01$). (D) Representative confocal images of 6-week-old abaxial rosette leaf 4 epidermis from six different genotypes. Images were taken under the same magnification. Scale bar, 20 μm. *hdg2* mutant rosette leaf epidermis shows characteristic delayed stomatal differentiation, a phenotype exaggerated in *hdg2 atml1* double mutant (brackets).

Figure 3-S6.

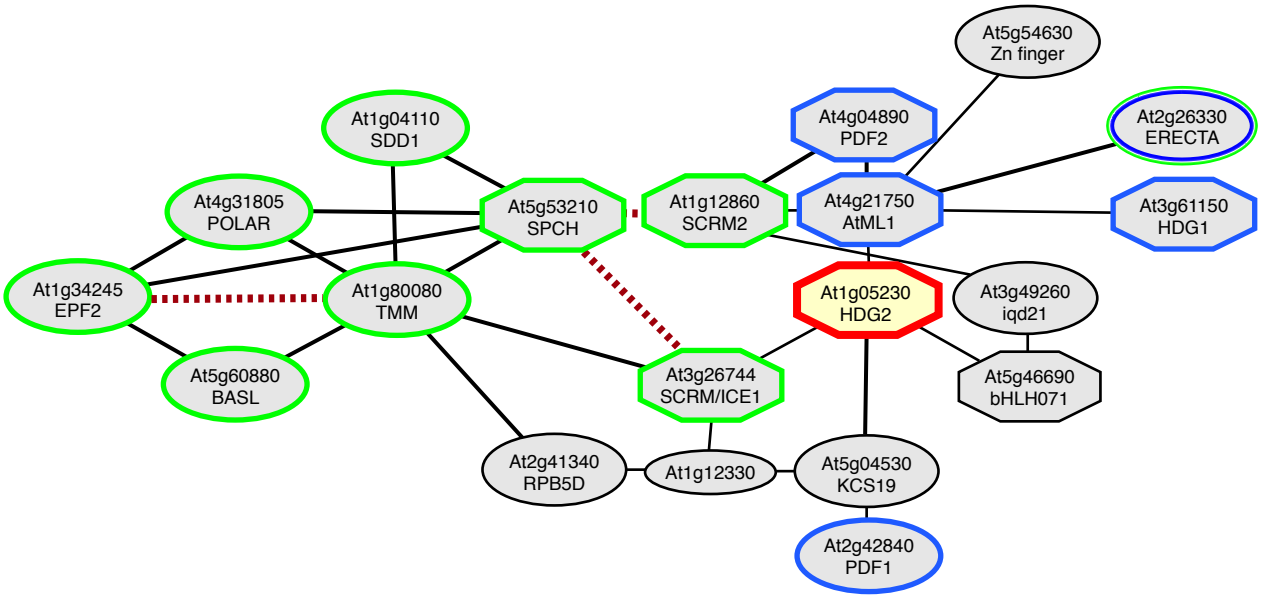


Figure 3-S6. Co-expression network of genes controlling stomatal development and protodermal differentiation.
The network diagram was generated using ATTED II network drawer (http://atted.jp/top_draw.shtml#NetworkDrawer). HDG2 is highlighted in red; known regulators of stomatal development are in green; known genes with protodermal expression are in blue; hexagons represent transcription factors. Thickness of lines (edges) represents mutual rank (MR) of co-expression; the higher the rank, the thicker the line. Dotted red lines indicate known protein-protein interactions.

Figure 3-S7.

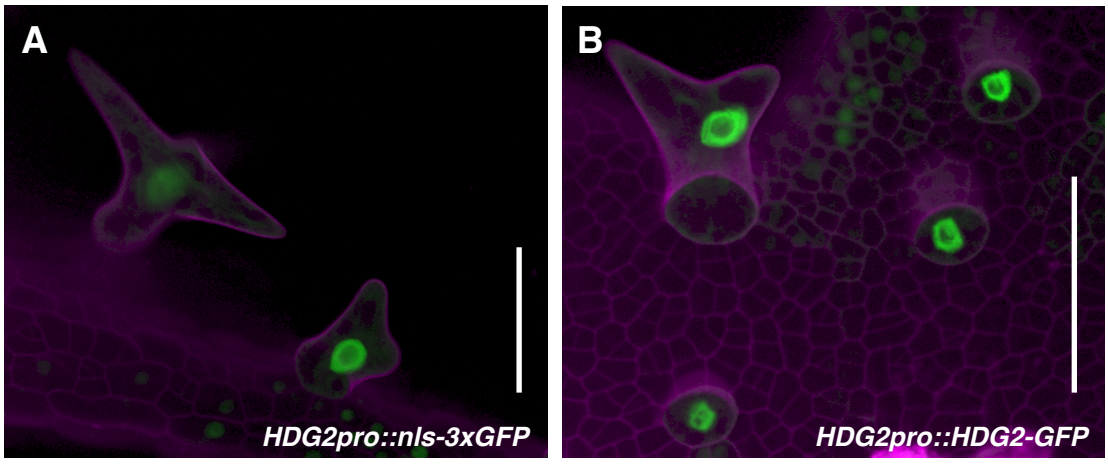


Fig. S7. Expression of transcriptional and translational reporters of *HDG2* in developing trichomes. Shown are confocal images of abaxial rosette leaf epidermis expressing *HDG2pro::nls-3xGFP* (A) and *HDG2pro::HDG-GFP* (B). Each developing trichome has a very large nucleus (fluorescing green). Scale bars, 50 μm.

Table 3-1.

Primer names	Sequences (5' to 3')	Notes
EcoRI-SpeI- L1C-XhoI-F	AATTCACTAGTTAAATGCATAAATGCATAAATGCA	Molecular cloning
EcoRI-SpeI- L1C-XhoI-R	AGCTTGCATTTATGCATTTATGCATTTAACTAGTG	Molecular cloning
EcoRI-SpeI- L1T-XhoI-F	AATTCACTAGTTAAATGTATAAATGTATAAATGTA	Molecular cloning
EcoRI-SpeI- L1T-XhoI-R	AGCTTACATTTATACATTTATACATTTAACTAGTG	Molecular cloning
EcoRI-SpeI- L1m-XhoI-F	AATTCACTAGTTAACGGCATAACGGCATAACGGCA	Molecular cloning
EcoRI-SpeI- L1m-XhoI-R	AGCTTGCCGTTATGCCGTTATGCCGTTAACTAGTG	Molecular cloning
EcoRI-SpeI-HAHR1T- XhoI-F	AATTCACTAGTCATTAAATGCATTAAATGCATTAAATG	Molecular cloning
EcoRI-SpeI-HAHR1T- XhoI-R	AGCTCATTAAATGCATTAAATGCATTAAATGACTAGTG	Molecular cloning
EcoRI-SpeI-HAHR1A- XhoI-F	AATTCACTAGTCATTAAATGCATTAAATGCATTAAATG	Molecular cloning
EcoRI-SpeI-HAHR1A- XhoI-R	AGCTCATTTAATGCATTTAATGCATTTAATGACTAGTG	Molecular cloning
EcoRI-SpeI-HAHR1m- XhoI-F	AATTCACTAGTCACTTAGTGCACTTAGTGCACTTAGTG	Molecular cloning
EcoRI-SpeI-HAHR1m- XhoI-R	AGCTCACTAAGTGCACTAAGTGCACTAAGTGACTAGT	Molecular cloning
13.SpeI-HDG2.gw.F	CACCGAACTAGTAGGAATGTTCGAGCCAAATATGC	Molecular cloning
14.HDG2-NotI.rc	TCTCTCGCGGCCGCTCTCTCTCAAGCAGTCTCACAAGAC	Molecular cloning
15.NtermHDG2-NotI.rc	TCTCTCGCGGCCGCTCTCTCCACGGATAAGTCGATGATG	Molecular cloning

16.SpeI.CtermHDG2.gw.F	CACCGAACTAGTAGGAGTTTCTAGAGCAATGAC	Molecular cloning
17.SpeI.actdomHDG2.gw.F	CACCGAACTAGTAGGAGCTGCAATGGAAGAGC	Molecular cloning
18.actdomHDG2-NotI.rc	TCTCTCGCGGCCGCTCTCTCCATCCCCGCGAAAATC	Molecular cloning
19.HDG2-SRDX.NotI.rc	TCGCGGCCGCTTAAGCGAAACCAAGACGAAGCTCAAGGTCAAGGTCGAGAGCAGTCTCAC	Molecular cloning
1g05230.-3000	ACGTACGTATAACTCCATCT	Cloning HDG2pro for KMP102
1g05230.-1.rc	CTTTTCTTATACTCCGGAAAC	Cloning HDG2pro for KMP102
1g05230.1.GW	CACCATGTTCGAGCCAAATATGCT	Cloning HDG2 genomic for KMP111
HDG2.XhoI.1	CCGCTCGAGATGTTCGAGCCAAATATGC	Cloning HDG2 cDNA for KMP144, RT-PCR
1g05230.835	CATATGCTTACGTTATGCAAATACG	Gentoyping SALK_127828 (T-DNA)
1g05230.1168	ACCAACTCCGTCTCGAAAATGC	Gentoyping SALK_127828 (endogenous allele), RT-PCR
HDG2.1750.rc	TGTTGGTGCAGTGATGGACT	RT-PCR
HDG2.2064.rc	ATTCTGCAACATCATATCAATGTTG	RT-PCR
1g05230.2359.rc	TGCTACTAGCCGGGGGATTTGGTT	Genotyping SALK_127828 (endogenous allele), RT-PCR
HDG2.3390	TGCTTAATGGAGGTGATCCAG	qPCR
HDG2.3568.rc	TGCTTCCATCAGGTTTTGCT	qPCR
1g05230.3634.rc	AGCAGTCTCACAAGACATTG	Cloning HDG2 genomic for KMP111
HDG2.SpeI.3634.rc STOP	GGACTAGTTCAAGCAGTCTCACAAGACATTG	Cloning HDG2 cDNA for KMP144
HDG2.SRDX.SpeI.3634.rc	GGACTAGTTTAAGCGAAACCAAGACGAAGCTCAAGGTCAAGGTCGAGAGCAGTCTCACAAGACATTG	Adding SRDX domain to HDG2 cDNA for KMP142
AtML1.-429	GAGGTTATGTCACGTAGTGGGTGC	Genotyping SALK_033408
AtML1.771.rc	GTTCTCGTGCCTCTCATGTTGTGC	Genotyping SALK_033408
AtML1.XhoI.1	CCGCTCGAGATGTATCATCCAAACATGTTCGAATCTC	Cloning AtML1 for KMP147

AtML1.1293	GTTTCCTAGAGGAATTGGACCG	Sequencing pKMP147
AtML1.1803	GGTTGTGGATGTCTCTTTGGACAG	Genotyping SALK_128172, RT-PCR
AtML1.2876.rc	AACATGTTGCTCTGCCCTGAGTTC	Genotyping SALK_128172, RT-PCR
AtML1.SpeI.3279.rc STOP	GGACTAGTTTAGGCTCCGTCGCAGGCCAG	Cloning AtML1 for KMP147
PDF2.2351.rc	TCCAGAATGAAGTAGCTGCAC	Genotyping SALK_109425, RT-PCR
PDF2.968	TGTAGGAGAAATGTATGGAAC	Genotyping SALK_109425, RT-PCR
LBa1	TGGTTCACGTAGTGGGCCATCG	Genotyping SALK lines
ACT2-3'-RT-F	GATGAGGCAGGTCCAGGAATC	qPCR normalization
ACT2-3'-RT-R	AACCCCAGCTTTTTAAGCCTTT	qPCR normalization
Act2-1	GCCATCCAAGCTGTTCTCTC	RT-PCR control
Act2-2	GCTCGTAGTCAACAGCAACAA	RT-PCR control
eIF4A 707 CDNAf	AGCCAGTGAGAATCTTGGTGAAGC	RT-PCR control
eIF4A 1248 CDNAr	CTAGTACGGCAGAGCAAACACAGC	RT-PCR control
HDG-123aa-GW-FP	CACCATGCACGAGCGGCATGAGAACTC	Cloning HDG2ΔHD
HDG2-RP	TCAAGCAGTCTCACAAGACATTGAAGC	Cloning HDG2ΔHD

Table 3-1. List of primer DNA sequences used for cloning, real-time quantitative RT-PCR, and RT-PCR.

Table 3-2.

Plasmid ID	Description	Insert	Vector	Bac R	Plant/yeast R
pCAB120	Gal4AD-HDG2 cDNA	HDG2	pDEST22	Amp	TRP
pCAB121	HAHR1-A box in pLacZi	HAHR1-A	pLacZi	Amp	URA
pCAB122	HAHR1-T box in pLacZi	HAHR1-T	pLacZi	Amp	URA
pCAB123	HAHR1-m box in pLacZi	HAHR1-m	pLacZi	Amp	URA
pCAB124	L1-C box in pLacZi	L1-C	pLacZi	Amp	URA
pCAB125	L1-T box in pLacZi	L1-T	pLacZi	Amp	URA
pCAB126	L1-m box in pLacZi	L1-m	pLacZi	Amp	URA
pCAB129	AtML1 cDNA in pKUT612	cAtML1	pKUT612	Kan	NA
pCAB130	Gal4AD-AtML1 cDNA	cAtML1	pDEST22	Amp	TRP
pCAB132	Gal4AD-SCRM cDNA	SCRM	pDEST22	Amp	TRP
pKMP102	Gateway: L4 - HDG2 promoter - R1	HDG2pro	pENTR 5' TOPO	Kan	NA
pKMP111	Gateway: L1 - HDG2 genomic - L2	HDG2 genomic	pENTR D-TOPO	Kan	NA
pKMP114	HDG2::HDG2-GFP	KMP102/KMP111	R4pGWB504	Strep/Spec	Hyg
pKMP139	HDG2pro::nls-3xGFP	KMP102/KUT612	AR132.9	Strep/Spec	Hyg
pKMP141	Estradiol-inducible 35S::HDG2 cDNA-SRDX	KMP142	pER8	Strep/Spec	Hyg
pKMP142	HDG2 cDNA-SRDX with stop	cHDG2-SRDX	pBSK II+	Amp	NA
pKMP144	HDG2 cDNA with stop	cHDG2	pBSK II+	Amp	NA
pKMP145	Estradiol-inducible 35S::HDG2 cDNA	KMP144	pER8	Strep/Spec	Hyg
pKMP147	AtML1 cDNA with stop	cAtML1	pBSK II+	Amp	NA
pKMP151	Estradiol-inducible 35S::AtML1 cDNA	cAtML1	pER8	Strep/Spec	Hyg
pLJP249	CaMV35S::HDG2	KMP102	pGWB2	Kan/Hyg	Kan/Hyg
pCS003	TMMpro in LUC reporter vector	pTMM	pGREEN-0800-LUC	Kan	NA
pCS004	MUTEpro in LUC reporter vector	pMUTE	pGREEN-0800-LUC	Kan	NA

pCS001	Empty LUC reporter vector	NA	pGREEN-0800-LUC	Kan	NA
pMK370	HDG2 cDNA exon1 for in situ probe	HDG2cDNA_exon1	pCR4-TOPO	Kan/Amp	NA
pMK371	HDG2 cDNA exon4 for in situ probe	HDG2cDNA_exon4	pCR4-TOPO	Kan/Amp	NA
pMK372	FDH cDNA exon3 for in situ probe	FDHcDNA_exon3	pCR4-TOPO	Kan/Amp	NA

Table 3-2. List of plasmid constructs generated in this study.

Name	Sequence
3xHAHR1-A	**CATTAAATG**CATTAAATGCATTAAATG
3xHAHR1-T	**CATTTAATG**CATTTAATGCATTTAATG
3xHAHR1-m	**CACTTAGTG**CACTTAGTGCACTTAGTG
3xL1-C	**TAAATGCA**TAAATGCATAAATGCA
3xL1-T	**TAAATGTA**TAAATGTATAAATGTA
3xL1-m	**TAACGGCA**TAACGGCATAACGGC

Table 3-3. DNA sequences used for yeast one-hybrid analysis.

Table 3-4.

Transcription factor / motif name	Promoters bound in subset	Promoters bound in genome	P-value	
TATA-box motif	88%	97%	0.297	
MYB1AT	88%	99%	0.290	
MYB4 binding site motif	83%	97%	0.299	
CARGCW8GAT	83%	88%	0.095	
GAREAT	77%	88%	0.126	
AtMYC2 BS in RD22	72%	74%	0.017	*
MYCATERD1	72%	74%	0.017	*
ARF binding site motif	72%	79%	0.019	*
W-box promoter motif	66%	95%	0.738	
MYB binding site promoter	61%	64%	0.036	*
T-box promoter motif	55%	91%	0.764	
BoxII promoter motif	55%	79%	0.473	
L1-box promoter motif	*55%*	*34%*	*0.003*	**
Ibox promoter motif	44%	79%	0.700	
MYB2AT	44%	64%	0.341	
SV40 core promoter motif	33%	52%	0.322	
AtMYB2 BS in RD22	33%	32%	0.077	
ATHB2 binding site motif	27%	23%	0.210	
CArG promoter motif	27%	18%	0.037	*
MYB1LEPR	27%	40%	0.456	
CCA1 binding site motif	27%	57%	0.845	
ABRE-like binding site motif	22%	42%	0.772	
ATHB6 binding site motif	22%	8%	0.026	*
DRE core motif	16%	52%	0.931	
CACGTG motif	16%	31%	0.758	
Gap-box motif	11%	27%	0.798	
RAV1-B binding site motif	11%	42%	0.800	
DREB1A/CBF3	11%	19%	0.561	
GADOWNAT	11%	18%	0.669	
MYB1 binding site motif	11%	21%	0.472	
ACGTABREMOTIFFA2OSEM	11%	30%	0.895	
MYB3 binding site motif	11%	13%	0.446	
LTRE promoter motif	11%	14%	0.391	
Hexamer promoter motif	11%	30%	0.681	
TGA1 binding site motif	5%	6%	0.544	
UPRMOTIFFIAT	5%	6%	0.544	
AGCBOXNPGLB	5%	8%	0.414	
MYB2 binding site motif	5%	4%	0.388	

EveningElement promoter motif	5%	15%	0.878	
AGL3 binding site motif	5%	0.4%	0.042	*
Z-box promoter motif	5%	6%	0.540	
TELO-box promoter motif	5%	22%	0.913	
GCC-box promoter motif	5%	25%	0.798	
AGL1ATCONSENSUS	5%	0.9%	0.090	
UPRE2AT	5%	1%	0.140	
RY-repeat promoter motif	5%	9%	0.680	
AGATCONSENSUS	5%	3%	0.299	

Table 3-4. Bioinformatic analysis of promoter motifs in stomatal development genes. Overrepresentation analysis of promoter elements in 18 genes involved in stomatal development (*EPF1*, *EPF2*, *ERECTA*, *ERL1*, *ERL2*, *TMM*, *SPCH*, *MUTE*, *FAMA*, *SCRM*, *SCRM2*, *HDG2*, *AtML1*, *POLAR*, *BASL*, *MYB88*, *FLP*, and *SDD1*) was performed with the Athena tool (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena).

Chapter 4.

Stomatal cell fate during cotyledon germination: a preliminary study.

Introduction

Stomatal development is a dynamic process in which cell fate is determined by cell-cell signaling and environmental conditions. Five basic helix-loop-helix transcription factors control the major developmental transitions during stomatal development, while global and local signaling modulate stomatal density (see Chapter 1). Genetic studies and reporter imaging have resulted in a clear model of the basic funciton of each of the stomatal master regulators, but their interactions and modulation by cell-cell signaling remain to be investigated.

Time-lapse microscopy is a powerful technique that allows direct examination of the relationship of gene expression and protein localization to cell fate. The time-lapse technique used for POLAR-GFP in Chapter 2 and later published in the *Journal of Visualized Experiments* (Peterson & Torii, 2012; see Appendix) provides a way to assess protein activity over time with direct relation to the eventual fates of cells. To confirm and extend the previous work on stomatal master regulatory transcription factors (see Figure 1-2 for summary), we explored their dynamics during epidermal development using the time-lapse protocol for germinating cotyledons.

This chapter provides preliminary evidence that division polarity and cell fate can be altered by changes in cell-cell signaling and by laser ablation of stomatal-lineage cells during development. In the Discussion, we outline hypotheses that will drive the remaining work as well as upcoming improvements and a powerful new technique for analysis.

Results

Germinating cotyledons express stomatal master regulators as predicted by developmental models

As reviewed in Chapter 1, *SPEECHLESS (SPCH)* is required for the asymmetric entry division to the stomatal lineage (Figure 4-1A). Under time lapse performed starting with germination, *SPCHpro::SPCH-GFP* appeared nuclear-localized as cotyledons emerged from dormancy, around 12-13 hours after germination. Around 75% of *SPCH*-positive cells underwent mitosis, while the remainder lost *SPCH* expression (93 mitoses of 125 nuclei on 3 abaxial cotyledons; see Movie 4-1). Divisions were asymmetric with respect to *SPCH* expression, in accordance with the model in which one daughter cell inherits the meristemoid state while the other loses stomatal lineage identity and dedifferentiates as a stomatal lineage ground cell (SLGC). Figure 4-1B shows a typical example of the expression of SPCH in cells with asymmetric division capability during the initial division of a meristemoid mother cell (MMC) and the genesis of a satellite meristemoid from the SLGC produced by the first division. These results are consistent with observations from rosette-leaf time lapse by Robinson et al. (2011).

MUTE falls next in the developmental sequence previously determined by mutant phenotype and cell-type-specific expression, and is expressed in the late meristemoid (Figure 4-2A; see Chapter 1). In germinating cotyledons, *MUTEpro::MUTE-GFP* appeared around four hours after SPCH-GFP. MUTE-positive cells maintained visible MUTE-GFP levels in single nuclei for approximately one day (Figure 4-2B, Movie 4-2). This result conforms to *MUTE*'s role in the single-cell transition from meristemoid identity to guard mother cell (GMC) identity.

MUTE expression and the GMC transition might be reasonable as a "point of no return" beyond which epidermal cells would not be able to dedifferentiate, but our observations suggest that MUTE-positive cells can lose their stomatal-lineage identities through endogenous cell-cell signaling. Specifically, in a pair of adjacent, initially MUTE-positive cells, we observed loss of MUTE-GFP over the course of around 12 hours (Figure 4-2C). Although MUTE is not known to be regulated directly by the receptor-kinase signaling cascade and lacks the MAPK target sites present in SPCH (see Chapter 1), MUTE was predicted to be phosphorylated by MAPK4 in *Arabidopsis* protein microarrays (Popescu et al. 2009).

MUTE-positive GMCs often divided symmetrically while retaining a low level of MUTE-GFP signal (Figure 4-2D, Movie 4-2). This observation may be due to slower protein degradation in some cases, since it does not occur in all MUTE-positive cells.

The final symmetric division of the GMC into guard cells (GC) is regulated by *FAMA*. Accordingly, *FAMApro::FAMA-GFP* has been reported to appear in GMCs and persist in guard cells, though not to maturity (Figure 4-3A; see Chapter 1) (Ohashi-Ito and Bergmann, 2006). All of these expectations were borne out in the germinating cotyledon (Figure 4-3B).

A dimerizing basic helix-loop-helix partner to the three master regulators, *SCREAM (SCRM)*, is expressed throughout stomatal development and in mature guard cells (Figure 4-4A; see Chapter 1). *SCRMpro::GFP-SCRM* was initially seen in the germinating cotyledon concurrently with SPCH-GFP, at around 12-13 hours (Movie 4-3). During asymmetric division, meristemoids remained SCRM-positive while SLGCs lost GFP-SCRM signal. GFP-SCRM disappeared from SLGCs within a few hours (Figure 4-4B, Movie 4-3), which may indicate specific degradation induced by stomatal cell-cell signaling (likely via EPF2 and EPF1; see Chapter 1). SLGCs that resumed stomatal

lineage identity also resumed GFP-SCRM. GMCs and guard cells continued to express SCRM strongly (Figure 4-4B, Movie 4-3).

Changes in bHLH stability and signaling levels alter stomatal patterning and gene expression

Our investigation of protein dynamics during real-time stomatal development led to the question of how changes in the activity of these bHLH transcription factors might affect cell fates. The semi-dominant mutation *scrm-D* converts nearly all epidermal cells to stomata. The scrm-D protein differs from the wild type by a single amino-acid change that alters a highly conserved motif, the KRAAM domain (Kanaoka et al. 2008). This change may stabilize the SCRM protein against degradation. Examining *SCRMpro::GFP-scrm-D* under time lapse revealed that the GFP-scrm-D protein appeared earlier than GFP-SCRM or SPCH-GFP (Figure 4-5), which is likely due to higher protein levels present during seed development. GFP-scrm-D nuclei differentiated directly into GMCs and divided symmetrically into GCs early in cotyledon development (Figure 4-5).

Another important way bHLH protein levels may be altered is through endogenous cell-cell signaling, such as EPF signaling (see Chapter 1). EPF peptides suppress stomatal lineage cell fates: specifically, EPF2 inhibits stomatal lineage identity at earlier developmental stages (Hara et al., 2009; Hunt & Gray, 2009), and EPF1 does so at later stages (Hara et al., 2007). Examining *SCRMpro::GFP-SCRM* in the *epf1* and *epf2* mutant backgrounds over time can provide information regarding the nature of these signals. Figure 4-6 shows that, in the absence of *EPF2*, SCRM-positive cells arose adjacent to each other and divided asymmetrically without regard to the positions of other SCRM-positive cells. Then, at later developmental stages, the stomatal-lineage cells so produced dedifferentiated to maintain the one-cell spacing rule. Thus, in this

preliminary analysis, the *epf2* mutant exhibited improper directionality in asymmetric divisions. However, *epf2* mutants also develop a larger total number of stomatal-lineage cells than wild type; a larger number of time-lapse experiments will determine whether this is solely due to a lack of lateral inhibition on MMCs or whether meristemoids may perform additional divisions. It will also be interesting to observe the dynamics of SCRM-GFP in the *epf1* mutant background, which produces stomatal pairing by failing to repress stomatal lineage identity at later developmental stages.

Laser ablation alters cell division in the stomatal lineage

Time lapse revealed a surprising expression pattern of GFP-SCRM in pairs of cells, one of which becomes a stoma while the other loses expression or divides away (Figure 4-7A). We collaborated with DRVision, a local software company, to adapt their lineage-tracking software, SVCell, to our time-lapse data. On average, GMC nuclei exhibited earlier GFP-SCRM fluorescence than the nuclei of asymmetrically dividing cells, and GMC nuclei were brighter (Figure 4-7B). We thus hypothesized that all earlier, brighter cells would represent GMC. Under this assumption, we ablated presumptive GMC with laser radiation to alter the behavior of the neighboring cells (Movie 4-4).

Without ablation, SCRM-positive nuclei adjacent to GMC took one of two developmental paths over the course of three-day time-lapse experiments: they lost GFP-SCRM or divided asymmetrically, in approximately equal numbers (39 GFP losses and 44 asymmetric divisions of 83 nuclei on 4 abaxial cotyledons). Asymmetric divisions were typically at the 60° angle described previously (see Chapter 1; data not shown).

At 20 hours post-germination (hpg), nonablated cells frequently divided in a plane nearly parallel to the ablated cell rather than at 60° away from it (Figure 4-7C,D).

Occasionally, the plane of cell division was perpendicular to the shared wall, but the meristemoid arose next to the ablated precursor (data not shown). In no ablation has the neighbor cell lost GFP-SCRM expression (15 nuclei on 9 abaxial cotyledons). To determine the statistical significance of these incidents, a larger number of ablation experiments needs to be conducted. However, our preliminary findings suggest that that the absence of a living GMC neighbor randomizes the polarity of the neighboring cell and allows it to maintain stomatal lineage identity.

No cell next to an ablated GMC divided symmetrically to form a stoma. This result may reflect an asymmetrically dividing cell state intrinsic to the MMC, or it may be that GMCs were not ablated sufficiently early in development for the neighbors' fate to be altered to that extent.

Discussion

Time-lapse experiments on stomatal master regulatory proteins support existing hypotheses about the developmental timing and functions of these genes, as reviewed in Chapter 1. This validation of thorough, time-tested genetic analyses and inferences from still-image reporter studies bears out the idea that time-lapse imaging is a powerful tool for accurately assessing stomatal development. Our preliminary results include improper division polarity in *epf2* mutant epidermis and the dynamics of scrm-D protein. Strikingly, we altered cell polarity and fate with laser cell ablation, an exciting result which merits further study. The cotyledon time-lapse method opens up many new avenues for the exploration of protein regulation and movement as well as direct cellular manipulation.

This investigation is poised to take several leaps forward in the next few months. One of the major drawbacks to the study so far is the lack of specific cell-state

information in determining targets for laser ablation, which will soon be rectified by the addition of double markers (GFP-SCRM with SPCH-, MUTE-, and FAMA-TagRFP, or vice versa). Currently cell state is estimated by timing and strength of GFP-SCRM fluorescence in stomatal lineage nuclei, which is relatively reliable. However, as shown in Figure 4-7B, this technique is not completely foolproof, as sometimes asymmetrically dividing nuclei exhibit GFP-SCRM early and may intermittently exceed the fluorescence of GMCs. Ablation results are very likely to depend on the precise developmental stages of both the ablated cell and its neighbor. Thus, the ability to visualize coexpressed stomatal development genes will help immensely in getting comparable, quantifiable results for depolarization studies and in ablating sufficiently early to redirect cell fate.

Double markers will also allow evaluation of the hypothesis that stomatal development is initiated in the overlap of *SPCH*- and *SCRM*-expressing cells. Currently, no definite evidence has been presented of the factor initiating the stomatal lineage; although *SPCH* is absolutely required, it is possible that another critical factor lies upstream. More likely, though, is that *SPCH* and *SCRM* are expressed somewhat stochastically in all epidermal cells and that a positive feedback loop supports stomatal development in the cells where both occur above baseline levels. Then signaling from the new stomatal precursors creates a nonrandom distribution of cells taking on MMC fate. Viewing *SPCH* and *SCRM* reporters together over time will provide evidence relevant to this important question.

A few other small questions can be solved as well. For instance, it has never been conclusively demonstrated that *SPCH* and *MUTE* follow sequentially in the embryo and germinating cotyledon, and that can be established easily with dual-color reporters. *SCRM* and *SCRM2* are thought to overlap in function but have primary expression

times that account for their mutant phenotypes; this will also be simple to address with a dual-reporter line.

Treatment with EPF peptides and inducible EPF expression are known to result in a no-stomata phenotype, so it is likely that they change gene expression promptly and quantifiably upon application. Application of peptides or drugs to cotyledons already undergoing time lapse would be difficult using the current protocol, because a protective layer of agar overlies a scattering of water-mounted specimens. However, microfluidics devices offer a great deal of promise for the instantaneous addition of bioactive molecules following baseline imaging without sample disruption. Devices would need to be adapted, but the RootChip (Grossmann et al., 2012) in particular has a great deal of promise if structures can be added to stabilize growing cotyledons in three dimensions.

The results of the initial *SCRMpro::GFP-SCRM epf2* trial indicate that time lapse can be used in mutant plants to better understand the developmental origins of mutant phenotypes. Detailed measurements of fluorescence intensity and angles of division can illuminate new aspects of cell-cell signaling, such as whether EPF2 affects cell polarity in stomatal lineage divisions. So far, it appears that EPF2 may itself act as a polarizing signal, which leads to questions about the localization and interactions of its receptor, ERECTA, during signal transduction, and about the dynamics of downstream interactors, all of which may be possible to address using the time-lapse technique. The directional randomization observed in ablated-cell neighbor division implies that neighbor cells detect which direction the EPF2 peptide is coming from, allowing spacing divisions to be correctly oriented away from existing stomatal precursors.

Many viable stomatal development mutants are available for transformation and subsequent cotyledon imaging with a variety of reporters. *SCRMpro::GFP-SCRM* is a

suitable marker for stomatal-lineage nuclei, but *POLARpro::POLAR-GFP* could also be used to track incipient cell divisions through its polarized localization. Mutants with similar phenotypes may arrive at them through different mechanisms, which close observation over time would illuminate.

Another important question regards the timing of polarity determination. To investigate this, laser ablation could be performed with asymmetrically localized markers such as POLAR and BASL to see whether already-polarized cells can have their division planes altered. There may be technical difficulties with photobleaching such reporters using the current method, due to their location at the cell periphery, but an argon laser's higher power could allow the use of only nonbleaching wavelengths. There are conflicting reports of possible polar localization in the receptor ERL1, as well, which may occur earlier and upstream of BASL and POLAR movement. Time lapse can clarify whether such polarization occurs and when. Perhaps such a polarized receptor would reorient when its neighboring GMC was ablated.

One of the major unreached goals remaining in this study is quantifiability of gene expression levels. The SVCell software will be extremely helpful in this regard. As yet, limitations of the software and experimental conditions restrict nuclear tracking to prealigned images, and many tracks require manual correction. However, this tool can be improved on an ongoing basis and has already resulted in large quantities of data. Quantification will allow statistically significant differences between ablated and unablated conditions to be demonstrated.

Some analytical methods must be developed to cope with the massive influx of data generated by SVCell. One of the most pressing involves direct comparisons among cell tracks: How should asynchronous time courses best be compared? The existing presentation of the data, such as Figure 4-7B, relies on the viewer's capacity to integrate

a small number of individual tracks, but that is untenable for larger sample sizes. It may be best to synchronize based on initial reporter visibility, or working backward from the eventual fate may be more relevant. It is important to check whether the time involved in developmental processes changes over the course of organ growth. Even as SVCell stands, it is certainly possible to quantify means and variances of parameters like time taken to division, which may be useful in creating computer models of development. This is a very promising tool with a great deal of potential going forward.

Materials and Methods

Plant material and growth conditions

Arabidopsis thaliana Columbia (Col) accession was used as a wild type; all mutants and lines used in this study were in the Col background. The *epf2-1* insertion line was obtained from ABRC. The following mutants and reporter lines were described previously: *epf2-1* (Hara et al., 2009), *scrm-D* (Kanaoka et al., 2008), *SPCHpro::SPCH-GFP* (Guseman et al., 2010), *MUTEpro::MUTE-GFP* (Pillitteri et al., 2007), *FAMApro::FAMA-GFP* (Pillitteri et al., 2007). *SCRMpro::GFP-SCRM* was introduced into *epf2-1* mutant background by genetic crossing. Seeds were sterilized using bleach solution (33% bleach and 0.05% Tween-20) for 15 min and washed five times with sterile water, then placed at 4°C for two days.

Time-lapse imaging

See Appendix for the full text and figures of the publication describing the time-lapse method for germinating cotyledons. Briefly, a solution of 0.5% Bacto Agar was prepoured into chamber slides (Lab Tek II chambered #1.5 German cover glass system; Fisher Scientific) at room temperature; dissected cotyledons were placed beneath the agar layer to immobilize and protect from desiccation, then imaged with an inverted

laser scanning confocal microscope. To avoid any complications due to circadian gene regulation, imbibed and refrigerated seeds were dissected at the same time (10 a.m.) across all experiments.

Microscopy

Confocal microscopy images were taken on the Zeiss LSM700: GFP excitation was performed at 488 nm and emission collected with a band-pass filter of 470 to 500 nm.

Laser ablation

Cells ablated by laser were imaged with the Zeiss LSM700 as above, then the nuclei were marked with circles of radius 4 in Zen 2009 software. Bleaching was performed using LED lasers of 405, 488, and 555nm at 100%, with 30 repetitions at a scan speed of 3.

Supplemental Movie Legends

Movie 4-1. Time-lapse imaging of *SPCHpro::SPCH-GFP* localization in germinating cotyledon. During the first two days of cotyledon development, SPCH-positive cells divide asymmetrically, consistent with SPCH's role in stomatal-lineage entry division. Z-stacks were taken every 30 minutes with 1μm intervals and visualized by maximum intensity projection.

Movie 4-2. Time-lapse imaging of *MUTEpro::MUTE-GFP* localization in germinating cotyledon. During the first two days of cotyledon development, MUTE-positive cells do not divide asymmetrically, consistent with MUTE's role in the single-cell transition from meristemoid to GMC cell state. Cells with a low level of MUTE-GFP occasionally divide symmetrically when protein persists into guard-cell mitosis. Z-stacks were taken every 30 minutes with 1μm intervals and visualized by maximum intensity projection.

Movie 4-3. Time-lapse imaging of *SCRMpro::GFP-SCRM* localization in germinating cotyledon. During the first two days of cotyledon development, SCRM-positive cells may divide either asymmetrically or symmetrically, consistent with SCRM's expression throughout stomatal development. Z-stacks were taken every 30 minutes with 1μm intervals and visualized by maximum intensity projection.

Movie 4-4. Laser cell ablation of *SCRMpro::GFP-SCRM* putative GMC in germinating cotyledon. During time-lapse microscopy, two SCRM-positive nuclei were ablated with laser radiation at 24 hours post-germination. Their neighboring SCRM-positive nuclei divided asymmetrically at atypical angles. Last frame shows cotyledon stained with propidium iodide to highlight cells killed by ablation (shrunken and highly stained on left, stained nucleus on right). Z-stacks were taken every 30 minutes with 1μm intervals and visualized by maximum intensity projection.

Acknowledgements

We thank A. Rychel for early assistance in developing the time-lapse protocol and Prof. C. Queitsch for comments on the chapter. K.M.P. was supported by the National Science Foundation Graduate Research Fellowship (DGE-0718124). K.U.T. is a Howard Hughes Medical Institute–Gordon and Betty Moore Foundation investigator.

References

Grossmann, G., Meier, M., Cartwright, H. N., Sosso, D., Quake, S. R., Ehrhardt, D. W., & Frommer, W. B. (2012) Time-lapse fluorescence imaging of *Arabidopsis* root growth with rapid manipulation of the root environment using the RootChip. *Journal of Visualized Experiments* 65: e4290.

Guseman, J.M., Lee, J.S., Bogenschutz, N.L., Peterson, K.M., Virata, R.E., Xie, B., ... Torii, K.U. (2010) Dysregulation of cell-to-cell connectivity and stomatal patterning by loss-of-function mutation in *Arabidopsis CHORUS (GLUCAN SYNTHASE-LIKE 8) Development* 137: 1731-1741.

Hara, K., Kajita, R., Torii, K.U., Bergmann, D.C., & Kakimoto, T. (2007) The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. *Genes & Development* 21: 1720-1725.

Hara, K., Yokoo, T., Kajita, R., Onishi, T., Yahata, S., Peterson, K.M., ... Kakimoto, T. (2009) Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2, in *Arabidopsis* leaves. *Plant & Cell Physiology* 50(6): 1019-1031.

Hunt, L., & Gray, J.E. (2009) The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Current Biology* 19: 1-6.

Kanaoka, M.M., Pillitteri, L.J., Fujii, H., Yoshida, Y., Bogenschutz, N.L., Takabayashi, J., ... Torii, KU. (2008) SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* 20: 1775-1785.

Peterson, K. M., & Torii, K. U. (2012) Long-term, high-resolution confocal time lapse imaging of *Arabidopsis* cotyledon epidermis during germination. *Journal of Visualized Experiments* 70: e4426.

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

Popescu, S.C., Popescu, G.V., Bachan, S., Zhang, Z., Gerstein, M., Snyder, M., & Dinesh-Kumar, S.P. (2009) MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes & Development* 23: 80-92.

Robinson, S., Barbier de Reuille, P., Chan, J., Bergmann, D., Prusinkiewicz, P., & Coen, E. (2011) Generation of spatial patterns through cell polarity switching. *Science* 333(6048): 1436-1440.

Figure 4-1.

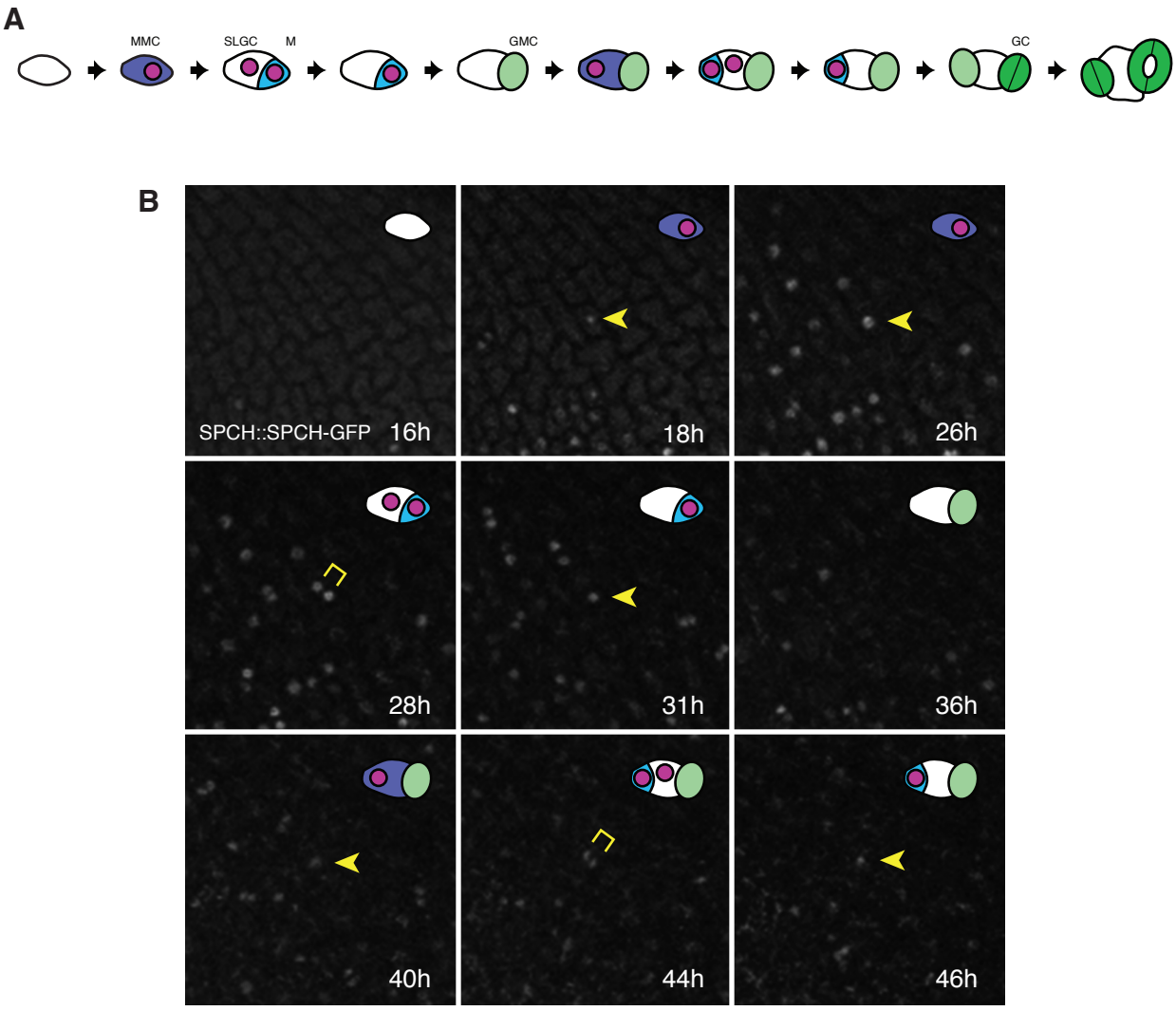


Figure 4-1. *SPCH* determines the asymmetric division state in the stomatal lineage.
(A) Detailed developmental diagram of expected SPCH occurrence during stomatal development. See also Figure 1-2. Meristemoid mother cells (MMC) express *SPCH* upon acquiring asymmetric division capability; when they divide, both daughter nuclei briefly retain SPCH protein but only the meristemoid (M) continues expression. *SPCH* is not expressed in differentiating guard mother cells (GMC). Non-meristemoid sister cells from asymmetric divisions are termed stomatal lineage ground cells (SLGC) and can resume stomatal lineage fate as MMC during their early development, in which case they express *SPCH* during asymmetric divisions as previously.
(B) Time series of *SPCHpro::SPCH-GFP* in a germinating cotyledon. Cell diagrams show comparable stages in (A). Arrowheads, *SPCH*-positive nuclei of interest; brackets, asymmetrically divided nuclei.

Figure 4-2.

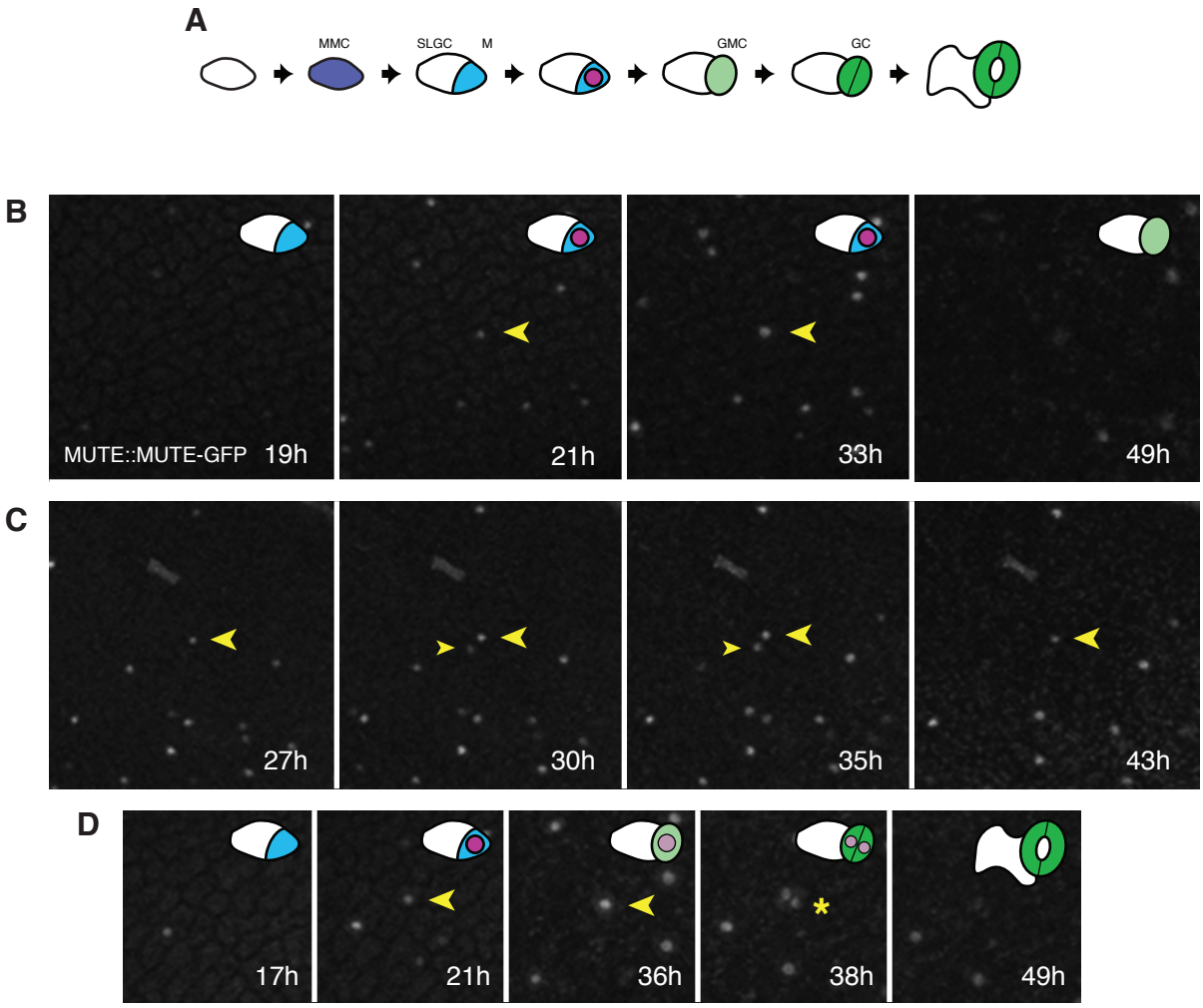


Figure 4-2. *MUTE* determines the transition from asymmetric division to symmetric division in the stomatal lineage.
(A) Detailed developmental diagram of expected MUTE occurrence (pink) during stomatal development. See also Figure 1-2. Meristemoids (M) express *MUTE* upon ceasing asymmetric division and changing to guard mother cells, which have symmetric division capability.
(B) Time series of *MUTEpro::MUTE-GFP* in a germinating cotyledon. Cell diagrams show comparable stages in (A). Arrowheads, *MUTE*-expressing nuclei of interest.
(C) Time series showing MUTE expression in nuclei of adjacent cells (arrowheads), one of which (smaller arrowhead) is repressed by cell-cell signaling (likely EPF1 secretion; see Chapter 1) to maintain the stomatal one-cell spacing rule.
(D) Cells containing MUTE-GFP are occasionally seen to divide symmetrically, but fluorescence rapidly fades in the young GCs. Cell diagrams show comparable stages in (A), with lighter pink nuclei indicating unexpected MUTE-GFP persistence. Arrowheads, *MUTE*-expressing nuclei of interest; bracket, symmetrically divided *MUTE*-expressing nuclei.

Figure 4-3.

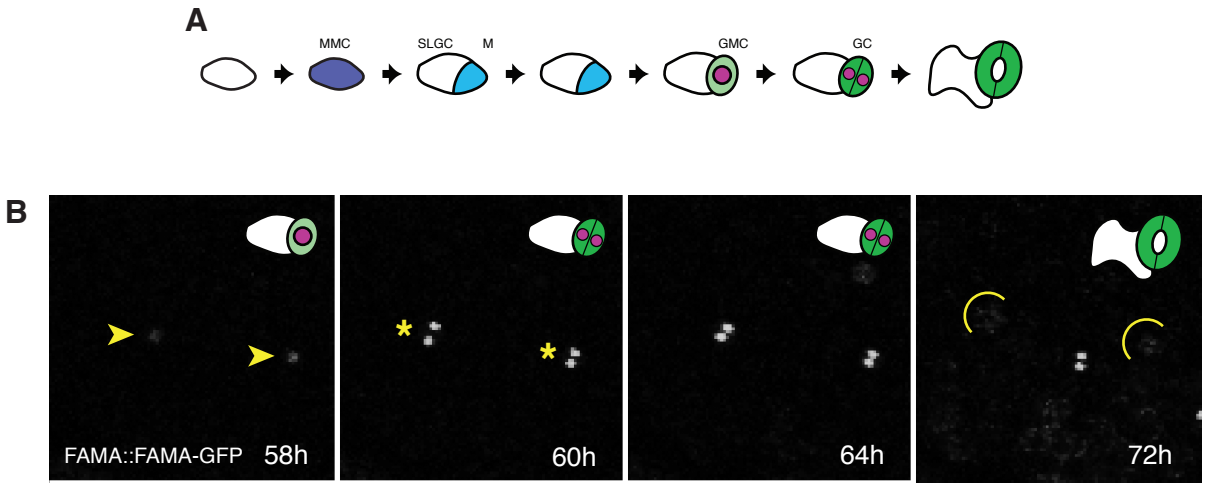


Figure 4-3. *FAMA* defines symmetric division and guard cell differentiation.
(A) Detailed developmental diagram of expected FAMA occurrence (pink) during stomatal development. See also Figure 1-2. Guard mother cells (GMC) express *FAMA* upon differentiation and while dividing symmetrically to form guard cells.
(B) Time series of *FAMApro::FAMA-GFP* in a germinating cotyledon. Cell diagrams reference comparable stages in (A). Arrowheads, *FAMA*-expressing nuclei of interest; asterisks, symmetrically divided *FAMA*-expressing nuclei; semicircles, mature guard cells no longer expressing *FAMA*.

Figure 4-4.

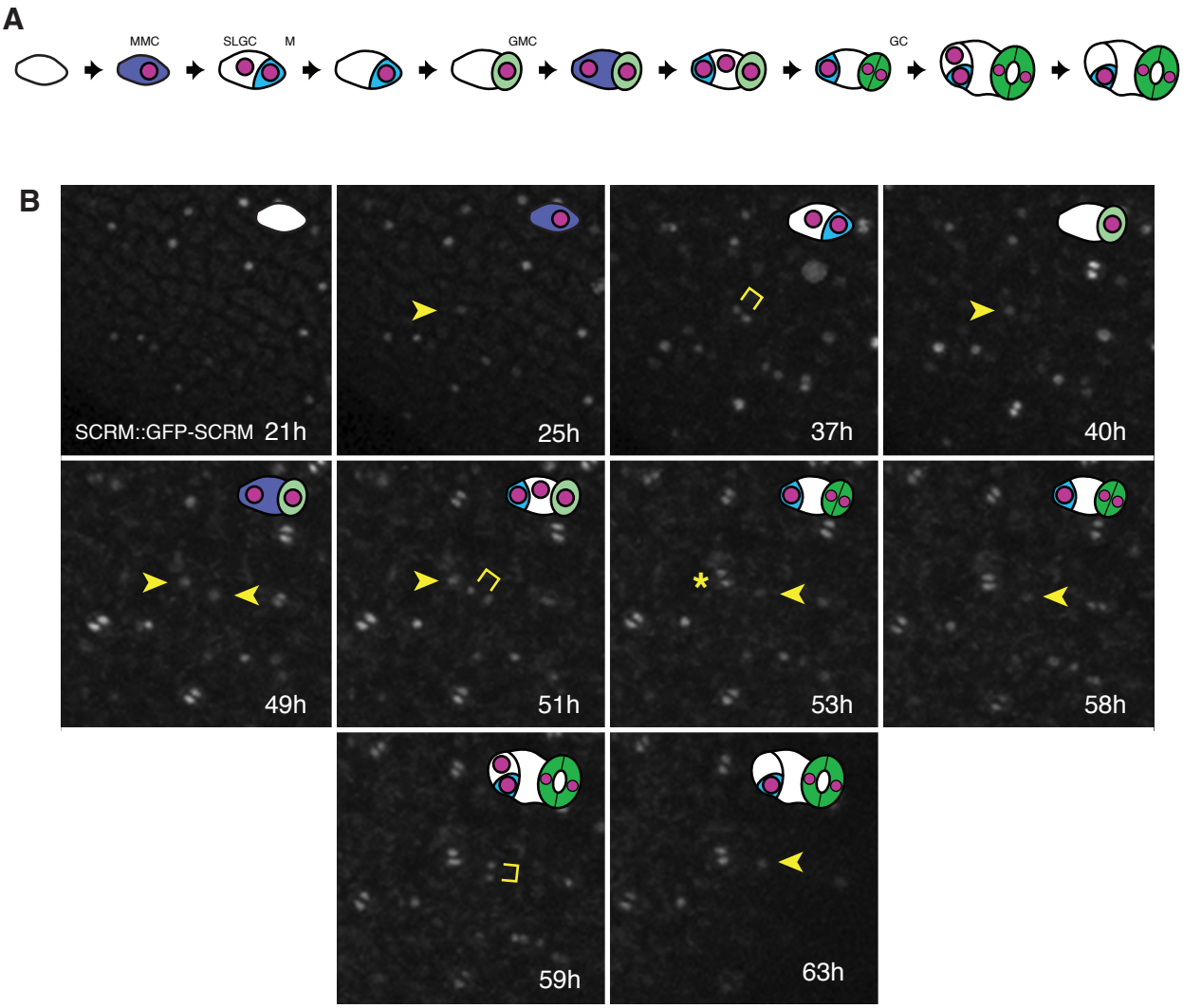
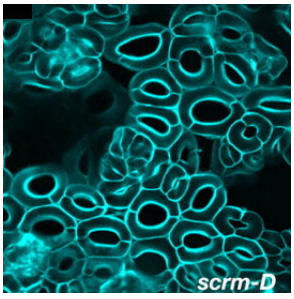


Figure 4-4. *SCRM* is expressed during all cell states in the stomatal lineage.
(A) Detailed developmental diagram of expected SCRM occurrence (pink) during stomatal development. See also Figure 1-2. Meristemoid mother cells (MMC) express *SCRM* upon acquiring asymmetric division capability; when they divide, both daughter nuclei briefly retain SCRM protein but only the meristemoid (M) continues to produce it. Meristemoids (M) express *SCRM* throughout asymmetric division and the transition to guard mother cells (GMC). GMC also maintain a high level of SCRM protein through their symmetric division and differentiation into mature guard cells (GC). Non-meristemoid sister cells from asymmetric divisions are termed stomatal lineage ground cells (SLGC) and can resume stomatal lineage fate as MMC during their early development, in which case they express *SCRM* throughout stomatal development as previously.
(B) Time series of *SCRMpro::GFP-SCRM* in a germinating cotyledon. Cell diagrams show comparable stages from (A). Arrowheads, *SCRM*-expressing nuclei of interest; brackets, asymmetrically divided nuclei; asterisk, symmetrically divided nuclei forming a stoma.

Figure 4-5.



scrm-D (Kanaoka et al. 2008)

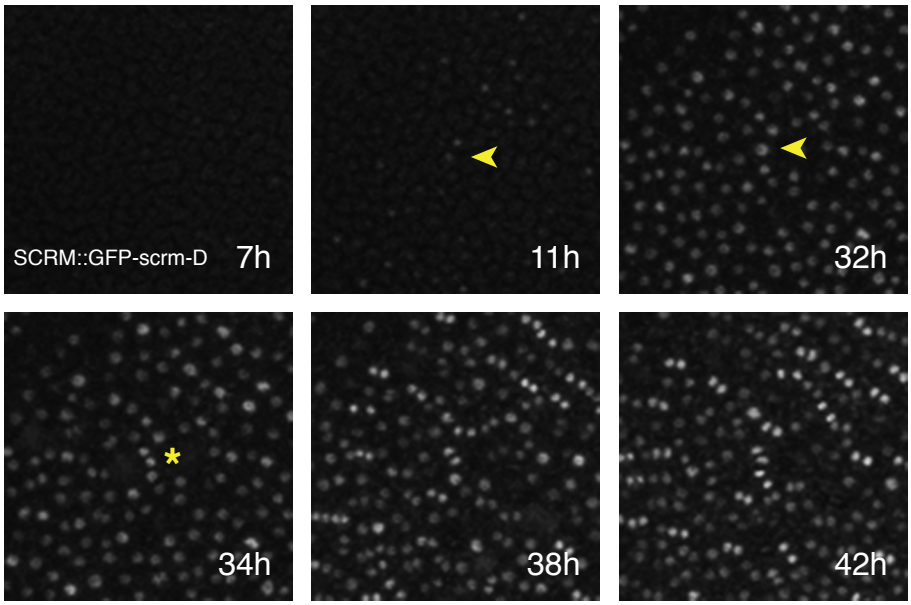
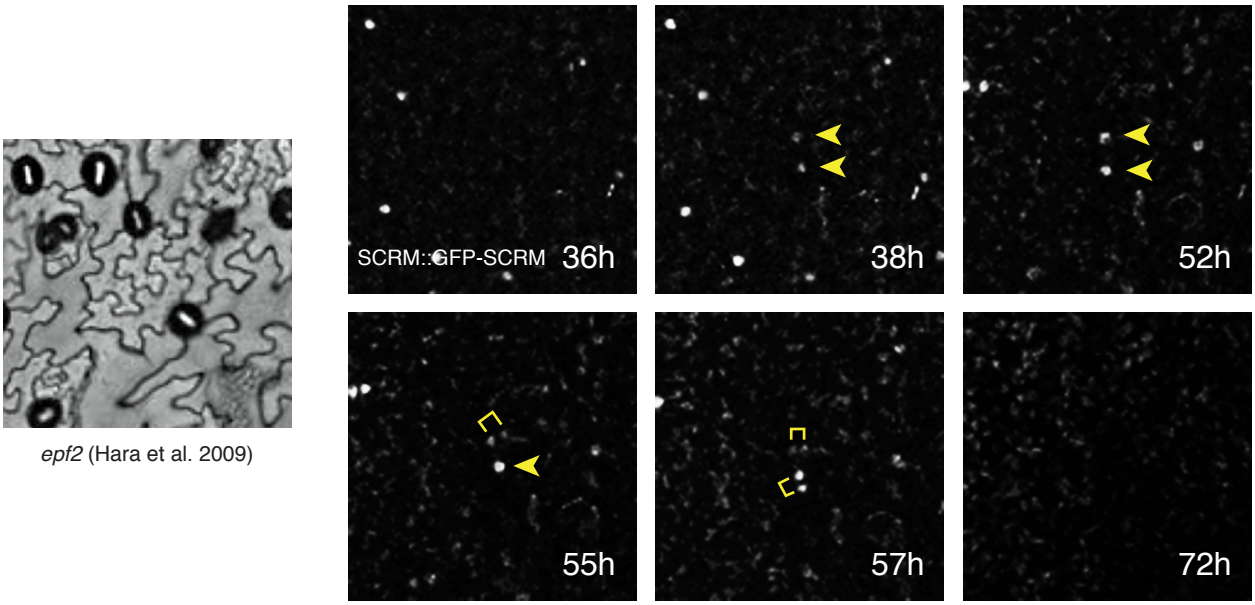


Figure 4-5. Stabilizing mutation *scrm-D* converts nearly all epidermal cells to stomata.
SCRMpro::GFP-scrm-D arises in nearly all epidermal nuclei rapidly after germination: faintly as early as 7 hours and clearly by 1-2 hours before SPCH or SCRM reporters are visible in the wild type. Cells positive for GFP-scrm-D quickly differentiate into guard mother cells (GMC) and then divide to form guard cells. By 42 hours, nearly all epidermal cells have become stomata. Arrowheads, nucleus of interest; asterisk, symmetrically divided nuclei forming a stoma.

Figure 4-6.



epf2 (Hara et al. 2009)

Figure 4-6. Loss of function in signaling components changes tissue patterning and expression of stomatal regulatory genes.
SCRMpro::GFP-SCRM marks stomatal-lineage nuclei. In the *epf*2 mutant background, SCRM-positive cells sometimes arise adjacent to one another, both divide asymmetrically, and then lose stomatal-lineage identity. Such a sequence would reasonably produce the *epf*2 mutant phenotype of small divided cells which then dedifferentiate into wavy-edged pavement cells (left).

Figure 4-7.

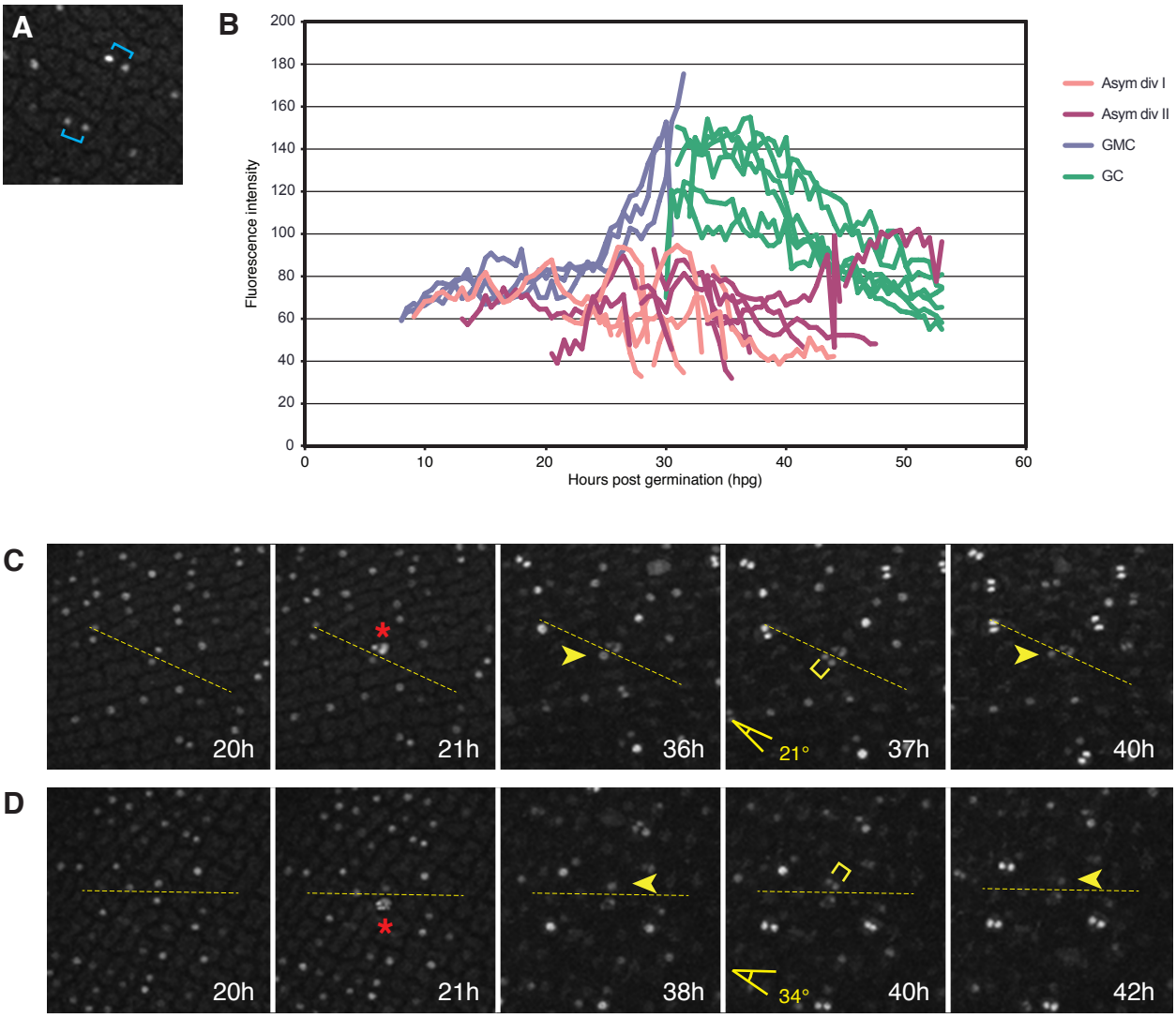


Figure 4-7. Laser ablation of SCRM-positive cells disrupts cell polarity.
(A) Adjacent pairs (brackets) of *SCRMpro::GFP-SCRM* nuclei in germinating cotyledon epidermis.
(B) Fluorescence intensity of paired 28 *SCRMpro::GFP-SCRM* nuclei over time, color-coded by cell fate. Asym div I represents nuclei dividing asymmetrically with a neighbor that also divided asymmetrically, while Asym div II nuclei had no asymmetrically dividing neighbors. GMC fluorescence tends to begin earlier and have higher intensity than asymmetrically dividing nuclei, while GC fluorescence is high (but declines artificially in later time points due to alignment required for cell tracking). Nuclear tracking and intensity measurements from SVCell software.
(C, D) Representative laser ablation experiments performed at 20h. In each, adjacent *SCRMpro::GFP-SCRM* nuclei were selected and the higher-fluorescing nucleus, a presumed GMC, was ablated. Neighbor cells divided 16-18 hours later at acute angles to the plane between the original two cells, creating a meristemoid adjacent to the ablated cell. Typical unablated division angle is 60°. Dashed line, cell-cell planes; red asterisks, ablated cells; arrowheads, neighbor nuclei; brackets, asymmetric divisions.

Chapter 5.

De novo stomatal pattern formation in the Arabidopsis embryo: a preliminary study.

Introduction

In Chapter 4, the results from a new method for time lapse in seedlings demonstrated that the expression of stomatal regulatory genes generally follows existing models based on genetic experiments and steady-state expression patterns. However, the time-lapse experiments also yielded entirely unexpected results. For instance, SCRM-positive cells were found to arise in pairs, and their fate and polarization were shown to depend on their neighboring cells. In this chapter, another unexpected result is explored further: some SCRM-positive cells in the germinating cotyledon assume GMC identity and divide symmetrically to form guard cells without a prior asymmetric division. This observation dramatically contradicts the current model of stomatal development, in which asymmetric division is absolutely required for entry into the stomatal lineage. How can these contradictory findings be reconciled?

In order to resolve this discrepancy, we hypothesized that stomatal patterning might begin in the embryo. The "green seeds" of some plants have been shown to carry out photosynthesis (Flint & Moreland, 1943), and some of these have integuments or inner silique surfaces with stomata to facilitate gas exchange (Jernstedt & Clark, 1979). Fully developed stomata in the sporophyte embryo are known as well, even in dry seeds, such as *Ceratonia siliqua* (Christodoulakis et al., 2001). Although *Arabidopsis* does not possess embryonic

stomata, these examples show that stomatal patterning may be initiated during embryonic development even in cases where seed dormancy occurs.

This chapter presents considerable preliminary evidence that stomatal development in *Arabidopsis* also begins in the embryo. Stomatal master regulatory transcription factors *SPCH* and *MUTE* are expressed during embryonic development, as is their partner *SCRM*. In contrast, we found no expression of *FAMA*, which governs the final symmetric division and terminal differentiation of guard cells. Genes controlling correct spacing of stomatal precursors, *TMM* and *EPF2,* were also verified at this stage by means of transcriptional reporters, and *MUTE* was accordingly expressed in a nonrandom pattern similar to that found in postembryonic stomatal development. *SCRM*-, *SPCH*- and *TMM*-expressing cells divided asymmetrically in the embryo, consistent with entry into the stomatal lineage. All indications so far point to the full regulatory program initating stomatal patterning in the cotyledon during embryonic development, prior to seed dormancy and germination. This resolves the contradiction of stomata in the germinating cotyledon appearing to perform an abbreviated version of stomatal development: de novo epidermal tissue patterning occurs in the embryonic epidermis.

Results

Post-germination SCRM-positive cells may become stomata directly

Initially, *SCRMpro::GFP-SCRM* was examined through time lapse in germinating cotyledons (see Chapter 4). Although SCRM is expected to be present throughout stomatal development, including the initial asymmetric entry division (Figure 5-1A, see Chapter 1), some cells in germinating cotyledons

became stomata after a single symmetric division (Figure 5-1B,C). *SPCHpro::SPCH-GFP* cells either divided asymmetrically or faded away (see Chapter 4), which implies that some SCRM-positive cells did not express *SPCH* after germination. It is possible that increased levels of functional SPCH protein might cause additional asymmetric entry divisions in *SPCHpro::SPCH-GFP*, but direct GMC differentiation after germination was seen in multiple genetic backgrounds: a subset of POLAR-positive cells became GMC without asymmetric division (see Figure 2-S5), as did a subset of epidermal cells in the wild-type background expressing the plasma membrane marker Q8 (Figure 5-S1).

However, the *spch* mutant phenotype has no stomatal lineage cells (MacAlister et al., 2007; Pillitteri et al., 2007), indicating that *SPCH* is absolutely required for the initiation of stomatal development. Therefore, *SPCH* must have been expressed in these cells, and asymmetric division must have occurred, at some earlier stage. We next set out to identify the developmental window during which this hypothesized earlier expression occurs.

Germination is directly preceded by seed dormancy, a desiccated state in which mitosis is unlikely. Prior to seed dormancy, the sporophyte embryo develops inside the ovule. Embryonic development is defined by active cell division, organ growth, and the establishment of the body plan. We decided to examine the expression of stomatal regulatory transcription factors during embryonic development.

Stomatal development genes are represented in embryo transcriptomes

Using available gene expression data, we investigated whether genes known to be involved in stomatal development were expressed during

enbryonic development (Figure 5-2). Epidermal identity gene *AtML1* is expressed throughout embryonic development and upregulated beginning with the globular stage, when it becomes restricted to the L1 layer. This accords with published reporter studies (Takada & Jürgens, 2007). *ERECTA* is also expressed in all early epidermal cells, and its expression is very high starting with the globular stage. Intriguingly, *BASL* was expressed at relatively high levels throughout embryonic development. *EPF2* and *POLAR*, both of which are associated with MMC and meristemoid division, were considerably upregulated in the bent-cotyledon stage, indicating likely stomatal development at that time. *TMM* and *STOMAGEN* were also increased at that stage, though to a lesser degree. No strong effect was evident in *SPCH*, *MUTE*, or *FAMA*, likely because of the low expression levels effective for transcription factor activity. The uniformly low level of *EPF1* in the embryo may indicate that only earlier stages of stomatal development occur, as *EPF1* is expressed by late stomatal precursors.

Knowing that epidermal gene expression begins at the globular stage and that stomatal gene expression in the embryo is strong at the bent-cotyledon stage, we focused our efforts on documenting stomatal gene expression in the heart through mature stages through live imaging of fluorescent reporters.

Stomatal regulatory genes execute known developmental processes in the embryo

The developmental series of *Arabidopsis* embryo stages used in these time-lapse experiments is shown in Figure 5-3A. Note that, because substantial development occurs between the torpedo and mature stages, we assessed the walking-stick and bent-cotyledon stages separately. Figure 5-3B depicts the

method by which embryos are removed from green siliques and mounted in ovule culture media for imaging (see Materials and Methods).

My efforts to conduct time-lapse microscopy on developing embryos were facilitated by collaboration with the Higashiyama ERATO Live-Holonics Project in Nagoya University (http://www.liveholonics.com/en/). I observed stomatal reporter lines in the embryo over 24 hours of development using a confocal spinning-disc microscope. These preliminary results reflect at least three surviving embryos per reporter genotype. Confirming these results with a larger number of embryos will be necessary in the future.

Embryos from parent plants homozygous for *SCRMpro::GFP-SCRM* showed GFP signal in epidermal nuclei, starting in the late torpedo to early walking-stick stage of development (Figure 5-4A). Many SCRM-positive nuclei divided asymmetrically (Figure 5-4B), consistent with initial entry divisions into the stomatal lineage. In the future, I will obtain a larger sample size and determine division frequency and polarity.

Nuclei positive for *SPCHpro::SPCH-GFP* were found as early as the late torpedo stage (Figure 5-5A). SPCH-positive cells also divided asymmetrically in the embryo (Figure 5-5B, Movie 5-1). SPCH is the earliest known determinant of the stomatal lineage (see Chapter 1) and is specific to asymmetrically dividing stomatal-lineage cells, so it is likely that stomatal development begins concurrent with the expression of SPCH. Generally, GFP becomes visible in these cells two to three hours prior to asymmetric cell division (Figure 5-5C).

MUTEpro::nucYFP shows *MUTE* promoter activity starting in the bent-cotyledon stage (Figure 5-6A), which agrees with the results from the protein fusion, *MUTEpro::MUTE-GFP* (Figure 5-6B). *MUTE*-positive cells do not divide

in the embryo, consistent with the single-cell M-to-GMC cell-state transition (Figure 5-6C, Movie 5-2). Strikingly, the characteristic nonrandom distribution of stomata across the cotyledon epidermis is reflected by both the MUTE-GFP protein (Figure 5-6D) and the transcriptional reporter (data not shown). In the future I will use a cell-membrane reporter or cell-wall stain to determine the number of cells per MUTE-positive cell and validate the stomatal one-cell spacing rule in this context.

Although the parent plants were confirmed homozygous and expressing the reporter appropriately, *FAMApro::FAMA-GFP* signal was never observed in embryos (n=20 mature embryos, Figure 5-7). This finding is consistent with the fact that symmetric division did not occur in SCRM- or MUTE-positive cells.

Stomatal signaling gene promoters are active in the embryonic epidermis

Based on the nonrandom spacing of MUTE-positive cells, embryonic stomatal development involves appropriate patterning. I next examined whether known stomatal spacing signals were present in the embryo. EPF2 is a small cysteine-rich peptide produced by stomatal precursors, which suppresses stomatal identity in the surrounding cells (see Chapter 1 for review; Hara et al., 2009; Hunt and Gray, 2009). *EPF2pro::erGFP* was visible in embryos beginning at the bent-cotyledon stage (Figure 5-8A). TMM, a leucine-rich repeat receptor-like protein, binds EPF2 (Lee et al. 2012), and its loss of function results in stomatal clusters. *TMMpro::GUS-GFP* appears in the late torpedo stage (Figure 5-8B), and TMM-positive cells divided asymmetrically during embryonic development (Figure 5-8C). This implies that the stomatal spacing signals known from postembryonic development also determine stomatal patterning in the embryo.

Discussion

Finding that a subset of cells in germinating cotyledons became guard mother cells and then stomata without asymmetric divisions (Figure 5-1), we investigated previous possible times that stomatal development might have occurred. Embryo transcriptome data showed that stomatal regulatory genes were expressed during embryonic development (Figure 5-2). In particular, *POLAR*, which is highly specific to stomatal development (see Chapter 2), and *EPF2* were differentially expressed in the bent-cotyledon stage.

Thus, we hypothesized that stomatal development might start in the embryo, using the same genes and following the same trajectory as in other shoot organs. We further hypothesized that development could be paused by seed dormancy to resume in the germinating cotyledons. I explored these hypotheses by investigating embryonic development in real time, which was possible thanks to the expertise of Dr. Daisuke Kurihara and the generosity of Professor Tetsuya Higashiyama of Nagoya University, as well as the University of Washington Biology Department's Washington Research Foundation-Benjamin Hall Fellowship, which supported my work in Nagoya during one academic quarter.

I found that stomatal master regulators are expressed in embryos in a manner consistent with the known mechanisms of stomatal development: SCRM-and SPCH-positive nuclei undergo asymmetric division (Figures 5-4 and 5-5), and MUTE-positive nuclei appear at later developmental stages and do not divide (Figure 5-6). In contrast to cotyledon development (see Chapter 4), symmetric divisions were never observed in SCRM- or MUTE-positive nuclei, implying that stomata develop only partially before seed germination in *Arabidopsis*.

It will be helpful to establish exact developmental timing for the expression of each master regulator. *SPCH* and *SCRM* are expected to coincide, since they require each other for function in stomatal development (Kanaoka et al., 2008). Basic helix-loop-helix transcription factors require a dimerization partner to bind DNA, and SPCH cannot dimerize with itself (Kanaoka et al., 2008). To evaluate this hypothesis, I plan to image SPCH and SCRM transcriptional reporters as well, to show the timing of promoter activity. Double reporters (e.g., TagRFP-SCRM/SPCH-GFP) will illuminate the developmental timing best, and can also address the important issue of whether both components are present in asymmetrically dividing cells.

Development in the embryo may differ in pace from postembryonic development. On average, the appearance of embryonic SPCH-GFP occurs only 2-3 hours before asymmetric division (Figure 5-5C). Embryonic development does occur at an increased pace compared to later growth processes, but expressing a critical transcription factor so soon before division is somewhat surprising, considering that SPCH expression in germinating cotyledons is visible twelve or more hours before division (see Chapter 4). Because a low SPCH protein level may have important effects before the translational fusion is visible under these conditions, the timing of *SPCH* promoter activity will also inform this question.

MUTE-positive nuclei appear at a later embryonic stage than SPCH-GFP, corresponding to their order of activity during stomatal development. So far, MUTE is the only stomatal gene that I have examined with both a transcriptional and a translational reporter. *MUTEpro::nucYFP* fades over embryonic development (Figure 5-6C, Movie 5-2); while this does not correspond to the

dynamics of MUTE protein, it does indicate that MUTE promoter becomes active for a specific length of time in each meristemoid. The lower signal level of *MUTEpro::MUTE-GFP* may not allow the most sensitive detection of early expression, but it does capture the degradation rate of a protein containing the full MUTE sequence. The combination of the two reporters should effectively capture both gene regulation and protein dynamics, and I plan to compare them thoroughly in the future.

The fate of *MUTE*-expressing cells which have had their stomatal-lineage identity repressed is intriguing. In Chapter 4, MUTE-positive cells were shown to lose *MUTE* expression in rare cases, though those cells were not followed to maturity in a fully expanded leaf. The question arises of whether such a cell might later resume MMC fate like an SLGC, which was presumably repressed in the same manner, or whether it has progressed too far by explicitly stopping its own asymmetric divisions with *MUTE*.

No FAMA-positive cells were found before germination (Figure 5-7), consistent with the absence of SCRM- and MUTE-positive symmetric divisions in the embryo. This leads to an intriguing question: must *MUTE* be reactivated in cells after germination, before symmetric division occurs? *MUTE* is required for the cell-state transition from asymmetrically dividing meristemoid to symmetrically dividing GMC, but cells in the germinating cotyledon express *MUTE* around 18 hours after germination, before division occurs, and then go on to form stomata. This could be addressed by imaging MUTE with a plasma membrane marker, to determine whether any stomata form without postembryonic MUTE expression, or by imaging MUTE-FAMA dual-color

reporters to determine whether they are always sequential in the germinating cotyledon.

Thus far, stomatal signaling genes appear to be expressed in the embryo as expected if the early stages of the known program of stomatal development are being followed (Figure 5-8). Further investigations could examine the expression of *EPF1*, which codes for a signaling ligand produced by GMCs, and the ERECTA family of receptor kinases, which are receptors for EPFs (Lee et al., 2012). Based on the embryonic transcriptome results, *EPF1* may not be expressed until after germination, but I expect to find *ERECTA* expression throughout the epidermis and *ERL1* in the stomatal lineage.

The finding that stomatal development begins in the embryo, while postgermination development resumes midway through the stomatal development sequence, implies that cell states persist through seed dormancy. If the *Arabidopsis* cotyledon after germination is carrying out halted developmental programs and expanding on prepatterning that was laid down previously, what are the options for seeing de novo stomatal pattern formation in real time? Robinson et al. (2011) observed the process in young rosette leaves, but there are several drawbacks: leaves are relatively difficult to access at very early stages, the developmental gradient along the length of the leaf is a complicating factor, and the picture is quite literally occluded by trichome development. The embryo, despite requiring dissection and special care, is actually a very tractable system for looking closely at the coalescence of pattern from an unpatterned field of cells. Many additional hypotheses could be explored using this system. For instance, *SCRM* and *SPCH* may overlap in expression to determine initial patterning of stomatal precursors. Investigating this in the cotyledon may not

give an accurate picture of the primary events, due to the embryonic prepatterning we have demonstrated in this study, but the embryo is virgin territory.

Furthermore, the retention of developmental cell state through seed dormancy implies that the "pairs" of SCRM-positive cells identified in chapter 4 may reflect M/SLGC cell pairs produced by embryonic asymmetric divisions. In fact, in the highly autofluorescent early stages of germination, such cell pairs frequently appear to have different sizes and shapes (see ablated pairs in Figure 4-7C,D). It may be possible to verify that conclusion using the embryo culture system. The cells expressing *MUTE* in the embryo must logically be the same ones that become stomata after germination without dividing asymmetrically, because *SPCH* is required for stomatal development but does not seem to be expressed in postgermination GMCs. I found that in a few cases during embryo time lapse, mature embryos seemed to lose GFP signal and become more autofluorescent in a way reminiscent of cotyledons at germination; this may have been seed dormancy. Had I followed those same embryos through a dormant state while still mounted, it would have been easy to tell whether the MUTE-positive nuclei before dormancy were still MUTE-positive after it. The ovule culture system, when mantained for long periods, can result in mature plants (Kurihara, pers. comm.), although dissected embryos are more fragile than those still protected in ovules.

Another way of addressing this question would be to use *Arabidopsis* seed-dormancy mutants. We have acquired mutants of *leafy cotyledon1 (lec1)*, *lec2*, and *fusca3*, which are defective in seed maturation and dormancy, and made crosses with stomatal bHLH reporter lines (currently at the F1 generation). The

homozygous mutants develop directly, without entering seed dormancy, so their stomata should also develop directly. Starting with a mature embryo, then, we would expect to see *FAMA* expression during time lapse.

It is fascinating to speculate on how developmental progress might be carried over from embryo to germinating seedling despite dormancy and seed desiccation. The two main candidates would seem to be chromatin state and protein or RNA retention. The early appearance and high levels of *SCRMpro:: GFP-scrm-D* in germinating cotyledon epidermis provide evidence that at least some proteins are retained (see Figure 4-5), and this could be further assessed using photoconvertible fluorophores attached to MUTE and SCRM.

Given the evidence of Chapter 4 and the robust expression of *MUTE* in the embryo, it is clear that *MUTE* is not a stomatal "point of no return." Although stomatal prepatterning occurs during embryonic development, postgermination application of peptides can result in no stomata (Lee et al., 2012), despite *MUTE* already being expressed in some cells. The next obvious candidate for irrevocable stomatal identity would be *FAMA* expression. However, *FAMA* is not required for symmetric division; *fama* mutant GMCs divide symmetrically to excess rather than failing to divide (Ohashi-Ito and Bergmann, 2006). An unknown factor downstream of *MUTE*, which induces symmetric division and perhaps *FAMA* expression, may be the true point of no return. Finding MUTE's direct targets through chromatin immunoprecipitation would help lead us to the missing link.

If MUTE induces the next, irrevocable step, though, how can stomatal development pause before seed dormancy? Because cells in the embryo express *MUTE* at different times and *MUTE* expression fades away asynchronously (see Movie 5-2), it is possible that a cell-autonomous "holding" mechanism exists.

Repression of downstream genes by an embryo-specific mechanism could serve this role. One possibility is epigenetic *FAMA* repression: in dry conditions, the *SPCH* and *FAMA* loci can undergo RNA-directed de novo methylation (Tricker et al., 2012), allowing developmentally plastic reduction of stomatal density. A similar mechanism could silence *FAMA* or other downstream genes in *MUTE*-expressing embryonic cells, and that repression could be removed immediately after germination. One way to investigate this would be by observing stomatal development in methylation and demethylation loss-of-function mutants. For instance, plants lacking methylation capability might express *FAMA* or even develop stomata in the embryo, while plants lacking demethylation might exhibit arrested GMC following germination. No stomatal phenotype has been reported in such mutants to date, but its detection would require close attention to phenotype and developmental timing.

Our results imply that, despite the different developmental origins and morphologies of cotyledons and true leaves, the development of cotyledon epidermis closely resembles the development of leaf epidermis. However, the cotyledon version of the epidermal development program does not include trichomes, which are very much present on *Arabidopsis* leaves. Trichomes may be specifically excluded from the cotyledon; indeed, *lec1* and other seed dormancy mutants exhibit leaflike cotyledons including trichomes (e.g., West et al., 1994), a phenomenon attributed to heterochronic gene expression (To et al., 2006).

In *Arabidopsis*, embryonic stomatal development likely serves to prepare the cotyledon for rapid photosynthetic activity upon emergence, while its arrest keeps the seedling safe from soil-borne pathogens by not forming vulnerable pores too soon. Although albino cotyledon mutants can germinate and grow in

lab conditions with supplemented sugar (e.g., Yamamoto et al., 2000; Albrecht et al., 2006), in nature no readily available sugars means no growth, which could easily lead to overshadowing or death by herbivory. The ability to immediately differentiate stomata during photomorphogenesis would be an advantage, so stomatal prepatterning in the embryo would increase evolutionary fitness.

Materials and Methods

Plant material and growth conditions

The following reporter lines were described previously: *SPCHpro::SPCH-GFP* (Guseman et al., 2010), *MUTEpro::MUTE-GFP* (Pillitteri et al., 2008), *FAMApro::FAMA-GFP* (Pillitteri et al., 2008), *EPF2pro::erGFP* (Hara et al., 2009), *TMMpro::GUS-GFP* (Nadeau & Sack, 2002). Seeds were sterilized using 33% bleach solution (bleach and 0.05% Tween-20) for 15 min and washed five times with sterile water, then plated on 1× or 0.5× Murashige and Skoog media and placed at 4°C for 2 d. Plates were placed under standard long-day conditions (16-h day/8-h night, 21°C), and seedlings were transplanted to soil after 10-14 days.

Embryo culture

Ovule culture media (5% trehalose, 1× Nitsch salts, 1× Gamborg's vitamins, 0.05% MES-KOH, pH 5.8) was sterile-filtered and stored at 4°C. Green siliques were slit open with sterile tools to expose ovules, then transferred to media. While immersed in media, ovules were removed from siliques and embryos extracted with dissecting needles. Embryos were moved into a four-well glass-bottomed dish (Matsunami Glass, Osaka, Japan) by pipetting. The dish was sealed with Parafilm or micropore tape and placed on an inverted confocal microscope for imaging.

Microscopy

Confocal microscopy images were taken with the Olympus CSU-X1 with 488 nm laser and 20x or 40x water objective, or the Zeiss LSM700 (Thornwood, NY, USA) with excitation at 488 nm and a band-pass emission filter of 470 to 500 nm and 20x objective. Z-stacks were made using optimized optical parameters for each system (CSU, 0.2-0.5μm; LSM, 1μm).

Supplemental Movie Legends

Movie 5-1. *SPCHpro::SPCH-GFP* expression in asymmetrically dividing embryonic cotyledon cells. Z-stacks were taken every 30 minutes over 24 hours of development.

Movie 5-2. *MUTEpro::nucYFP* expression in non-dividing embryonic cotyledon cells. Z-stacks were taken every 30 minutes over 24 hours of development.

Acknowledgements

We thank Prof. T. Higashiyama, Prof. M. Kanaoka, D. Kurihara, and Y. Hamamura for providing lab space, embryo protocols, and support during this project, and Prof. C. Queitsch for comments on the chapter. K.M.P. was supported by the Washington Research Foundation–Benjamin Hall Fellowship from the University of Washington Department of Biology. K.U.T. is a Howard Hughes Medical Institute–Gordon and Betty Moore Foundation investigator.

References

Albrecht, V., Ingenfeld, A., & Apel, K. (2006) Characterization of the *Snowy Cotyledon 1* mutant of *Arabidopsis thaliana*: the impact of Chloroplast Elongation Factor G on chloroplast development and plant vitality. *Plant Molecular Biology* 60(4): 507-518.

Christodoulakis, N.S., Menti, J., & Galatis, B. (2001) Structure and development of stomata on the primary root of *Ceratonia siliqua* L. *Annals of Botany* 89: 23-29.

Flint, L.H., & Moreland, C.F. (1943) Note on photosynthetic activity in seeds of the spider lily. *American Journal of Botany* 30(4): 315-317.

Guseman, J.M., Lee, J.S., Bogenschutz, N.L., Peterson, K.M., Virata, R.E., Xie, B., ... Torii, K.U. (2010). Dysregulation of cell-to-cell connectivity and stomatal patterning by loss-of-function mutation in *Arabidopsis CHORUS (GLUCAN SYNTHASE-LIKE 8) Development* 137:1731-1741.

Hara, K., Yokoo, T., Kajita, R., Onishi, T., Yahata, S., Peterson, K.M., Torii, K.U., and Kakimoto, T. (2009). Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR2, in *Arabidopsis* leaves. *Plant Cell Physiol*. 50: 1019–1031.

Hunt, L., and Gray, J.E. (2009). The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Curr. Biol.* 19: 864–869.

Jernstedt, J., & Clark, C. (1979) Stomata on the fruits and seeds of *Eschscholzia* (Papaveraceae). *American Journal of Botany* 66(5): 586-590.

Kanaoka, M.M., Pillitteri, L.J., Fujii, H., Yoshida, Y., Bogenschutz, N.L., Takabayashi, J., Zhu, J.K., and Torii, K.U. (2008). *SCREAM/ICE1* and *SCREAM2* specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* 20: 1775–1785.

Lee, J.S., Kuroha, T., Hnilova, M., Khatayevich, D., Kanaoka, M.M., McAbee, J.M., Sarikaya, M., Tamerler, C., and Torii, K.U. (2012) Direct interaction of ligand-receptor pairs specifying stomatal patterning. *Genes & Development* 26: 126-136.

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

Nadeau, J.A., & Sack, F.D. (2002) Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science* 296: 1697–1700.

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

Robinson, S., Barbier de Reuille, P., Chan, J., Bergmann, D., Prusinkiewicz, P., & Coen, E. (2011). Generation of spatial patterns through cell polarity switching. *Science* 333(6048): 1436-1440.

Takada, S. & Jürgens, G. (2007) Transcriptional regulation of epidermal cell fate in the *Arabidopsis* embryo. *Development* 134: 1141-1150.

To, A., Valon, C., Savino, G., Guilleminot, J., Devic, M., Giraudat, J., & Parcy, F. (2006) A network of local and redundant gene regulation governs *Arabidopsis* seed maturation *Plant Cell* 18(7): 1642-1651.
Tricker, P.J., Gibbings, J.G., Rodríguez López, C.M., Hadley, P., & Wilkinson, M.J. (2012) Low relative humidity triggers RNA-directed *de novo* DNA methylation and suppression of genes controlling stomatal development *Journal of Experimental Botany* 63(10): 3799-3813.
Xiang, D., Venglat, P., Tibiche, C., Yang, H., Risseeuw, E., Cao, Y., Babic, V., Cloutier, M., Keller, W., Wang, E., Selvaraj, G., & Datla, R. (2011) Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in *Arabidopsis*. *Plant Physiology* 156: 346-356.
Yamamoto, Y.Y., Puente, P., & Deng, X.-W. (2000) An *Arabidopsis* cotyledon-specific albino locus: a possible role in 16S rRNA maturation. *Plant & Cell Physiology* 41(1): 68-76.
Yang, M., & Sack, F.D. (1995) The *too many mouths* and *four lips* mutations affect stomatal production in Arabidopsis. *Plant Cell* 7(12): 2227-2239.

Figure 5-1.

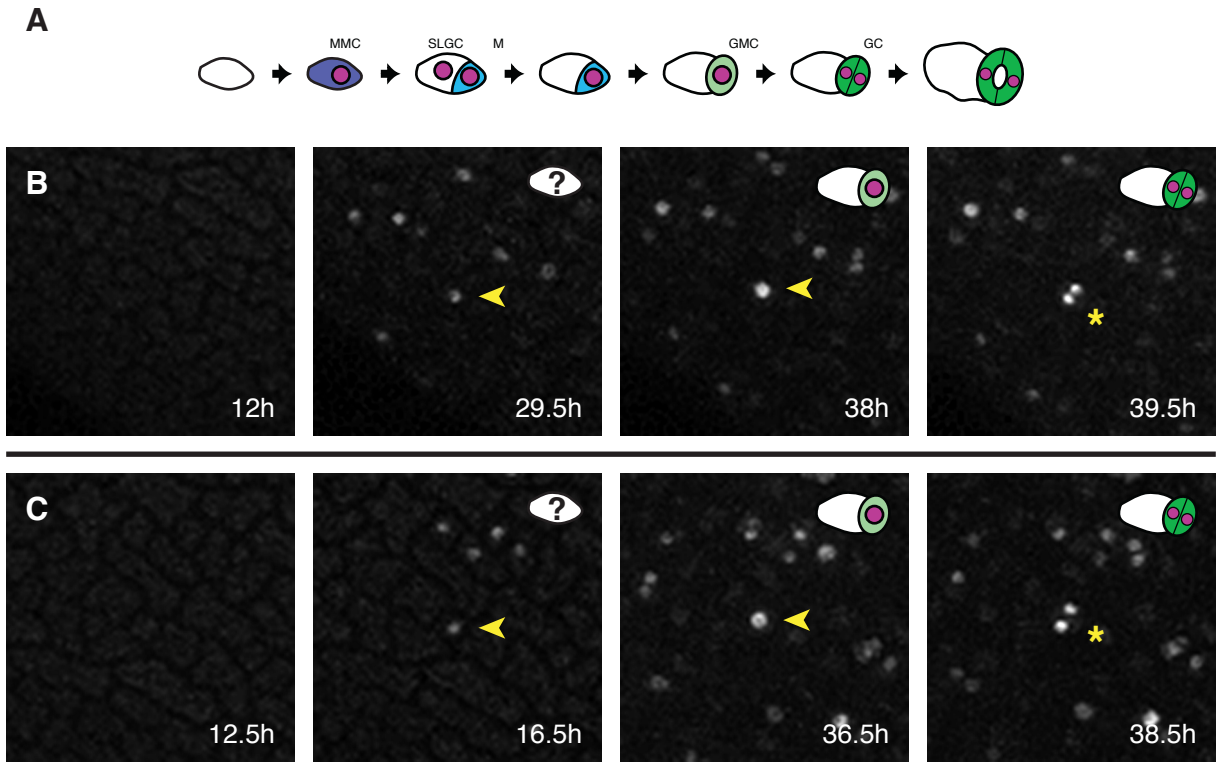


Figure 5-1. A subset of *SCRM*-expressing cells differentiate directly to GMC following germination.
(A) Developmental diagram of known sequence of stomatal development and expected SCRM occurrence (pink) during stomatal development. See Figure 4-4 for representative *SCRMpro::GFP-SCRM* data following the expected progression.
(B, C) Two time series of individual cells starting from germination demonstrate that some *SCRM*-positive cells may differentiate directly into GMC without asymmetric division and produce stomata. Arrowheads, SCRM-positive nuclei of interest; asterisks, symmetrically divided guard cell nuclei.

Figure 5-2.

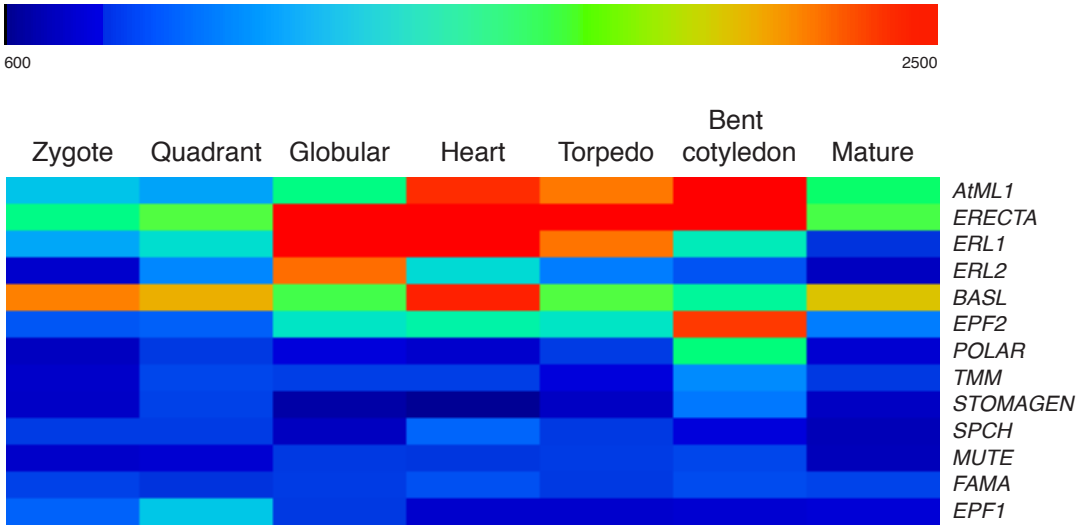


Figure 5-2. Heat map of stomatal development genes during embryonic development.
Based on publicly available expression data from http://www2.bri.nrc.ca/plantembryo (Xiang et al., 2011) .

Figure 5-3.

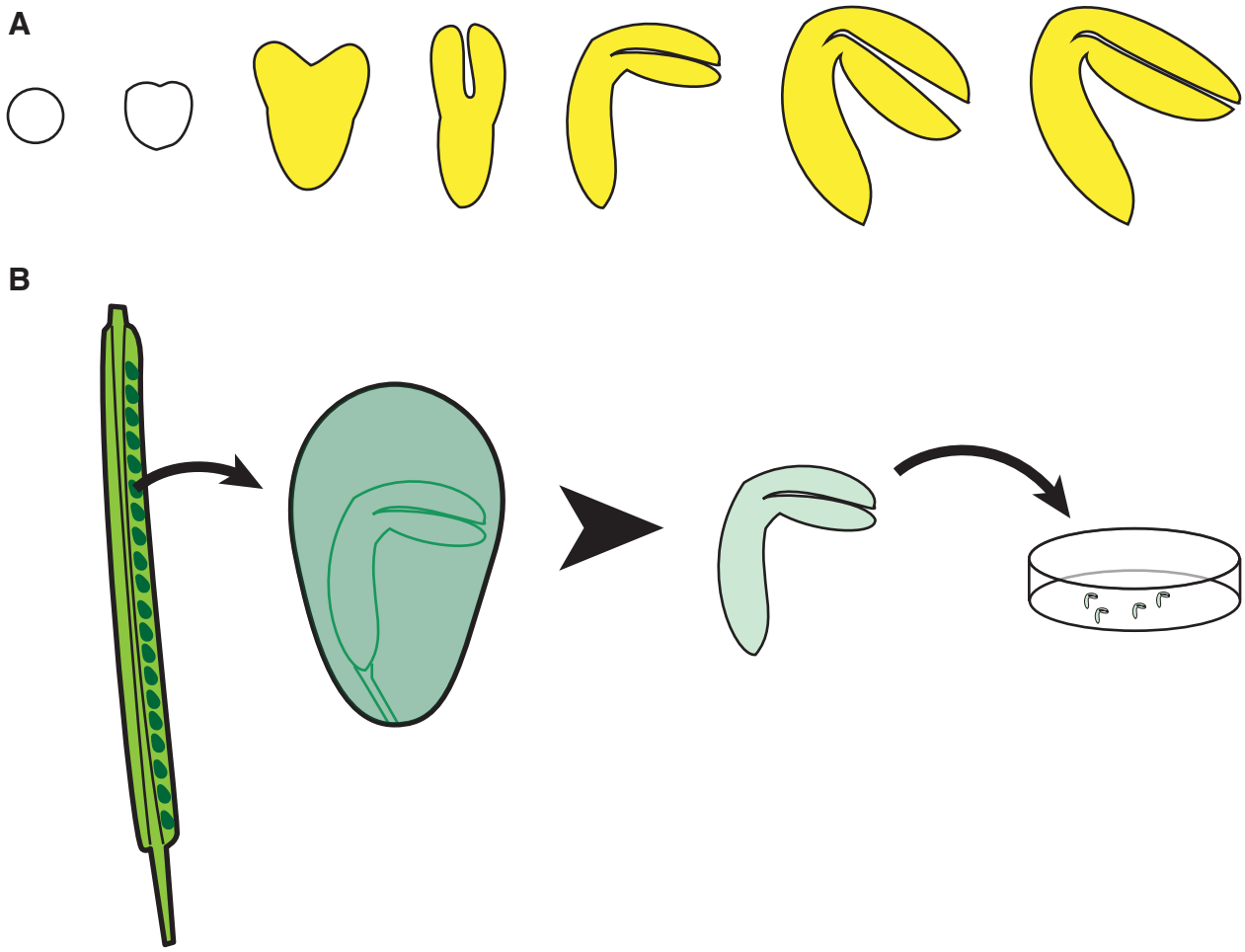


Figure 5-3. Embryo time lapse method.
(A) Diagram of embryo stages in Arabidopsis (not to scale). White fill: globular and early heart stages. Yellow-filled stages were imaged to assay stomatal gene reporters: late heart, torpedo, walking stick, bent cotyledon, and mature stages.
(B) Ovules are dissected from green siliques and embryos removed with dissecting needles. Embryos are pipetted into a glass-bottomed dish and placed on an inverted confocal microscope for imaging.

Figure 5-4.

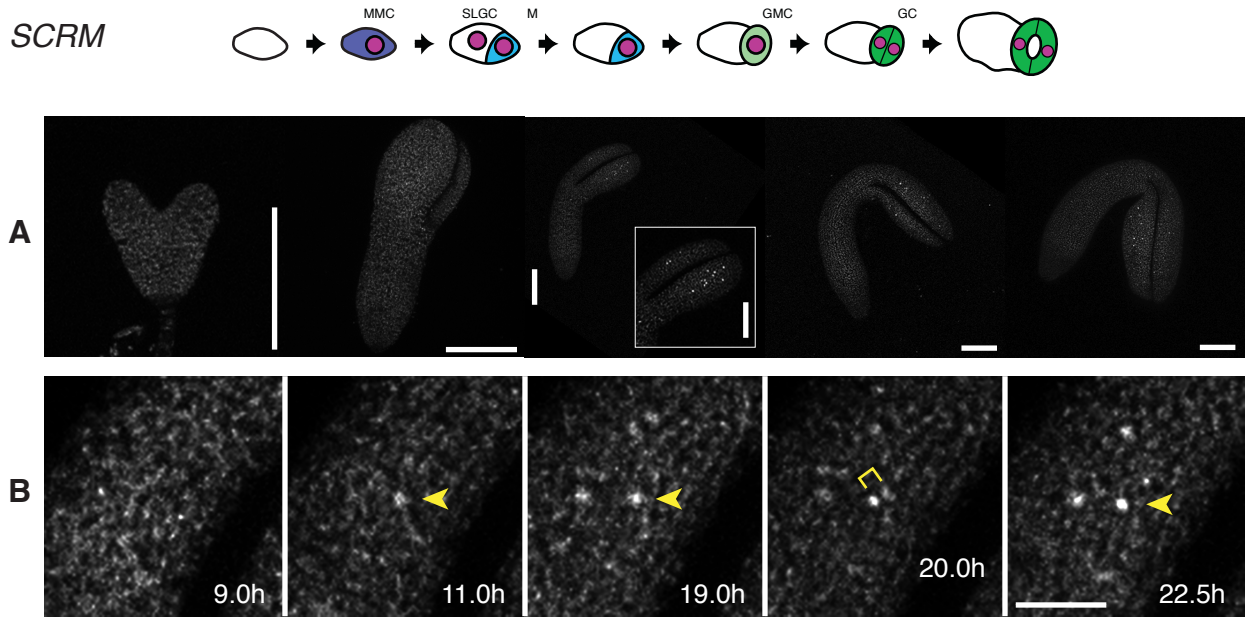


Figure 5-4. SCRM is expressed in asymmetrically dividing epidermal cells in embryos.
Developmental diagram of known sequence of stomatal development and expected SCRM occurrence (pink) during stomatal development.
(A) Series of embryos from a homozygous parent plant demonstrating that expression of *SCRMpro::GFP-SCRM* begins in the late torpedo to early walking-stick stage. Scale bars, 100um except insets, 50um.
(B) Time lapse imaging of walking-stick embryo expressing *SCRMpro::GFP-SCRM* shows appearance of GFP and asymmetric cell division. Times indicate hours elapsed since embryo dissection. Scale bar, 50um.

Figure 5-5.

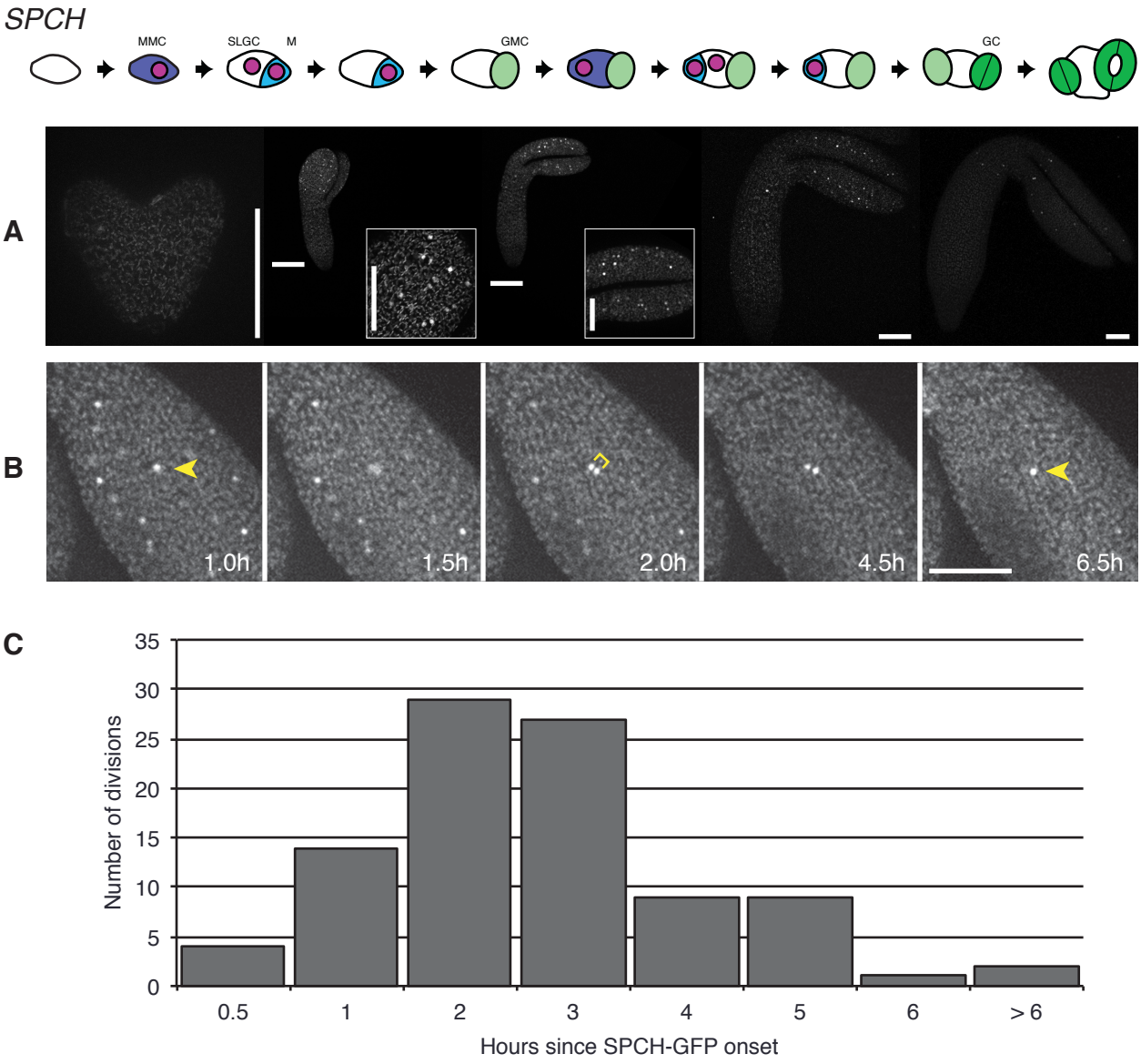


Figure 5-5. SPCH-expressing cells divide asymmetrically in the embryo.
Developmental diagram of known sequence of stomatal development and expected SPCH occurrence (pink) during stomatal development.
(A) Series of embryos from a homozygous parent plant demonstrating that expression of *SPCHpro::SPCH-GFP* is first observed at the late torpedo stage. Scale bars 100um except insets, 50um.
(B) Time lapse imaging of nearly mature embryo expressing *SPCHpro::SPCH-GFP* shows asymmetric cell division. Arrowheads, SPCH-positive nuclei of interest; bracket, asymmetric division. Scale bar, 50um.

Figure 5-6.

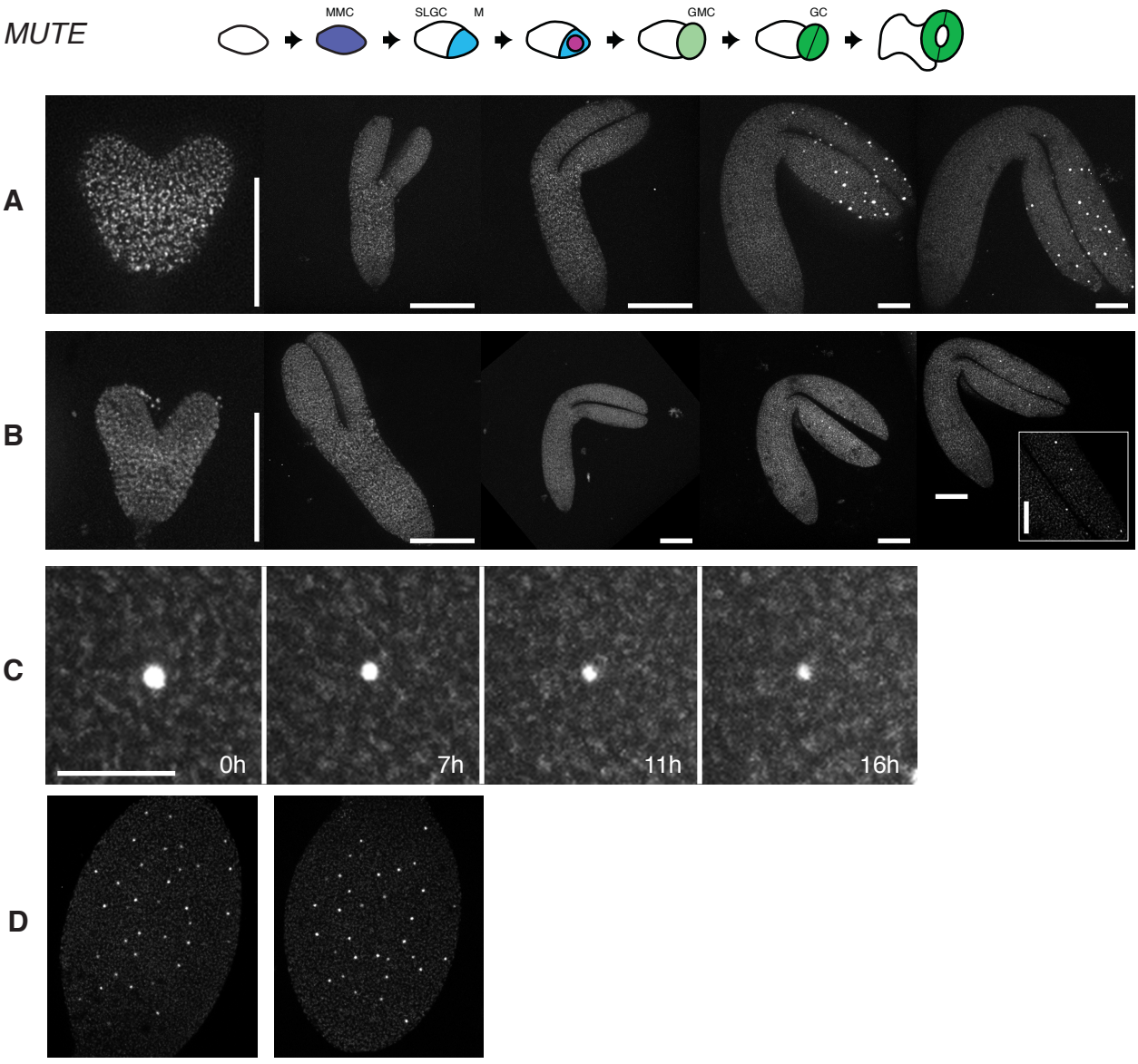


Figure 5-6. MUTE is expressed in single, nonrandomly spaced cells in the embryo.
Developmental diagram of known sequence of stomatal development and expected MUTE occurrence (pink) during stomatal development.
(A) Series of embryos from a homozygous parent plant demonstrating that expression of *MUTEpro::nucYFP* begins at the bent-cotyledon stage. Scale bars, 100um.
(B) Series of embryos from a homozygous parent plant demonstrating that *MUTEpro::MUTE-GFP* is also visible at the bent-cotyledon stage. Scale bars, 100um.
(C) Time lapse imaging of mature embryo expressing *MUTEpro::nucYFP* shows no cell divisions in *MUTE*-positive cells and reporter fluorescence fading over time in the presumed meristemoid-to-GMC transition. Scale bar, 30um.
(D) Detached cotyledons from mature embryos show *MUTEpro::MUTE-GFP* nuclei reflecting the characteristic nonrandom distribution of stomata.

Figure 5-7.

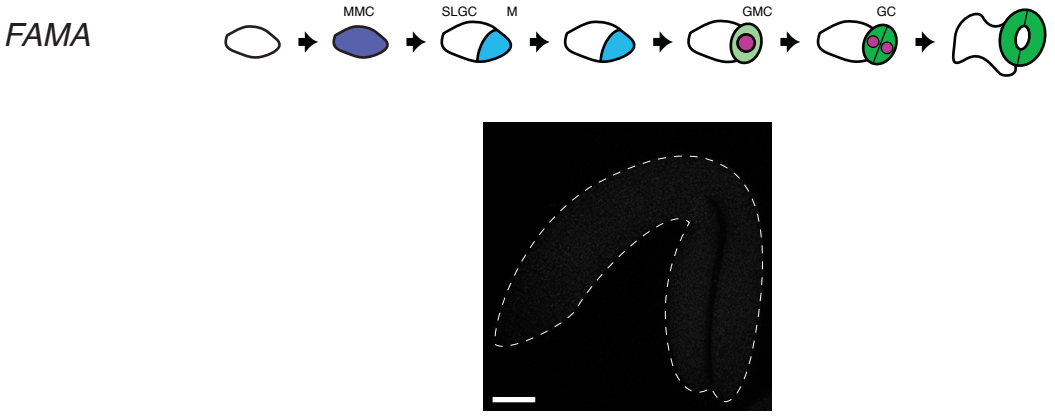


Figure 5-7. FAMA is not found in embryos.
Developmental diagram of known sequence of stomatal development and expected FAMA occurrence (pink) during stomatal development.
Over 20 mature embryos were examined for FAMA expression under conditions that allowed GFP visualization of other embryo reporters and parental *FAMApro::FAMA-GFP* plants. No GFP fluorescence was observed.

Figure 5-8.

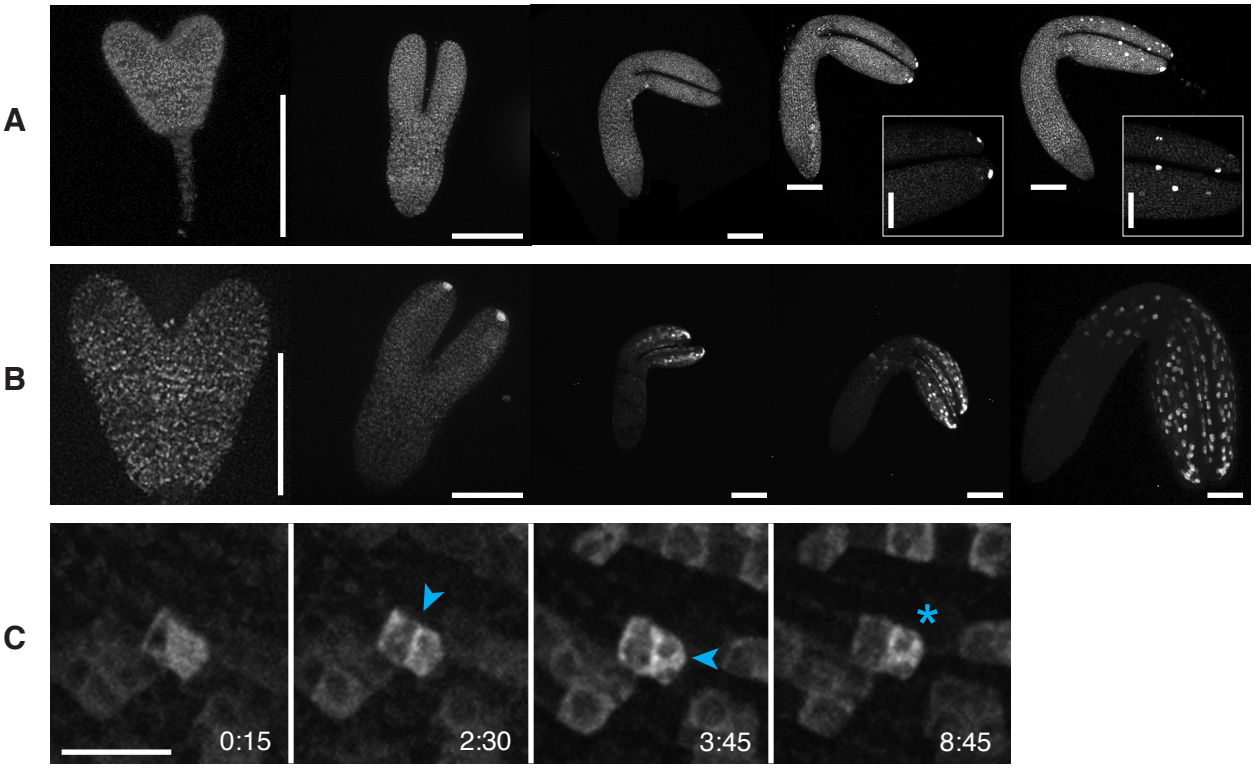


Figure 5-8. Stomatal spacing genes are expressed during embryonic development.
(A) Series of embryos from a homozygous parent plant demonstrating that expression of *EPF2pro::erGFP* begins at the bent-cotyledon stage. Scale bars, 100um. Inset scale bar, 50um.
(B) Series of embryos from a homozygous parent plant demonstrating that expression of *TMMpro::GUS-GFP* begins at the late torpedo stage. Scale bars, 100um.
(C) Time lapse imaging of mature embryo expressing *TMMpro::GUS-GFP* shows asymmetric cell divisions (arrowheads) and continuing higher expression in the meristemoid (asterisk). Scale bar, 30um.

Figure 5-S1.

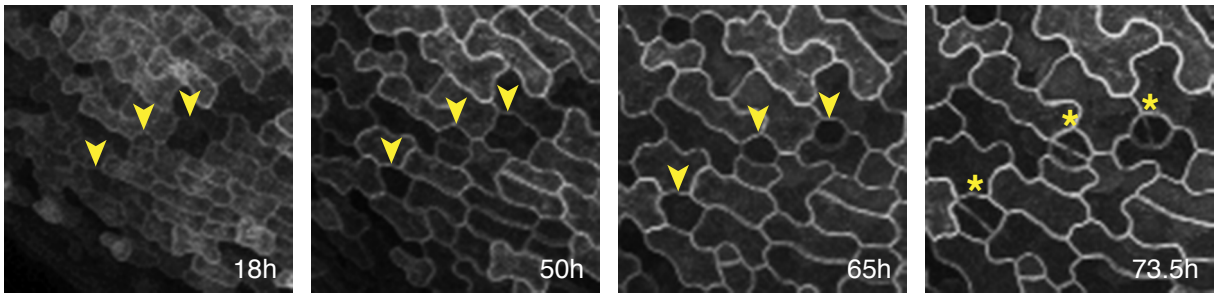


Figure 5-S1. GMC in the wild-type background differentiated without asymmetric divisions.
Time-lapse imaging was performed on cotyledons expressing the plasma membrane marker Q8. Cells do not divide asymmetrically before taking on rounded GMC morphology around 65 hours post-germination. The GMC divides symmetrically to produce a stoma around 73 h. Arrowheads, cells of interest; asterisks, newly formed stomata.

Chapter 6.

Conclusions.

The theme of this dissertation is the concept of cellular differentiation into patterned, functional tissues, which I have chosen to study in the highly tractable system of *Arabidopsis* epidermis. So, how do protodermal cells and their descendants organize themselves and create pattern to produce functional epidermis?

Both the tough, waxy epidermal pavement cells and the controllable stomatal pores are necessary to create a functional epidermis: a cell layer protective against desiccation that nonetheless permits photosynthesis as appropriate to environmental conditions. In *Arabidopsis*, the stem-cell-like meristemoid is critical to the process of stomatal development. Evolutionarily, an asymmetrically dividing meristemoid cell state seems to occur in land plants with wide leaf blades that expand in two dimensions as they develop to maximize light harvesting. Plant stem cells are ancestrally found in the shoot and root meristems, which are sometimes simply asymmetrically dividing apical cells. Because light harvesting leads directly to growth and the ability to make more seeds, meristemoids may have borrowed existing regulatory modules from stem-cell maintenance to support an important function in leaf blade expansion. What aspects of stem-cell nature do plant meristemoids have in common with the shoot and root meristems? In Chapter 2, we found a number of commonalities, including cytokinin response, LRR-RLK receptors, and CLE peptides.

Meristemlike procambium cells evolved to allow expansion of the vasculature, which lets plants form wood and grow to enormous size, again helpful in competing for light. In the future it may become possible to add the transcriptome of vascular cambium cells to our comparisons, but that tissue is embedded within the stem, leading

to considerable technical difficulties in extracting cells for profiling. We were able to enrich meristemoid-expressed gene signatures through an insightful combination of mutants, so perhaps at some point cambium-enriched comparisons might be made similarly.

In Chapter 2, we also examined genes specific to the meristemoid stage of stomatal development, including a previously unknown protein, POLAR. Study of such novel genes, as well as genes newly found to be involved with stomatal asymmetric divisions, like *SCHIZORHIZA*, will allow a deeper understanding of this important process of plant growth. One gene highly enriched in meristemoids was *HDG2*. This transcription factor, closely related to epidermal determinants *AtML1* and *PDF2*, has proven to be important for stomatal development. Based on the mutant phenotype as examined in Chapter 3, we can see that, in the absence of just one meristemoid-expressed gene, tissue patterning is disrupted. Although the phenotype has low penetrance, it can be enhanced by eliminating *AtML1*. These facts tell us that the program of stomatal development is highly tuned to produce a desired result and that there is redundancy in the mechanism, so development can be both repeatable and robust.

The *HDG2* study also allows us to evaluate the introgression of tissue-specific cell types into a neighboring tissue. Stomata in the mesophyll have no apparent function; in fact, they have likely converted existing cells that could instead have performed photosynthesis. The essential nature of a tissue is to contain different cell types in a spatial arrangement that allows them to work together, like circuit boards or clock gears. Placing a gear on a motherboard yields either no result or damage. The question of how tissues hierarchically govern the cell types they require is fascinating and, at least in this case, not entirely spatially determined.

There is also a question of mixed cell state in HDG2-OX. The cells showing epidermal endoreduplication markers in the mesophyll could be entering a pavement-cell state, or they could be endoreduplicating as typical mesophyll cells with superfluous epidermal characteristics. Some TMM-positive mesophyll cells in HDG2-OX had autofluorescent chloroplasts, strongly implying mixed mesophyll and epidermal traits. When the regulatory hierarchy is violated, cells seem able to activate multiple developmental programs at once. What would have happened to those cells over time remains unclear; does a correcting mechanism exist?

Further studies of *HDG2* should take into account its link with cell-cycle genes. Preliminary results (not shown) of expression levels in several cyclin and cyclin-dependent genes under *HDG2-OX* seemed to show *HDG2* affecting cell cycle genes, and the knockout phenotype's low but consistent production of single guard cells implies a role in cell division as well. Stomata differ from other epidermal cell types in that they do not endoreduplicate their genomes, which may provide a unique link between stomata and the cell cycle in the epidermal layer. It would also be interesting to investigate determine how *HDG2*, despite its high level of meristemoid enrichment, is required for the maturation of other epidermal cell types, as it has been shown to affect trichome development as well.

Live imaging can provide a new understanding of developmental processes, particularly stomatal development, which is conveniently located on the shoot epidermis. In Chapter 2, the translational fusion protein POLAR-GFP was shown to relocalize dynamically during the process of asymmetric division in the stomatal lineage. Then, in Chapter 4, stomatal development genes were examined in germinating cotyledons, and we demonstrated developmental plasticity in division patterns and cell fates. Time lapse imaging has been used on a physiological level since the invention of

cameras, but there is comparatively little previous work in plant protein visualization. We were asked to publish my protocol in the *Journal of Visualized Experiments*, and there has been some demand for information and training.

Seeing stomatal determinants over time allows us to look at the relationship of gene expression to developmental events in a way that is not otherwise possible. Still images can show *SPCH* in cells of a particular shape, and not present in morphologically different cells. Based on other data, such as the *spch* mutant phenotype of no stomatal lineage cells, we can infer that *SPCH* expression is required in cells entering the stomatal lineage. However, time lapse microscopy is the first technique allowing investigation of, for instance, how long *SPCH* expression is present in a cell before asymmetric division occurs, how long the protein persists in SLGCs, and whether all SPCH-positive cells indeed divide asymmetrically. Interesting future possibilities lie in timed and photoconvertible fluorophores that allow visualization of protein lifetimes, showing whether stomatal determinants are held over or remade at each asymmetric division.

In Chapter 4, I evaluated the dynamic expression of tagged stomatal determinants and tested cell-cell communication by laser ablation of putative guard mother cells (GMC). One worrisome discovery was that stomatal development in the cotyledon did not follow the complete picture of stomatal development that I had expected based on previous work: those GMC should not have been present without a preceding asymmetric division. I therefore investigated stomatal gene expression in the previous stage of development, the embryo.

Our unpublished data (Pillitteri, Peterson) had shown that late embryo stages included active *SPCH* and *MUTE* promoters in the epidermis, with *MUTE* in a punctate pattern resembling the nonrandom distribution of stomata after germination. I planned

to document the exact stages in developing embryos over time, extending my time lapse work to earlier stages of the same cotyledon organ I had already studied. I proposed this project to the Department of Biology and was awarded the Washington Research Foundation - Benjamin Hall Fellowship to spend an academic quarter in Nagoya University, studying with the Higashiyama ERATO Live-Holonics Project. The results are presented in Chapter 5.

Embryonic development is thrilling. Organs spring out of masses of undifferentiated cells. A featureless field on the outside of the embryo gives rise to an organized tissue, the epidermis, which is then held in an immature state until germination. Stomatal development in the embryo is clear and elegant: de novo cell patterning in an accessible system. So far, my work confirms that stomatal master regulators and signaling genes are active in the embryo. When combined with multiple reporter assays and peptide treatments, this technique will become a powerful tool for investigating and manipulating stomatal development.

Advances in understanding stomatal development can be adapted and generalized to other organisms. Many eudicot crop plants have similar stomata and general cell divisions compared to the model plant; although others are grasses, similar principles hold, and some genes are used in common (see Chapter 1). Understanding land plants and their relatives at a molecular level will help scientists to support our food supply through climate change and, ideally, to remediate elevated CO_2 levels as much as possible by increasing photosynthetic sequestration.

Asymmetric division is also a fundamental mechanism for creating diverse cell types and enabling growth across kingdoms. The basic plan of related bHLH transcription factors enabling sequential developmental steps in a single process appears in myogenesis and neurogenesis as well as stomatal development, and though

the LRR family in plants is unusually large compared to other kingdoms the general nature of ligand-receptor signaling is critical to multicellular organisms. Understanding stomatal development and patterning in depth will add to an overall fund of knowledge applicable to all development.

Appendix.

Long-term, high-resolution confocal time lapse imaging of *Arabidopsis* cotyledon epidermis during germination

This appendix was previously published as the text component of:
Kylee M. Peterson & Keiko U. Torii (2012) *Journal of Visualized Experiments* 70: e4426.

Introduction

Imaging in-vivo dynamics of cellular behavior throughout a developmental sequence can be a powerful technique for understanding the mechanics of tissue patterning. During animal development, key cell proliferation and patterning events occur very quickly. For instance, in *Caenorhabditis elegans* all cell divisions required for the larval body plan are completed within six hours after fertilization, with seven mitotic cycles (Sulston *et al.*, 1983); the sixteen or more mitoses of *Drosophila* embryogenesis occur in less than 24 hours (Foe *et al.*, 1993). In contrast, cell divisions during plant development are slow, typically on the order of a day (Vincent *et al.*, 1995; Beemster *et al.*, 2002; Grandjean *et al.*, 2004). This imposes a unique challenge and a need for long-term live imaging for documenting dynamic behaviors of cell division and differentiation events during plant organogenesis. *Arabidopsis* epidermis is an excellent model system for investigating signaling, cell fate, and development in plants. In the cotyledon, this tissue consists of air- and water-resistant pavement cells interspersed with evenly distributed stomata, valves that open and close to control gas exchange and water loss. Proper spacing of these stomata is critical to their function, and their development follows a sequence of asymmetric division and cell differentiation steps to produce the organized epidermis (Figure A-1).

This protocol allows observation of cells and proteins in the epidermis over several days of development. This time frame enables precise documentation of stem-cell divisions and differentiation of epidermal cells, including stomata and epidermal pavement cells. Fluorescent proteins can be fused to proteins of interest to assess their dynamics during cell division and differentiation processes. This technique allows us to understand the localization of a novel protein, POLAR (Pillitteri *et al.*, 2011), during the proliferation stage of stomatal-lineage cells in the Arabidopsis cotyledon epidermis, where it is expressed in cells preceding asymmetric division events and moves to a characteristic area of the cell cortex shortly before division occurs. Images can be registered and streamlined video easily produced using freely available shareware to visualize dynamic protein localization and cell types as they change over time.

Protocol text:

1.) Seed sterilization

1.1) Prepare seed sterilization solution: 33% household bleach, 0.1% Triton X-100.

1.2) Place seeds carrying desired fluorescent reporter construct(s) and genotype(s) in 1.7mL tube and apply 1mL of sterilization solution. Incubate on nutator for 15 minutes.

1.3) In a sterile hood, use a pipettor to remove sterilization solution from tube, leaving seeds behind. Rinse with 1mL sterile water. Repeat four times.

1.4) Incubate at 4 °C for two days or more.

2.) Preparation of chamber slide media

2.1) Mix 20mL of 0.5% solution of Bacto Agar (not pure agarose) in water in a 200mL flask, and microwave until dissolved. Be cautious when boiling the solution.

2.2) Cool the solution to around 60 °C.

2.3) Collect 1mL of agar solution with a pipettor and slowly apply to the cover glass inside the chamber slide (Lab-Tek II Chambered #1.5 German Coverglass System). Solution which is too hot or applied too quickly may melt glue or crack glass.
2.4) Immediately add a second mL of agar solution. Agar media should now completely cover the bottom of the slide chamber. This minimal media is sufficient to support development in a plant cotyledon, which contains stored nutrients, for four to five days by maintaining moisture and allowing gas diffusion.
2.5) Cool slide on benchtop with chamber covered. If flask is covered tightly, solution may be saved and reheated for future use.

3.) Seed dissection

3.1) Remove tube of sterile seed from 4 °C storage.
3.2) Place a paper tissue on the stage of a dissecting microscope and moisten with water. Adjust tissue to create a smooth surface. Tissue is used to stabilize seeds for dissection.
3.3) Pipet 20-30 seeds from the tube to the tissue, being careful to place them in the dissecting scope's field of view.
3.4) Using sharp forceps (Roboz, #5 biology tip), carefully remove the seed coat from the seedling inside. Both outer and inner integuments must be removed.
3.5) When imaging the cotyledon, it is beneficial to remove the hypocotyl and radicle. Cotyledons contain sufficient nutrients to develop for several days under this protocol. If intact, the hypocotyl will straighten and lengthen dramatically, moving the cotyledon out of the time lapse field of view. Use a scalpel to slice the cotyledons free of the hypocotyl.
3.6) Hold dissected cotyledons immersed in water, in a microtube or similar.

3.7) Repeat 3.4-3.6 until the desired number of cotyledons is reached. 15-20 successful dissections are recommended before mounting.

4.) Mounting cotyledons in the chamber slide

4.1) Using a small (18mm^2) cover slip, cut through the agar media and lift the entire layer up from the chamber slide's glass base.

4.2) Pipet dissected cotyledons with a minimum of water from holding tube into the chamber slide, under the lifted agar layer.

4.3) Gently lower the agar onto the cotyledons. If excess water is present, soak it away with a tissue, being careful not to disturb the cotyledons. Specimens should no longer move easily when mount is complete.

4.4) Place cover on chamber slide and move slide to microscope stage.

5.) Time lapse imaging

5.1) Set up inverted confocal microscope to image the desired fluorophore(s). LSM700 parameters used: for GFP, excitation at 488 nm and collection with a band-pass filter at 440-530 nm; for RFP, excitation at 555 nm and collection with a band-pass filter at 570-610 nm.

5.2) Inspect mounted specimens for damage. Program the locations of intact, properly mounted cotyledons into microscope software. In Zen 2009: Focus on a cotyledon with 20x objective and move objective to center of chamber slide. Under *Acquisition Mode*, change *Zoom* to 0.5 and *Frame Size* to 48x48 pixels. Select *Positions* checkbox and *Scan overview image...* button under *Positions*, then increase *Horizontal* and *Vertical* tile numbers to 30-35. When scan is complete, click the *Positions* button under the *Dimensions* tab and place crosshairs at each cotyledon to observe.

5.3) Set up z series encompassing the cotyledon epidermis of all samples.

In Zen 2009: Click the *Z-Stack* checkbox. Select each position under *Position List* and click the *Move to* button. Focus on the uppermost plane of the cotyledon epidermis and click the *Set First* button under *Z-Stack*. Focus to the lowest position at which epidermis is usefully visible and click the *Set Last* button. Click the button next to *Optimal*, which shows the best z-slice thickness, to set. Check each cotyledon to confirm that these settings encompass all samples; the z parameters cannot be set for individual positions in Zen 2009.

5.4) Set up time series at desired resolution and length. Intervals of 15-30 minutes for three days are effective for protein dynamics in epidermal cell division. This interval can be changed as necessary. In Zen 2009: Click the *Time Series* checkbox. Under *Time Series*, set *Cycles* for the number of images desired using the slider or typing in the text field. Set *Interval* to 30 and use the pull-down menu to select *min*.

5.5) Begin xyzt multi-position scan. Note that the number of positions must be limited by the scan specifications; if total scan time is greater than the desired interval, the time points will not be correct. A maximum of six simultaneous positions are possible when imaging GFP and RFP at 59 z-slices in a 30-minute interval using our system. In Zen 2009: Change *Frame Size* to desired resolution and click *Start Experiment*.

6.) Video editing

6.1) From the confocal data file, make a maximum intensity z-projection of each time/position combination and export as image files in a lossless format such as TIFF. File names should be in one series per position (e.g., GFP-POLAR_pos1_0001, GFP-POLAR_pos1_0002, etc., not GFP-POLAR_0001_pos1) for software to combine them into a movie. In Zen 2009: Use *Copy:Subset* under the *Processing* tab to split positions into separate files, then *Maximum Intensity Projection*. Finally, export as TIFF (*File:Export, Full resolution image window - series*).

6.2) Use FIJI (Schindelin, 2008; Abramoff *et al.*, 2004) to align successive images by running the bUnwarpJ macro (Arganda-Carreras *et al.*, 2006) (*Plugins: Registration:bUnwarpJ*). Change *Registration Mode* to *Mono*. All *Advanced Options* can use default values. For long time lapse sequences, download and install the macro "Affine + Consistent Elastic 2D Image Registration," which will apply bUnwarpJ to a series of images automatically, and open your data using *File:Import:Image Sequence* to use it. Uncheck *MOPS landmarks extraction* and run.

6.3) Use FIJI/ImageJ or Quicktime Pro to open the sequence of aligned images and save in the desired video format (AVI is a good choice).

Representative results

A set of informative time points collected with this method is shown in Figure A-3. Cell membranes are labeled with RFP (pm-rb) and GFP is fused to POLAR protein under its native promoter (*POLARpro::POLAR-GFP*). At the 30-minute time scale, we see cell divisions along with the changes in protein localization preceding them. Asymmetric cell divisions in the stomatal lineage form stem-cell-like stomatal precursors called meristemoids, which retain the ability to undergo further asymmetric divisions, and their sister cells, called stomatal lineage ground cells (SLGCs). SLGCs do not assume a stomatal precursor fate; they frequently transdifferentiate into pavement cells, but may resume asymmetric division capability to produce another meristemoid spaced away from the first as the leaf expands. All of these fates are reflected in the expression and localization of POLAR-GFP (Figure A-3, Video A).

Video Legend

Video A: Streamlined, registered video of *POLARpro::POLAR-GFP* localization during amplification and spacing of one stomatal precursor cell.

Initially, POLAR-GFP appears uniformly in the cell; by 39 hours after germination, it polarizes strongly, just before a division (40.0h) placing a smaller meristemoid at the opposite end of the cell. The larger cell at right also regains stomatal lineage identity, and by 45 hours POLAR-GFP localization moves adjacent to the stomatal precursor. The division at 47 hours produces a new meristemoid (right) oriented away from the existing meristemoid (left) from the first. The end result is two stomata separated by one cell. (Images missing from the video were not collected and do not represent significant changes.)

Discussion

This time-lapse confocal technique allows longitudinal studies of fluorescently tagged protein expression and localization in individual cells of the *Arabidopsis* cotyledon epidermis, which in the case of POLAR and other dynamically changing proteins is critical to a proper understanding of their function. Previously, sustained time-lapse imaging has been used to examine *Arabidopsis* root fungal infection (Czymmek *et al.*, 2007) and meristem growth (Campilho *et al.*, 2006), but adding the cotyledon epidermis expands the versatility of this technique and allows its use for additional proteins, particularly assisting the expanding field of stomatal development.

The protocol's main limitation is that it is currently restricted to cotyledons, which contain the nutrients they need for early development. Further, cotyledon development is not exactly identical to that undergone in air because the gel medium restricts gas exchange. Scanning laser exposure and room lights are adequate for photomorphogenesis, inducing cotyledon expansion and chloroplast maturation, but stomata may occasionally develop in adjacent pairs due to low CO_2 concentration. This tendency can be reduced by using only a thin layer of media and minimal mounting water as directed in this protocol.

Fluorophore selection is also important for success with this technique. Both GFP and RFP work well, and allow double labeling to visualize the cell periphery or assess protein co-localization. Custom filters greatly reduce autofluorescence and allow sensitive detection of fluorescent protein tags even when chlorophyll is present. However, even filtered CFP autofluorescence in germinating cotyledons is strong enough to preclude visibility of almost all signal using the LSM700. For instruments with a spectral unmixing capability, a wider selection of fluorophores may be feasible.

Acknowledgments

We thank Amanda Rychel for assistance in developing the time lapse protocol and Lynn Pillitteri for constructing *POLARpro::POLAR-eGFP*. We are also grateful to ABRC for providing the *pm-rb* construct. This protocol was developed through a support from the PRESTO award from Japan Science Technology and Agency. Research on POLAR was also supported by the University of Washington Royalty Research Fund (RRF-4098) and the National Science Foundation (MCB-0855659). KMP was an NSF Graduate Research Fellow (DGE-0718124), and KUT is an HHMI-GBMF investigator.

Disclosures: We have nothing to disclose.

Table of specific reagents and equipment

Name of reagent	Company	Catalog Number	Comments
Bacto Agar	BD	214010	
One-chamber slide	Nunc (Thermo Scientific)	155360	Or two-chamber (155379)
Laser scanning confocal microscope	Zeiss	LSM700	Zen 2009 software
20x objective lens	Zeiss	420650-9901	NA 0.8, Plan-APOCHROMAT
Dissecting microscope	Benz (National)	431TBL	Illuminates from below
#5 forceps, biology tip	Roboz Surgical Instrument	RS-4978	Very fine tips are critical

References

Abramoff MD, Magalhaes PJ, Ram SJ. 2004. Image processing with ImageJ. *Biophotonics International* 11(7):36.

Arganda-Carreras I, Sánchez Sorzano CO, Marabini R, Carazo JM, Ortiz-de Solorzano C, Kybic J. 2006. Consistent and elastic registration of histological sections using vector-spline regularization. *Computer Vision Approaches to Medical Image Analysis*, ser. *Lecture Notes in Computer Science*, vol. 4241. Heidelberg: Springer Berlin. 85-95.

Beemster GTS, De Vusser K, De Tavernier E, De Bock K, Inzé D. 2002. Variation in growth rate between *Arabidopsis* ecotypes is correlated with cell division and A-type cyclin-dependent kinase activity. *Plant Physiology* 129(2):854.
Campilho A, Garcia B, Toorn HV, Wijk HV, Campilho A, Scheres B. 2006. Time-lapse analysis of stem-cell divisions in the *Arabidopsis thaliana* root meristem. *Plant Journal* 48 (4):619.
Czymmek KJ, Fogg M, Powell DH, Sweigard J, Park S-Y, Kang S. 2007. In vivo time-lapse documentation using confocal and multi-photon microscopy reveals the mechanisms of invasion into the *Arabidopsis* root vascular system by *Fusarium oxysporum*. *Fungal Genetics and Biology* 44(10):1011.
Foe V, Odell G, Edgar BA. 1993. Mitosis and morphogenesis: point and counterpoint. *Development of Drosophila melanogaster*. Ed. M. Bate, A. Martinez Arias. Plainview, NY: Cold Spring Harbor Laboratory Press. Chapter 3.
Grandjean O, Vernoux T, Laufs P, Belcram K, Mizukami Y, Traas J. 2004. In vivo analysis of cell division, cell growth, and differentiation at the shoot apical meristem in *Arabidopsis*. *Plant Cell* 16(1):74.
Pillitteri LJ, Peterson KM, Horst RJ, Torii KU. 2011. Molecular profiling of stomatal meristemoids reveals new component of asymmetric cell division and dommonalities among stem cell populations in *Arabidopsis*. *Plant Cell* 23(9):3260.
Schindelin J. November 2008. Fiji Is Just ImageJ (batteries included). http://fiji.sc/wiki/index.php/Fiji
Sulston JE, Schierenberg E, White JG, Thomson, JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology* 100:64.
Vincent CA, Carpenter R, Coen ES. 1995. Cell lineage patterns and homeotic gene activity during *Antirrhinum* flower development. *Current Biology* 5:1449.

Figure A-1.

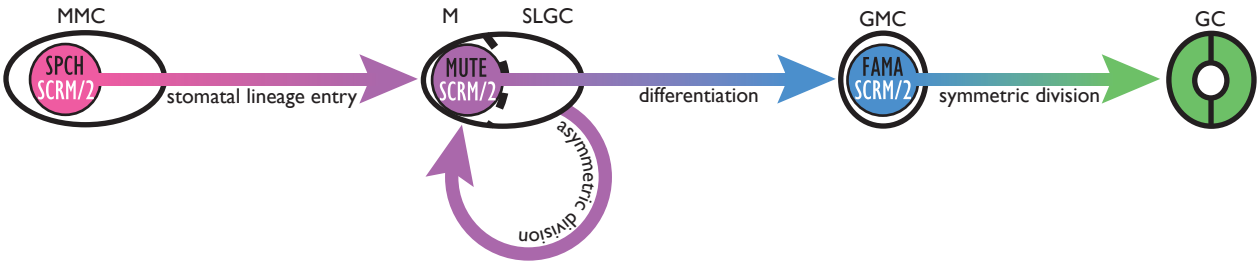


Figure A-1. Schematic of stomatal development.
Protodermal cells enter the stomatal lineage as meristemoid mother cells (MMC) in a step controlled by the basic helix-loop-helix transcription factors *SPEECHLESS (SPCH), SCREAM (SCRM)*, and *SCRM2*. MMCs divide asymmetrically to produce a meristemoid (M) and stomatal lineage ground cell (SLGC); this step may occur multiple times and provides one mechanism by which stomata are spaced apart. The cell-state transition of the meristemoid to the guard mother cell (GMC) identity is directed specifically by the bHLH *MUTE* as well as *SCRM*/2. A final symmetric division of the GMC and its differentiation into two mature guard cells (GC) requires the bHLH *FAMA* with *SCRM*/2 (reviewed in Peterson et al. 2010).

Figure A-2.

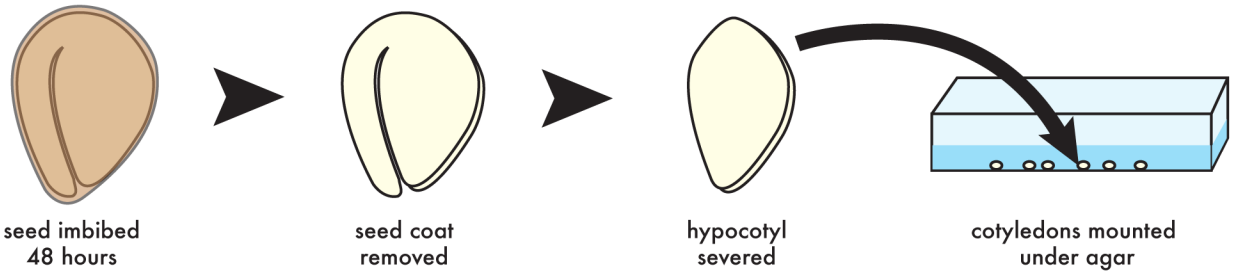


Figure A-2. Diagram of seed dissection and mounting procedure.
Seeds are sterilized and held at 4° C, the seed coats and hypocotyls removed, and the cotyledons mounted under agar media in a chamber slide for imaging on an inverted confocal microscope.

Figure A-3.

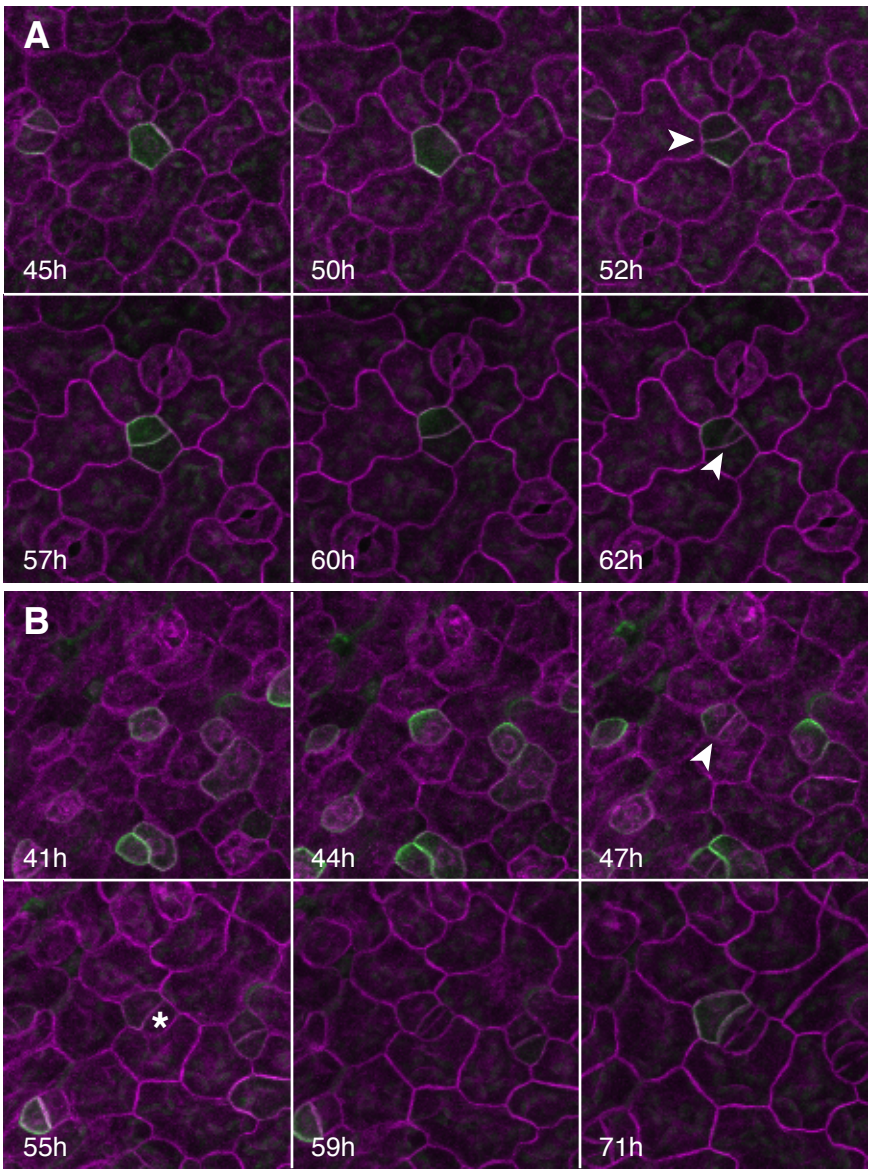


Figure A-3. Time-lapse imaging demonstrates *POLARpro::POLAR-GFP* distal localization preceding asymmetric cell division.
Series A: POLAR-GFP initially appears evenly distributed, then approximately two hours before asymmetric divisions (arrowheads) POLAR-GFP segregates away from the site of the incipient meristemoid. Series B: Following the initial asymmetric division (arrowhead), POLAR-GFP disappears in both cells, implying rapid transition to guard mother cell (GMC) state by the meristemoid. The GMC (asterisk) divides symmetrically and differentiates to form stomatal guard cells, while the sister cell regains POLAR-GFP expression, presumably presaging a later asymmetric division.