2008

Control Of Stomatal Formation In Arabidopsis Thaliana Inflorescence Stem

Nalini Odapalli
University of Central Florida

CONTROL OF STOMATAL FORMATION IN *ARABIDOPSIS THALIANA* INFLORESCENCE STEM

by

NALINI DEVI ODAPALLI
B.Sc and PGDMISCA, Osmania University, 2000

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science from the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2008

ABSTRACT

Stomata are cellular structures that control water loss and gas exchange through the plant epidermis. Stomata arise from special stem cells called meristemoids through a series of programmed asymmetric divisions that are controlled by cell signaling, or via multitude of regulatory pathways and intercellular communication between epidermal cells. In *Arabidopsis thaliana,* stomata are spaced non-randomly in the epidermis by cell-cell signaling of the receptor-like protein TOO MANY MOUTHS (TMM) as well as other proteins. Point mutation of the *TMM* gene prevents the development of stomata in some tissues like inflorescence stems. Investigation of *tmm* mutant stems showed that self-renewing stem cell-like precursors form by dividing asymmetrically but fail to form stomata. This is further supported by molecular markers of stomatal cell fate that show stomatal precursors form but do not differentiate as stomata. Therefore, TMM signaling is likely required to control expression of genes that are essential for the formation of stomata in stems. As a second approach, gene expression profiling was used to identify candidate genes involved in stomatal biogenesis. Differentially expressed genes were categorized by gene ontology and analyzed for statistically overrepresented classes to gain insight into functional processes. Comparison of stem expression data with previously published microarray data was used to narrow the list to genes involved in stomatal patterning. Mutants in these target genes have been obtained and phenotypic analysis revealed new stomatal regulators. Comparison of epidermal cells of the stem tip region from wild-type and *tmm* revealed that there are significantly more meristemoids formed in *tmm* stems compared to wild-type stems. In addition, the

orientation of meristemoids formed in wild-type stems was random with respect to stem polarity and followed a spiral pattern of asymmetric divisions similar to leaves. This showed that stomatal patterning in dicots does not follow orientation in asymmetric cell division for spacing the adjacent stomata like monocots.

ACKNOWLEDGEMENTS

I would like to thank Dr. Jeanette Nadeau for all the aid and advice for helping me achieve my goals. I dedicate my thesis to my beloved daughter Siri.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GENERAL INTRODUCTION

Background and significance

Stomata are microscopic pores on plant organs that regulate the exchange of CO_2 and H_2O with the surrounding environment. *Arabidopsis*, the model plant in this study is an annual flowering plant that belongs to the Brassicaceae (mustard or crucifer) family. Stomatal patterning in *Arabidopsis* is emerging as a significant model system for the study of molecular level regulation and genetics of cell differentiation, pattern formation in plants. Revealing the molecular nature of stomatal regulators will help us understand cell differentiation and utilize that knowledge to produce plants with better water use efficiency, ultimately improving the productivity of plants. Plants with more abundant stomata could reduce atmospheric CO_2 may serve as a tool in reducing global warming and thus the study of genes that control stomatal development would be very useful.

In wild-type Arabidopsis plants stomata are arranged in a nonrandom pattern. Most stomata are separated by at least one non-stomatal cell on leaves (Yang and Sack, 1995; Berger and Altmann, 2000). The TOO MANY MOUTHS mutant disrupts the 1-cell spacing rule in leaves, leading to the formation of large clusters of stomata. Also, in *tmm* mutant plants there are too many stomata on cotyledons, leaves and some other organs of the plant whereas no stomata on inflorescence stems. The *tmm* mutation is not deleterious as plants grown in laboratory conditions look normal and healthy except for reduced size of siliques and flower stalks (Geisler *et al*., 1998).

Stomatal development

Stomatal development has been described in *Arabidopsis* leaves but not in stems. In a newly formed leaf there are undifferentiated protodermal cells among which some assume the stomatal lineage and form meristemoid mother cells (MMC). It is not known how MMCs are chosen from the protodermal cells of the epidermal layer. First, a meristemoid mother cell undergoes asymmetric division to produce a larger sister cell and a smaller cell, the meristemoid (Fig. 1A). The meristemoid may further divide asymmetrically several times to form neighbor cells (NCs) of the same cell lineage. These series of divisions can be termed amplifying divisions as they contribute to the total number of epidermal cells. These new NCs not only surround future stomata, but also space adjacent stomata to maintain the 1-cell spacing rule (Fig. 2A). Eventually the meristemoid stops asymmetric division and differentiates into a guard mother cell (GMC). Finally the GMC divides symmetrically to form two cells that differentiate as guard cells (GC). Young GCs develop a pore between them, thus forming the stomata (Nadeau and Sack, 2002).

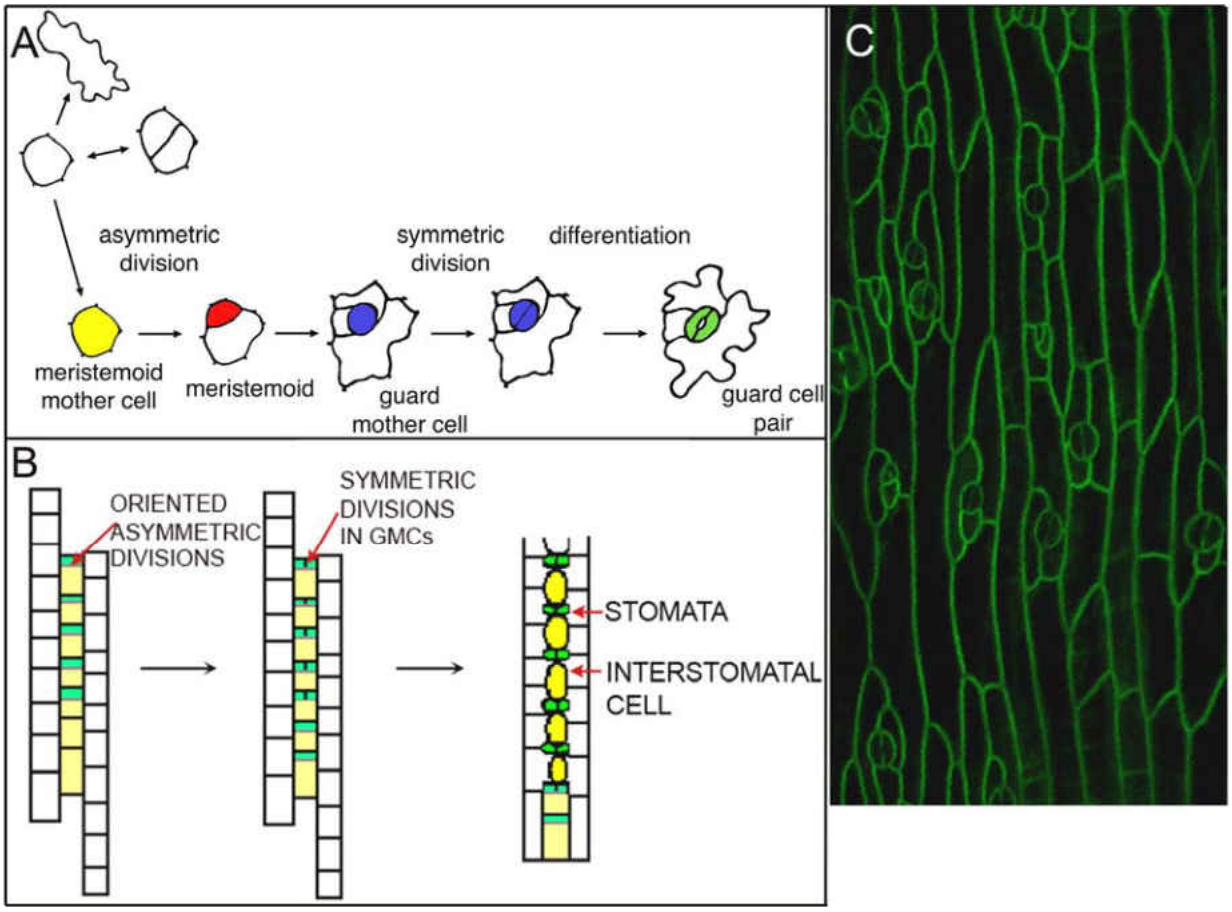


Figure 1. Stomatal development in Arabidopsis and maize leaves

(A) Stomatal development in Arabidopsis leaves occurs through a series of asymmetric divisions that may require cell-cell communication to control patterning.
(B) In monocot leaves only one asymmetric division occurs relative to the polarity of the organ, and cell lineage spacing may play a role in stomatal patterning. This diagram is modified from Hernandez *et al*. (1999).
(C) The mature stem epidermis in Arabidopsis shows linear files of cells similar to the monocot leaf.

CHAPTER 1:
STOMATAL DEVELOPMENT IN *ARABIDOPSIS* INFLORESCENCE STEMS

Introduction

Stomatal patterning

The meristemoid is a type of plant stem cell that has limited capacity for self-renewal and ultimately differentiates into stomata. In the C24 ecotype of *Arabidopsis* leaves, the meristemoid divides three times asymmetrically in sequence, every time producing a larger sister or subsidiary cell and a smaller meristemoid (Berger and Altmann, 2000; Serna and Fenoll, 2000). Some of the neighbor cells retain the potential to re-enter the cell cycle and form satellite meristemoids (Larkin *et al*., 1997). Thus during the growth of the epidermis meristemoids may arise adjacent to GCs or other meristemoids. A meristemoid in contact with one GC or precursor can retain stomatal cell fate but the next asymmetric division would be oriented away from the pre-existing stomatal cell resulting in new meristemoid separated from the stomata by a pavement cell (Geisler *et al*., 2000). For this to occur, cellular interactions among cells in various stages of stomatal lineage and also between stomatal cells and pavement cells are required for generating spaced stomata. These cellular interactions occur through cell signaling. In plants containing mutations in stomatal patterning genes such as *tmm,* the signaling pathway responsible for orientation of the division plane is defective so it results in clusters of adjacent stomata on leaves (Fig. 2B) and loss of stomata on inflorescence stems (Fig. 6D).

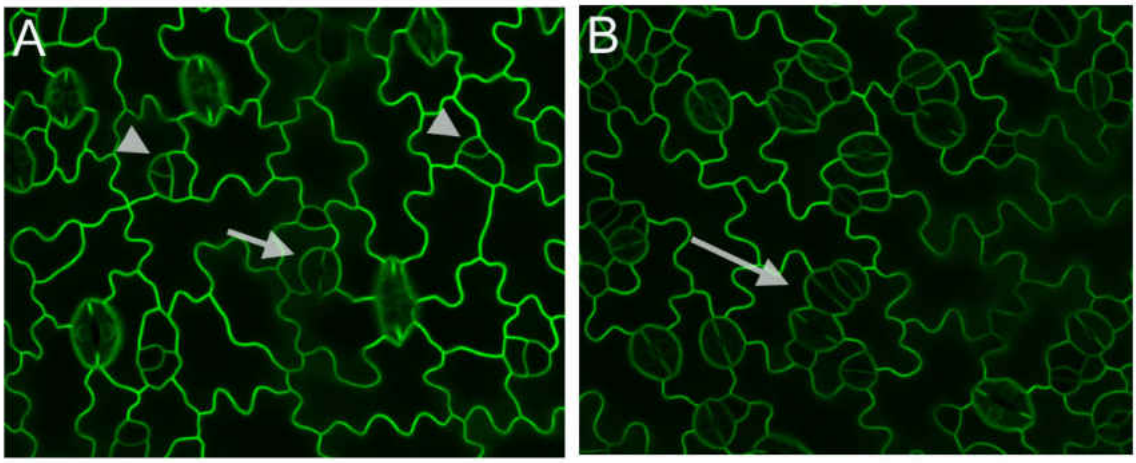


Figure 2. Plasma membrane marker fused to GFP showing the leaf epidermis
(A) Spaced stomata in a wild-type leaf.
(B) Clusters of stomata in *tmm* leaf.
Meristemoids indicated by arrowheads and stomata by arrows.

TMM regulates asymmetric divisions and orientation of division plane in meristemoids

In the leaf, *TMM* marker expression is seen in some epidermal cells but absent from the shoot apical meristem (SAM) suggesting TMM has no role in development of protodermal cells (Nadeau and Sack, 2002). Weak TMM expression in older neighbor cells and strong expression in young NCs that are going to divide or have undergone a division suggest TMM may be expressed in cells that are division competent (Fig. 3). The presence of TMM in cells between two stomata or precursors that would not undergo any cell division suggests that TMM may help in prohibiting cell division in certain spatial locations. Therefore, TMM is suggested to detect extracellular cues to regulate asymmetric divisions and orientation of the division plane that forms

meristemoids to avoid formation of adjacent stomata. The expression of TMM in stem tips is presented in my results.

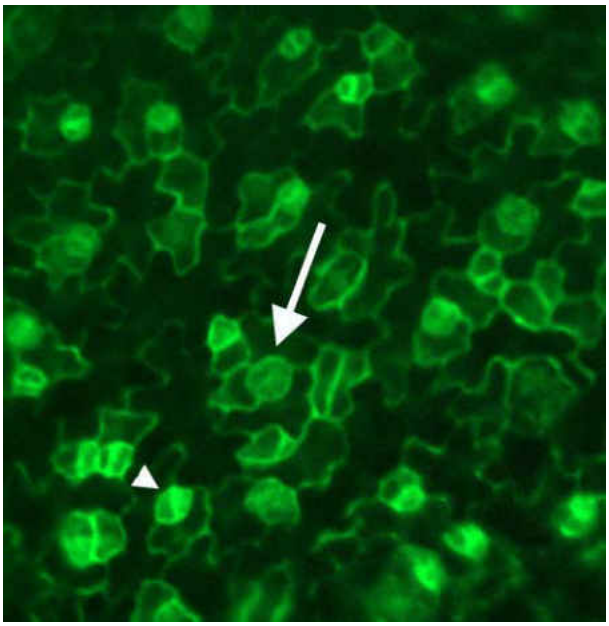


Figure 3. *TMMpromoter::GFP* expression pattern in the wild-type leaf
Bright fluorescence in stomatal lineage cells. Strong expression in meristemoids (arrowhead) and faint expression in stomata (arrow) and neighbor cells.

TMM negatively regulates stomata in leaves and positively regulates stomata in stems

Stomatal development and patterning has been studied intensively in *Arabidopsis* leaves to lay a background for understanding the function of genes involved in the formation of stomata in dicots (Yang and Sack, 1995; Larkin *et al*., 1997; Nadeau and Sack, 2002). Mutation of *TOO MANY MOUTHS* had opposite effects in the *Arabidopsis* stem and leaf so that stomatal clusters form in the leaf but no stomata form in the inflorescence stem (Geisler *et al*., 1998). Stomata are also lost from other regions such as hypocotyls and the adaxial side of the sepal. The number of stomata is

reduced in siliques, cauline leaves and base of flower stalks. On the other hand the number of stomata increased in cotyledons, anthers and abaxial side of the sepal, and these organs exhibit varying degrees of clustering. In conclusion, some organs such as leaf do not require TMM activity for stomatal formation but it is absolutely essential in stems. Therefore TMM is formally a negative or positive regulator of stomatal formation in different regions of the plant, most likely depending on different molecular interactors. There is a need to understand the similarities and differences in epidermal layer development and patterning of *Arabidopsis* stems and leaves to understand better the roles of *TMM*.

Comparison of epidermal patterning in the monocot leaf and dicot stem

Comparing the stomatal patterning mechanisms found in monocotyledonous leaves to the *Arabidopsis* stem is a potential way to understand common features. It also allows us to understand questions related to evolution of land plants because stomata are the primary structures that allowed plants to colonize land. Both monocot and dicot classes belong to the angiosperm phylum or flowering plants. Monocots (grasses, lilies, orchids, etc.) are a monophyletic group, whereas dicotyledonous plants are a paraphyletic group. Grasses represent a clade of the monocot lineage, and have strap shaped leaves with rectangular epidermal layer cells arranged in longitudinal files. In contrast, many dicots have round or oval shaped leaves. The epidermal layer of many dicots has puzzle-piece shaped pavement cells. Here I used Arabidopsis as a representative of the dicot class and grasses from the monocot class.

The epidermal layer of the dicot leaf and stem has different patterns of epidermal development that may be due to differences in organ morphology. The dicot leaf is flat and oval with randomly distributed stomata and puzzle piece or irregularly-shaped pavement cells that make up the epidermal layer. In contrast the stem is cylindrical and elongated, its epidermal layer is made up of stomata and rectangular or trapezoidal shaped long pavement cells in files (Fig. 1C).

There is a resemblance in the organization of the cells of the epidermis in the dicot stem and the monocot leaf. During development, there is a continuum of stomatal lineage cells in the *Arabidopsis* stem epidermis and the monocot leaf epidermis, with young cells at the growing end and mature stomata at the other end. The *Arabidopsis* stem tip has meristematic cells at the tip that add new cells to the growing stem and at the base of the stem there are mature stomata and pavement cells. This situation is opposite from the monocot leaf, which has mature cells and stomata towards tip of the leaf blade and the intercalary meristem where stomatal initials are formed at the basal end. Here, stomata are formed in linear files of cells that originate from a narrow intercalary meristem at the base of the leaf (Fig. 1B). Finally, guard mother cells are always oriented towards the leaf apex so stomata are not formed in contact with each other (Hernandez *et al*, 1999). This contrasts with the *Arabidopsis* leaf epidermis, where asymmetric cell divisions forming the meristemoid do not occur relative to the leaf axis polarity (apical-basal). The orientation of asymmetric division in the *Arabidopsis* leaf is controlled by cell-cell communication to achieve non-random stomatal patterning during mosaic growth (Fig. 1A). To understand the stomatal

patterning mechanism in the *Arabidopsis* stem I hypothesized that *Arabidopsis* stem epidermal cells, similar to monocot leaves, undergo asymmetric divisions relative to the stem apical-basal axis.

While it is unknown how stomatal patterning is controlled in monocot leaves, it was proposed that it is dependent on the position of cells within the intercalary meristem when they reach a specific phase of mitosis (Charlton, 1990). This contrasts with *Arabidopsis* leaf stomatal patterning, which requires cell-cell communication as new stomata fill-in between old (Nadeau and Sack, 2002). Stomatal patterning mechanisms have not been determined in stems and cylindrical organs of dicots, where linear files of cells more similar to grass leaves are found. Similarities and differences in the monocot leaf, dicot stem and dicot leaf may be informative in understanding mechanisms involved in stomatal patterning. The big question is how the undifferentiated meristematic cells at the tip of *Arabidopsis* stem acquire and maintain the stomatal fate. It is not possible to address this question at a mechanistic level, but an attempt has been made to understand if stomatal precursors follow a preplanned or oriented (relative to stem polarity) divisions to create the minimum one cell spacing between adjacent stomata.

Methods

GUS staining

GUS staining solution was prepared using 8.8 mL ddH_2O, 1mL of 0.5 moles phosphate buffer at pH 7.0, 0.1mL of 0.1M potassium ferricyanide, 0.1 mL of 10% triton and 10 mg of X-Gluc powder (5-Bromo-4-chloro-3indolyl β-D-glucuronide cyclohexylamine salt) by Rose Scientific Limited. Fresh tissue from the inflorescence stem and rosette leaf or cotyledon were soaked in GUS staining solution and vacuum suction was applied for 20-60 minutes to infiltrate staining solution into the tissue until tissue looked wet. The tissue was then incubated overnight in 37ºC air incubator. The amount of time vacuum applied varied depending on the strength of the expression of the reporter. After staining, stems were destained in ethanol series of increasing concentrations 30%, 50%, 70%. Before mounting on slide tissue was rehydrated in 50% and 30% ethanol and then to water.

Microscopy

After GUS staining and destaining the wet mount of tissue was made for microscopic observations. An Olympus BX60 compound microscope was used for observations. Depending on the requirement 20X U Plan FI, 40X U Plan FI or 100X U Plan FI objectives were used. Pictures taken with transmitted light were captured with a Carl Zeiss camera, Serial number 242042995 and processed by using AxioVs40 AC V 4.3.0.101 software.

Plants that expressed Green Fluorescent Protein (GFP) as a reporter were observed using an Olympus FV300 confocal scanning laser microscope. The 488nm wavelength of an Argon laser was used to excite GFP, and emission wavelengths captured through 505-525 nm filter. All pictures were taken with 40X U PlanAPO objective. Images were processed by the Olympus Fluoview version 4.3 software.

Stomatal lineage molecular markers

GRL2 is a molecular marker for meristemoids (Kim *et al*., 2003). Wild-type *Arabidopsis* plants containing the transgene (GUS reporter fused to the *GRL2* promoter) were crossed with *tmm-1* plants to obtain mutant plants containing this molecular marker. This molecular marker was used to test whether meristemoids were formed in *tmm* mutant stems. Tissues were stained using GUS staining solution prepared as described above except that 20 mg/ml of X-Gluc powder was used. The tissue was vacuum infiltrated in GUS staining solution for 1 hour because this molecular marker was faintly expressed. Other steps of GUS staining and mounting were as described above.

ET1967 is an enhancer trap line (Sundaresan *et al*., 1995) that serves as a molecular marker for stomatal lineage cells such as meristemoids, guard mother cells and guard cells. Neighbor cells produced by asymmetric division within the stomatal lineage also stain weakly. Wild-type *Arabidopsis* plants containing the transgene were crossed with *tmm-1* mutant plants to obtain homozygous mutants containing this molecular marker. The tissue was stained using GUS staining solution prepared as

described above and vacuum was applied for 5-10 minutes only, because this tissue stained easily.

KAT1 encodes a voltage-gated inward-rectifying potassium channel (Anderson *et al*., 1992; Nakamura *et al*., 1995). The transgenic *Arabidopsis* plants containing a *KAT1* promoter fragment fused to GUS that is mainly expressed in guard cells were used. Hence *KAT1* was selected as a molecular marker of guard cells to show at what stage stomatal development arrests in the *tmm* stems. These transgenic plants (GUS reporter fused to *KAT1* promoter) were crossed with *tmm-1* mutant plants to obtain homozygous mutants containing this molecular marker. The tissue was stained using GUS staining solution prepared as described above and vacuum was applied for 5-10 minutes only because this molecular marker was expressed strongly. Other steps of GUS staining and mounting were as described above.

Counting the number and orientation of meristemoids in wild-type and tmm stems

Arabidopsis Columbia plants containing the transgene PIP2 (water channel in plasma membrane) fused to GFP reporter line Q8 was used to visualize cell outlines to locate asymmetric divisions in the stem. Cutler *et al* obtained subcellular markers by random fusion GFP::cDNA library of *Arabidopsis* (Cutler *et al*., 2000). Q8 plants were grown for approximately 30-34 days until a stem of ~4 cm was available for counting the number and orientation of meristemoids. To observe the stem epidermis under the microscope, first all the branches, cauline leaves, flowers from base of pedicel and buds were removed using #5 tweezers. Stem tips from six Columbia (wild-type) and *tmm-1*

plants were imaged using confocal microscopy. The 40X U PlanAPO objective lens was used to visualize meristemoids. For the purposes of this experiment, I defined three regions in a growing stem. The regions defined as the “tip” has small, rectangular cells arranged in files. In wild-type and *tmm* stems the ~2mm tip region was selected for counting meristemoids and their orientation (Fig. 4). In wild-type stem tip region some meristemoids have been formed through asymmetric division, but no guard mother cells are found. The region defined as “middle” has meristemoids that are progressing to form GMCs, and has some interspersed young stomata. The region defined as “mature” is closest to the base, with mature stomata spaced by fully expanded pavement cells. In the tip region all the epidermal cells were of approximately same size as meristemoids. In the middle region the newly formed meristemoids were smaller but the cells other than meristemoids (pavement cells) were intermediate in size. The pavement cells were elongated in the basal region of the stem. Therefore epidermal cell size was used as a marker of stem developmental stage. In the tip region all the cells were of similar size so morphological characteristics of meristemoids were used to differentiate them from other cells. Stomatal precursors undergo series of asymmetric divisions to add cells to the epidermis, so a meristemoid is sometimes surrounded by related cells that are larger in size. The morphological characteristics of meristemoids include triangular or rectangular shape and contain denser cytoplasm. If a meristemoid was formed towards the tip of the stem it was considered to be apical (Fig. 6I), while if it was oriented towards the base of the stem it was considered as basal (Fig. 6I). If the meristemoid was towards the side then it was

counted as a lateral meristemoid. Puffed and oval appearance of GMCs (Fig. 6I) helped to differentiate them from triangular meristemoids. The figure 6I is taken from middle region of stem and was used to show GMCs, meristemoids and stomata in one picture. But the middle region was not selected for counting meristemoids.

The lateral meristemoids were not counted when stomatal development in the stem was compared with monocot leaf. But lateral meristemoids were added to the total meristemoid count when number of meristemoids formed in Arabidopsis leaf and stem were compared.

Because I observed that the shape of the cells in the tip region of wild-type stem were different from *tmm,* I also used pictures of the plants containing the *TMM* promoter driving expression of a GFP reporter that is expressed in stomatal precursors. I examined cells in which GFP was more abundant, which are meristemoids, and used this as a guide to define the appearance of meristemoids in both wild-type and *tmm-1* plants.

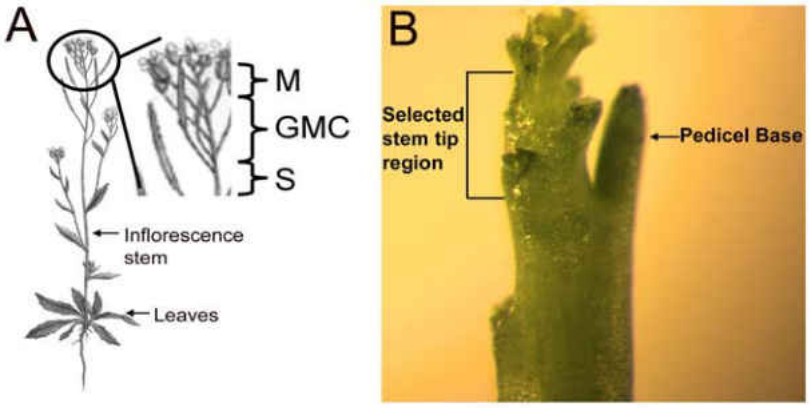


Figure 4. Region of the stem selected for microarray experiments and meristemoid quantitative analysis

(A) Arabidopsis plant illustrating the inflorescence stem with the tip region expanded to show the region containing a gradient of stomatal lineage cells. Meristemoids (M) are found at the tip, followed by occurrence of guard mother cells (GMC) and guard cells (GC) further down.
(B) The tip of the inflorescence stem magnified (10 X) to show the tip region containing meristemoids, with buds and pedicels removed.

Counting the orientation of symmetric division in GMCs of wild-type stems

Arabidopsis Columbia plants containing the transgene *PIP2* (water channel in plasma membrane) fused to a GFP reporter (line Q8) was used to locate symmetric divisions in the GMCs. The morphological characteristics used to locate GMCs are; round in shape and showing a symmetric cell division wall. The middle region has meristemoids that are progressing to form GMCs, so the middle region pictures were chosen for counting. The symmetric cell division in GMCs that is no more than 2° off from parallel to the stem longitudinal axis was defined as a longitudinal symmetric division. The symmetric cell division in GMCs that is out of this range was defined as a non-longitudinal cell division. The number of longitudinal and non-longitudinal symmetric cell divisions in GMCs were counted manually and recorded for analysis.

Image processing

Confocal microscopy was used for taking pictures of Q8 wild-type and *tmm* stem tips to count number and orientation of meristemoids. A 40X objective was used to take pictures. Images were captured and saved as tiff files by the Olympus Fluoview software. These were converted to grayscale images using Adobe Photoshop. Images were printed out and glued together in a complete series from the tip to the middle region of the stem. A paper mask of 5x8 cm was used for selecting the area on the pictures in which meristemoids were counted. The mask area of 5x8 cm is equal to $2.6x10^4$ μm^2 on the stem. Counting was performed manually on the printed images and the numbers were recorded. The statistical software package SPSS 11.5 was used to perform two-tailed t-test to find if there is a significant difference between the number of apical and basal meristemoids in wild-type stems, and if there is a significant difference in the total number of meristemoids between wild-type and *tmm-1* stems. Sigmaplot 2004 (version 9.01) was used for producing graphs.

Results

Asymmetric divisions in Arabidopsis wild-type stems do not occur relative to the apical-basal axis of stem

In wild-type stems, roughly square or trapezoidal shaped meristemoid mother cells (MMC) undergo an initial asymmetric division to form smaller cell (meristemoid) that assumes the stomatal fate. In order to understand the stomatal patterning in wild-type *Arabidopsis* stems I determined if meristemoids are formed with polarity determined by the organ apical-basal axis. First, asymmetric divisions were classified into three categories based on their orientation relative to the stem axis: apical, basal and lateral. Counting of meristemoids from Q8 stem tip pictures showed that there were ~33 apical, ~27 basal and ~8 lateral meristemoids in $2.6x10^4$ μM^2 area on a wild-type stem epidermis. Although there are more apical than basal meristemoids, statistical analysis on these numbers showed that there is no significant difference between the number of apical and basal meristemoids (Fig 5B, 6A). This showed that asymmetric divisions are not formed relative to the stem axis. It also showed that *Arabidopsis* stems do not follow a stomatal patterning mechanism similar to the monocot leaf blade, where in the asymmetric divisions are oriented to the organ axis (Table 1).

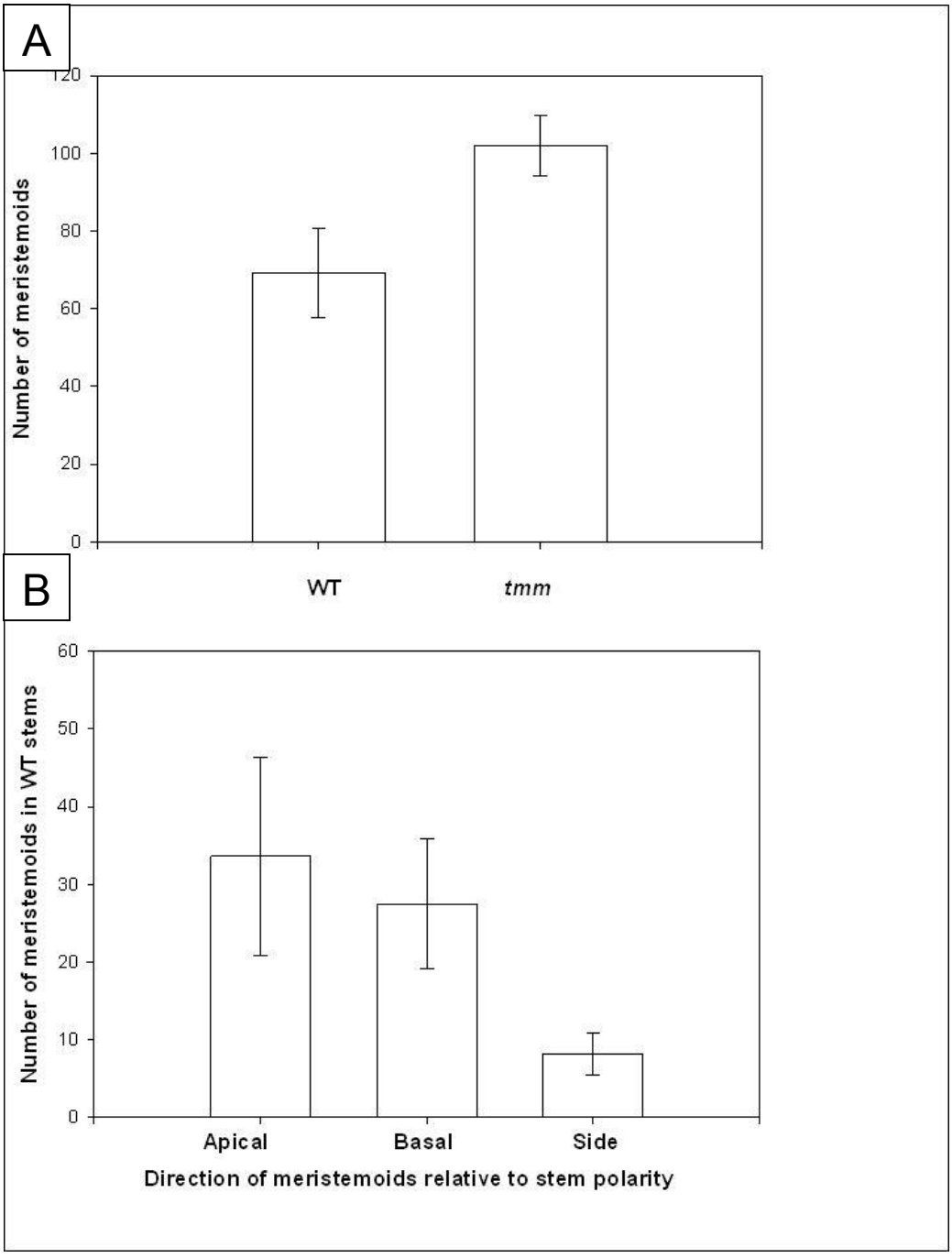


Figure 5. The number and orientation of meristemoids in wild-type and *tmm* stems

Statistical analysis software SPSS was used.

A. Significantly more meristemoids are formed in *tmm* stems than wild-type stems.
B. No significant difference in orientation of meristemoids relative to the organ axis was found in wild-type stems.

Amplification divisions occur in the Arabidopsis stems

To understand if stomatal development and pattering in wild-type stems is similar to the leaf, I looked for similar developmental features or milestones between them. Meristematic cells add more cells to the epidermis during growth, and among them some assume MMC fate. The newly formed MMCs undergo a series of 3 to 4 asymmetric divisions to form pavement cells. These series of asymmetric divisions are termed amplification divisions when described in *Arabidopsis* leaves. Our static observation of meristemoids in the tip region of *Arabidopsis* stem (Fig. 6B) showed that they also undergo amplification divisions as that in leaf. Also, each asymmetric division is oriented at an oblique angle so that the meristemoid is almost always produced at the inside of a spiral series of divisions. In conclusion one similarity between wild-type leaf and stems is that spiral amplification divisions occur in their meristemoids.

Another similarity is that stems also have satellite meristemoids (Fig. 12B-C) but are fewer than leaves. This could be the reason for fewer clusters of stomata in mature stems than leaf and indeed less need for cell-cell communication for spacing the stomata.

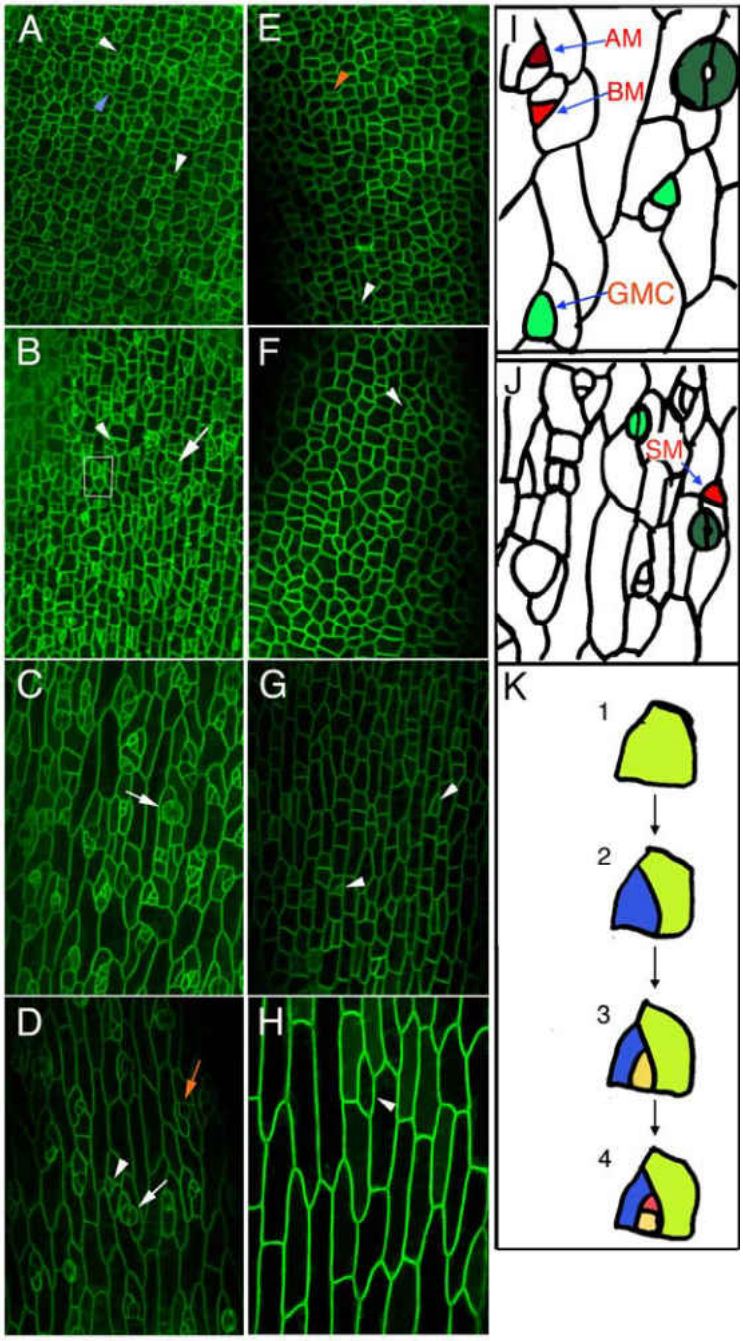


Figure 6. Wild-type and *tmm* stem epidermis from the tip, middle and mature regions

(A-D) Wild-type stem from tip (A, B), middle (C) and mature (D) regions showing gradual difference in epidermal layer cells during development.
The initial asymmetric division of a MMC (blue arrowhead), asymmetric amplification division of a meristemoid (white arrowhead), GMC that will divide symmetrically (orange

Figure 6. Continued
arrow) and stomata with vertical pore (white arrow) are shown. The enclosed box shows an example of an inward spiral division.
(E-H) *tmm* stem epidermis from tip (E, F), middle (G) and mature (H) regions shown for comparison with wild-type. Asymmetric division of a MMC (orange arrowhead) and meristemoids (white arrowheads) that do not differentiate to form stomata are shown.
(I) In wild-type stems, epidermal cells undergo asymmetric divisions to form triangular meristemoids (dark and light red). If the smaller cell formed towards apex of the stem then it was considered to be an apical meristemoid (AM) but if it formed towards the base of the stem then it was defined as a basal meristemoid (BM). The meristemoids were distinguished from the GMCs (light green) by their oval shape.
(J) Satellite meristemoids (SM) are formed in wild-type stems.
(K) Inward spiral divisions occur in wild-type stems. A MMC (green) undergoes a formative asymmetric division to produce a primary meristemoid (blue) that undergoes amplification asymmetric divisions to renew the meristemoid (yellow) towards inside of the cell lineage. An additional asymmetric division orients the new (red) meristemoid towards the interior of the cell lineage. Ultimately the meristemoid (red) differentiates into a GMC mostly or completely surrounded by clonally related cells (green, blue & yellow).

Stomata are oriented relative to stem axis in the Arabidopsis stem

Stomatal pores in wild-type stems are always parallel to the long axis of the stem. Quantification of the orientation of symmetric cell division in GMCs was performed to determine if a stereotype division was responsible for longitudinally oriented pores. For this, the wild-type pictures of the stem middle region were chosen because in this region there are GMCs that are undergoing symmetric divisions. My observation showed that approximately 84% of symmetric divisions were parallel to the longitudinal axis of stem. This significant number of GMCs undergoing longitudinally oriented asymmetric divisions (Fig. 6C) showed that orientation of these divisions with respect to stem long axis is responsible for longitudinally oriented stomatal pores in stems (Fig. 6D), rather than by a mechanism that adjusted pore angle after GMC division.

Table 1. Comparison of the characteristics of the dicot leaf and stem with the monocot leaf

Feature	Dicot leaf	Dicot stem	Monocot leaf blade
Morphology	Flat and oval	Cylindrical and long	Flat and long
Spread of stomatal initials in young organ	Randomly distributed	At the tip region	At the base region
Time period of existence of stomatal initials	Longer	Shorter, during initial stages of growth	Shorter, during initial stages of growth
Pavement cell shape	Puzzle-piece or irregular	Rectangular or trapezoidal	Rectangular or trapezoidal
Asymmetric division forming meristemoid	Oriented with respect to surrounding cell	Oriented with respect to surrounding cell, not with respect to stem polarity	Polarized and oriented, not sure if they are meristemoids
Mature stomatal arrangement with respect to stomatal initials	Randomly distributed and mixed	Continuum, initials at the tip and mature stomata at the base	Continuum, initials at the base and mature stomata at the tip
Orientation of stomatal pore	Random relative to organ axis	Most of them are parallel to longitudinal axis of stem	Parallel to longitudinal axis of organ
Amplification divisions in meristemoids before formation of stomata	Present	Present	Not found in grasses
Probable Stomatal patterning mechanism	Cell-cell communication	Not known	Cell lineage mechanism
Minimum 1-cell spacing between two adjacent stomata	Present	Present	Present
Satellite meristemoids	Present	Present, fewer compared to leaf	None

Stomatal precursors are formed in tmm stems

To determine why there are no stomata in *tmm* stems I determined if there was initiation of the stomatal lineage. Asymmetric cell divisions in the epidermal layer are indicators of the stomatal lineage initiation. I used ten *tmm* plants containing plasma membrane marker Q8 and chose tip region pictures to see if asymmetric divisions occur (Fig. 6E-F). This showed that asymmetric divisions do occur in *tmm* stems as they do in wild-type stems. The smaller cells that formed from asymmetric cell divisions were recognized to be meristemoids by their cytological characteristics such as small size and triangular shape (Zhao and Sack, 1999). Hence, I conclude that meristemoids are formed in *tmm* stems.

Significantly more meristemoids are formed in tmm than in wild-type stems

Though stomata are lost in the *tmm* mutant stems (Fig. 6H), the presence of stomatal precursors raised additional questions. In order to determine if there are more meristemoids in *tmm* stems, as there are in the *tmm* leaf, I quantified number of meristemoids in *tmm* and wild-type stem tips. Stem tips were taken from ten wild-type plants and ten mutant plants containing the plasma membrane GFP marker Q8 for counting meristemoids. I counted the stomatal precursors in a defined area of *tmm* stems and wild-type stems, and found that there is a significant difference in the number of stomatal precursors. In wild-type stem tips, on average there were 69 meristemoids in a 2.6×10^4 μm^2 area while a similar region in *tmm* stems contains 102 meristemoids in

a $2.6x10^4$ μm^2 area. So, there are ~43% more meristemoids in *tmm* stems (Fig. 5A). Significantly more meristemoids in *tmm* stems indicate that TMM may negatively regulate formation of meristemoids in wild-type stems. It suggested that in stems (as in leaves) TMM regulates entry into the stomatal pathway and formation of meristemoids.

Asymmetric divisions in tmm stem tips are oriented towards the stem apex

To determine if the meristemoids in *tmm* stems are formed with polarity determined by the stem apical-basal axis, meristemoid division orientation was quantified. For this I used ten *tmm* plants containing plasma membrane marker Q8 and chose pictures from the tip region of the stem for assessing the orientation of meristemoids. First, asymmetric divisions were classified into three categories based on their orientation relative to the stem axis: apical, basal and lateral. This showed that there are a significant number of apical meristemoids compared to basal and lateral meristemoids. It demonstrates that in the absence of functional TMM, the asymmetric divisions in *tmm* stem tips show an orientation bias. Orientation of cell divisions is required to create one cell space between adjacent stomata or to avoid cluster formation in leaves, but it is not clear why meristemoid formation would be oriented in *tmm* stems though they do not differentiate into stomata (Fig. 6E).

Tracing of stomatal lineage cells in tmm stems

I assayed several markers of stomatal cell lineage identity in the *tmm* shoot epidermis to determine at which stage the stomatal development arrests. The molecular character of stomatal development lineage cells was investigated using markers specific

for stomatal cell types. To investigate whether stomatal developmental initiation occurred in *tmm* stem epidermis I examined the expression of stomatal lineage markers such as *TMM, ET1967* and *GRL2. KAT1* was used to find if there are any GMCs. TMM, ET1967, GRL2 and KAT1 expression patterns are discussed in detail below.

TMM promoter driving GFP reporter expression confirmed meristemoids are formed in tmm stems

To determine whether the smaller cells formed from asymmetric divisions in *tmm* stems have the molecular identity of meristemoids, I observed the expression of *TMMpro::GFP* in *tmm*. The epifluorescence microscope pictures of wild-type (Fig. 7A, C) and *tmm* (Fig. 7B, D) stem tips showed that the *TMM* promoter is active in small cells undergoing asymmetric divisions that resemble meristemoids in geometry. Based on two observations, that asymmetric divisions are present in *tmm* stems and that the same cells strongly express the *TMM* promoter (Fig. 7B,D), I conclude that meristemoids are formed in *tmm* stems. But the absence of stomata in *tmm* stems indicates that these meristemoids do not differentiate into stomata.

Comparison of wild-type and *tmm* stem tips shows that more cells in *tmm* stems express this marker. This serves as evidence for supporting meristemoid count studies that show more meristemoids are formed in *tmm* stems.

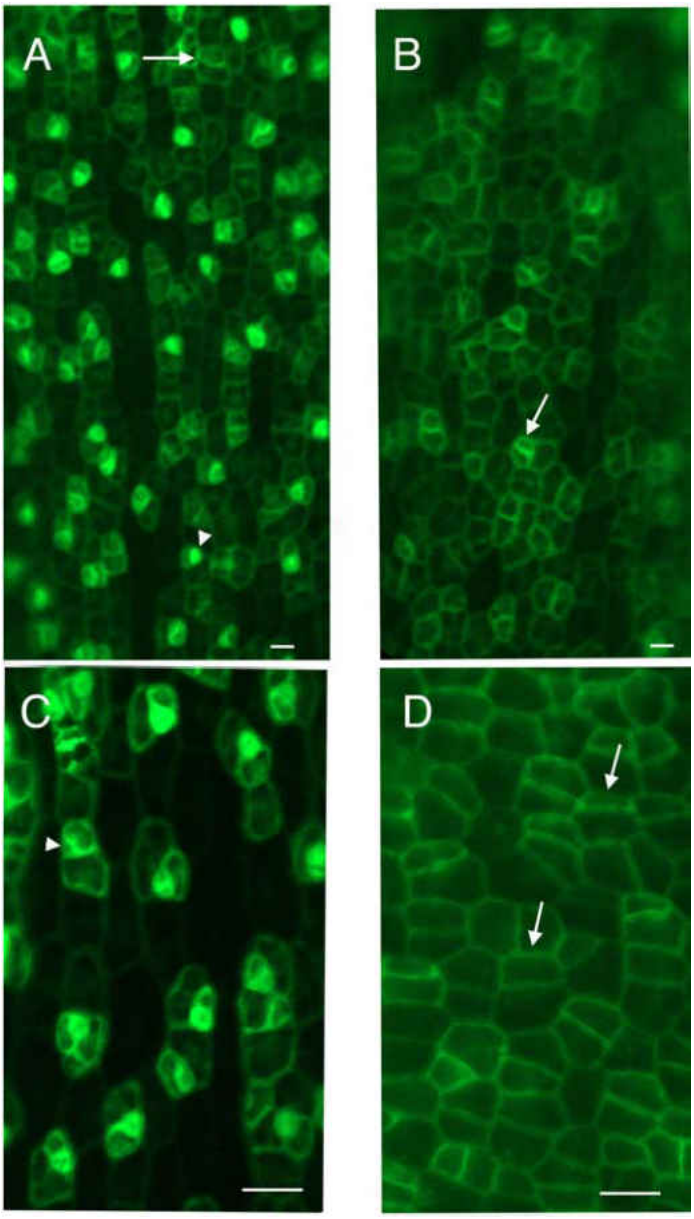


Figure 7. Wild-type and *tmm* stems expressing GFP from the *TMM* promoter (*TMMpro::GFP*)

(A, C) Wild-type stem tips.
(B, D) *tmm* stem tips. GFP expression in small triangular cells of *tmm* stems showed that they have molecular identity of meristemoids.
Confocal microscope images showing meristemoid mother cells undergoing first asymmetric division (arrows) and meristemoids undergoing amplification asymmetric division (arrowheads). A and C were seen with 20X. B and D were seen with 40X.

The stomatal lineage cell marker ET1967 is expressed in tmm stems

A second molecular marker that has an expression pattern similar to TMM (Enhancer Trap line 1967) was used for independent assessment of asymmetric cell divisions in *tmm* mutants. Wild-type stems (Fig. 8A, D, G) and cotyledons (Fig. 8E) containing ET1967 transgene fused to GUS reporter were used as positive controls. Like TMM even ET1967 is also expressed most intensely in meristemoids but less intensely in recent sister cells and young stomata of wild-type stems and cotyledons. In *tmm* stem tip and middle region, appearance of darker GUS staining in smaller cells resulting from asymmetric divisions suggested that meristemoids are formed (Fig. 8B, C, F). Using ET1967 marker it was shown that meristemoids are probably formed in *tmm* stems

In *tmm* stems, all sister cells in tip, middle region and mature region showed diffused expression of this marker (Fig 8B, C, F, H). This could be a technical problem, or it may reflect a difference between genotypes. In wild-type stems, the middle region and mature region did not show non-specific expression.

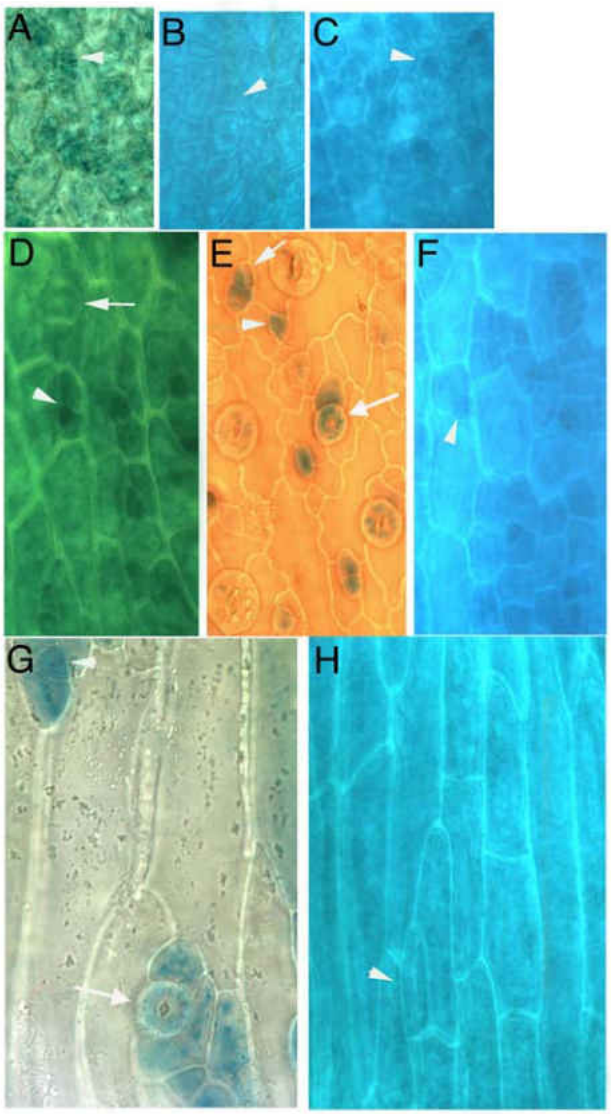


Figure 8. GUS staining of stomatal lineage cells in wild-type and *tmm* stems containing ET1967

ET1967 fused to a GUS reporter was used as a marker to show that the stomatal lineage is initiated in *tmm* stems.
(A, D, G) Wild-type stems from the tip, middle and mature regions, respectively. ET1967 expression is present only in stomatal lineage cells and absent from other epidermal cells.
(B, C, F, H) *tmm* stems from the tip (B,C), middle and mature regions, respectively. ET1967 is expressed in all cells but is darker in stomatal lineage cells.
(E) Wild-type cotyledon used as a positive control for GUS staining.
The meristemoids are darkly stained (arrowheads) but GMCs and GCs (arrows) are more faintly stained.

The meristemoid specific marker GRL2 is expressed in tmm stems

To make sure that the smaller cells formed from asymmetric divisions are definitely meristemoids I chose a meristemoid specific marker, *GRL2*. Transgenic plants containing the meristemoid marker (GRL2 promoter driving a GUS reporter) were expressing this marker strongly in meristemoids of wild-type (Fig. 9A, B) and *tmm* cotyledons (Fig. 10A). Faint expression was observed in larger sister cells that formed from asymmetric division and young stomata. Darkly stained small triangular cells resulting from asymmetric division in the tip region of wild-type (Fig. 9D) and *tmm* (Fig. 10B) stems were observed. This again confirmed that meristemoids are formed in *tmm* stems (Fig. 10). Staining experiment repeated many times and always the staining was weak, may be due to technical problem.

The middle region in wild-type (Fig. 9C) and *tmm* (Fig. 10C) stems showed expression of GRL2:GUS in meristemoids. In mature regions of wild-type stems (Fig. 9E), young stomata showed faint expression whereas meristemoids showed darker expression. In mature regions of *tmm* (Fig. 10D) stems there was no expression of the marker.

Comparison of wild-type and *tmm* stem tips shows that more cells in *tmm* stems express this marker. This also supports meristemoid count studies.

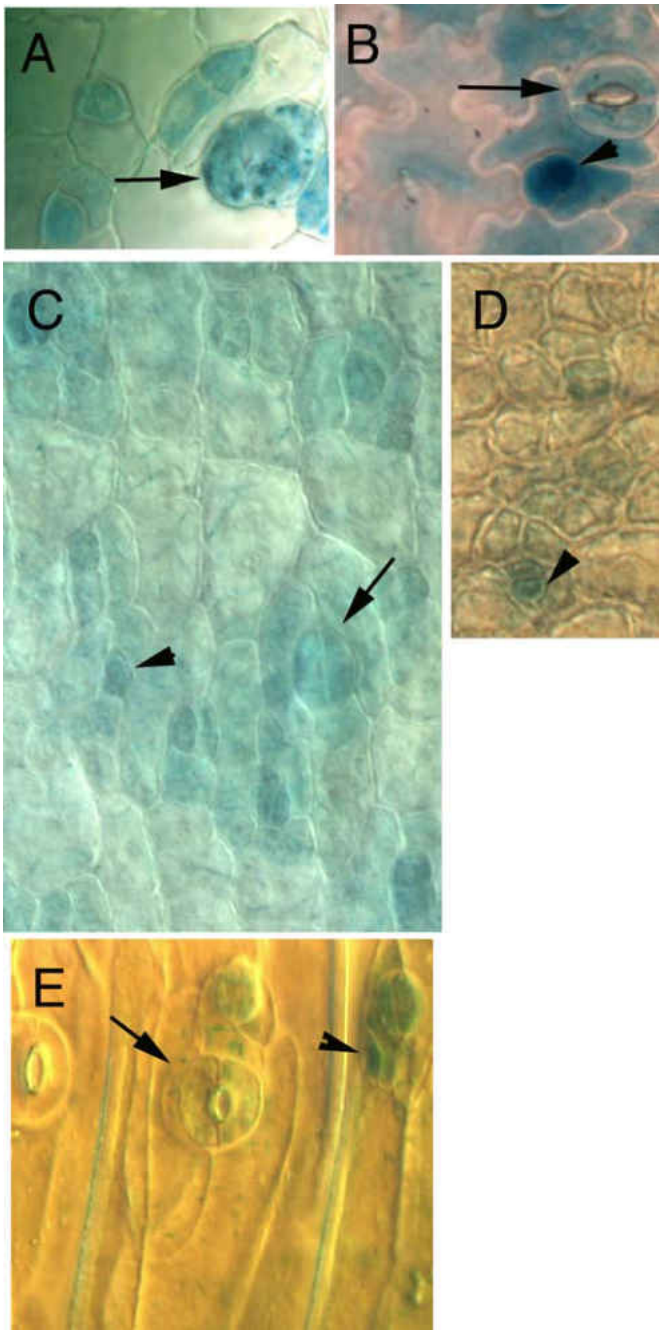


Figure 9. *GRL2pro::GUS* expression in wild-type stems and cotyledons

(A, B) The abaxial side of the wild-type cotyledons seen with 100x and 40X objectives, respectively. *The GRL2* promoter fused to a GUS reporter is expressed in meristemoids in this tissue.
(C, D, E) Wild-type stem epidermis in the middle, tip and mature regions, respectively seen with 100X objectives. As is stems, GRL2 is expressed darkly in meristemoids (arrowheads) and faintly in GMCs and GCs (arrows).

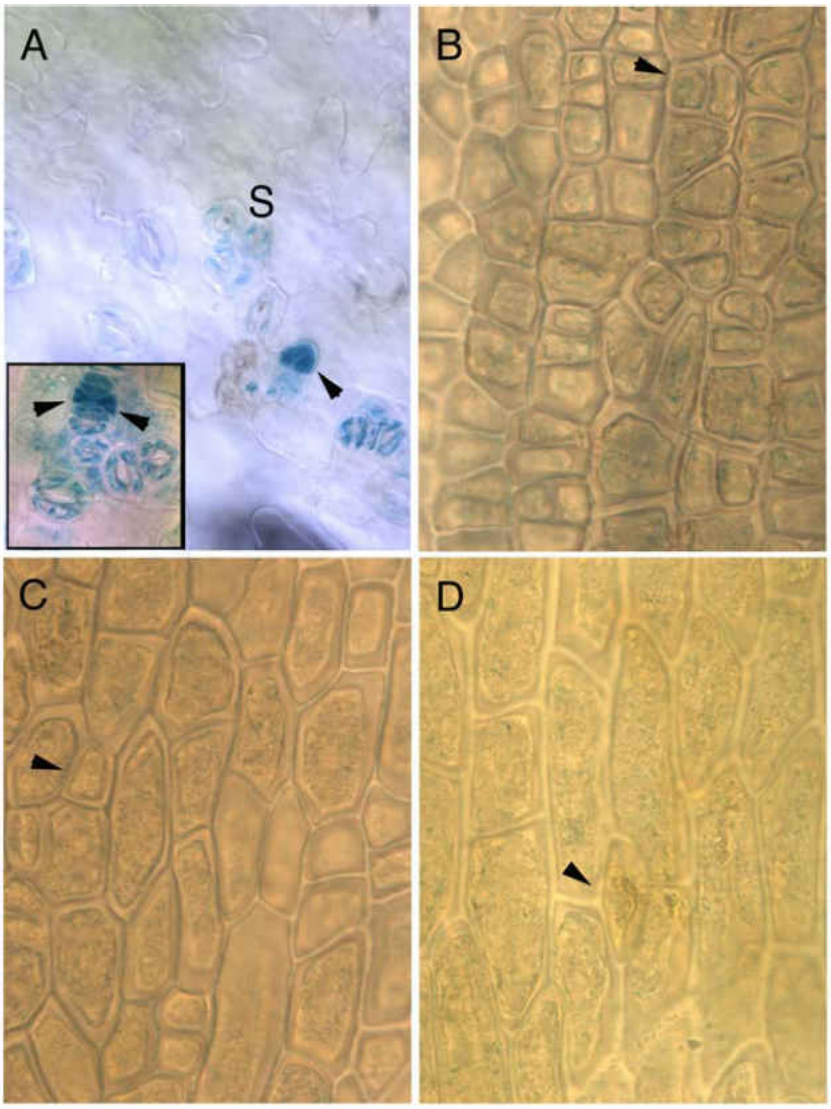


Figure 10. *GRL2pro::GUS* expression in *tmm* stems and cotyledons

The GRL2 promoter fused to a GUS reporter is expressed in meristemoids, confirming that the meristemoids are formed in *tmm* stems.
(A) The abaxial side of wild-type cotyledon seen with 40X objective used as a positive control. Inset showing meristemoids in a stomatal cluster. GRL2 is expressed darkly in meristemoids (arrowheads) and faintly in GMCs and GCs (S).
(B, C, D) *tmm* stems from the middle, tip and mature regions, respectively seen with 100X objective. Faint staining in the putative meristemoids in tip region (B) compared to other cells confirmed that the stomatal lineage is initiated in *tmm* stems.

The guard cell marker KAT1 is not expressed in tmm stems

The *KAT1pro::GUS* is a molecular marker for GMCs that I used to assess at what stage stomatal developmental arrests in *tmm* stems. The *tmm* stems containing this marker do not show GUS staining in tip (Fig. 11F), middle (Fig. 11G) or mature (Fig. 11H) regions. This indicates that stomatal development has been initiated in *tmm* stems but likely did not proceed to the GMC stage. The cotyledons of *tmm* and wild-type plants, and wild-type stems were used as positive controls. Wild-type (Fig. 11D) and *tmm* cotyledons (Fig. 11E) showed KAT expression darkly in mature stomata and very lightly in young stomata. In wild-type stems, there was very faint expression in GMCs in the middle region (Fig. 11B) and strong expression in guard cells in the mature region (Fig. 11C). Absence of *KAT1::GUS expressions* in *tmm* stems showed that stomatal precursors are formed but did not proceed to guard mother cell stage (Fig. 11).

On the basis of stomatal cell molecular marker studies I conclude that meristemoids are formed in *tmm* stem epidermis but they do not progress to the guard mother cell stage. This indicates that TOO MANY MOUTHS is required for progression of meristemoid to GMC in stomatal development in stem epidermis.

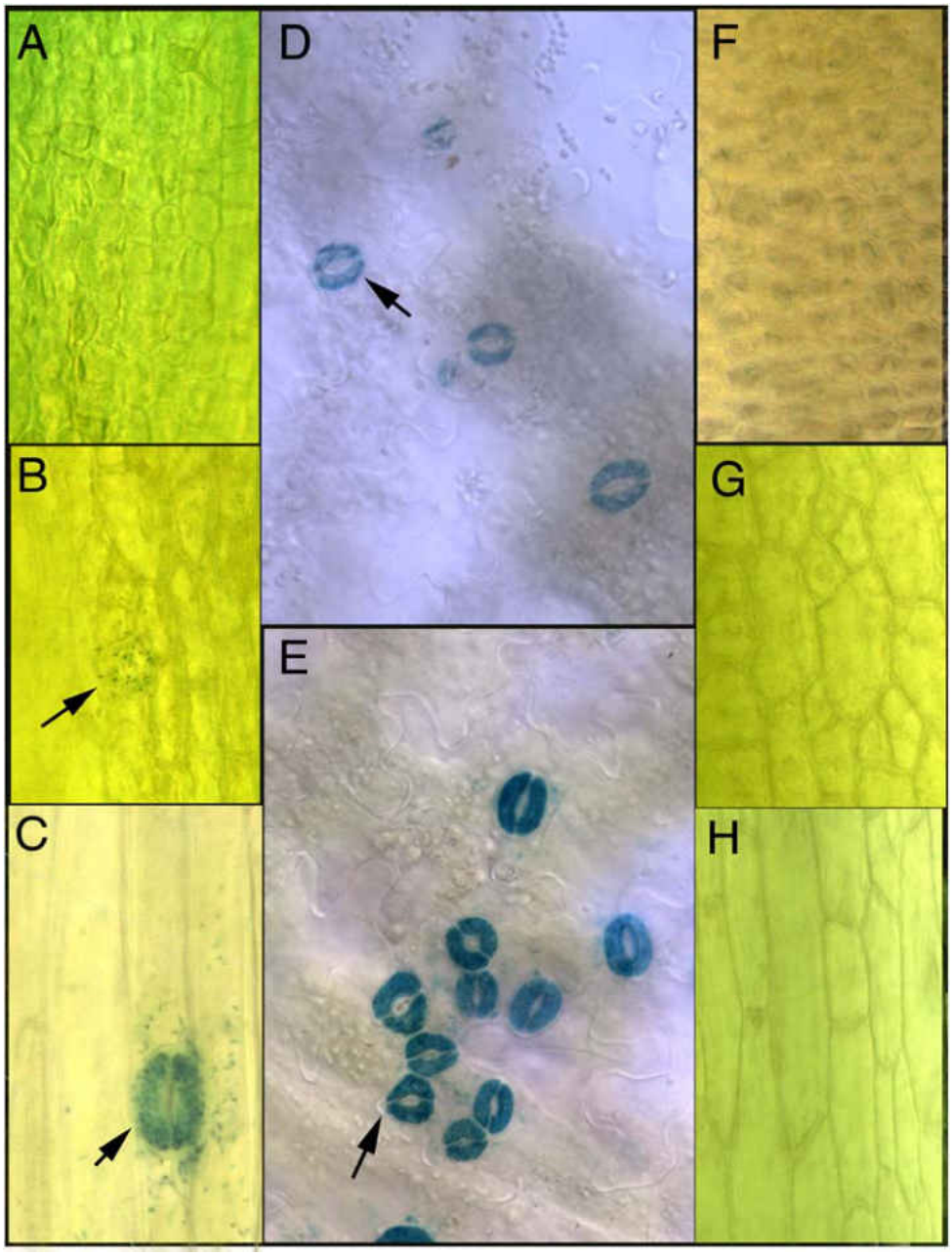


Figure 11. *KAT1pro::GUS* expression in wild-type and *tmm* stems

KAT1 is expressed in only in wild-type stems, showing that no GMCs are formed in *tmm* stems.
(A, B, C) Wild-type stems from the tip, middle and mature regions, respectively.
(F, G, H) *tmm* stems from the tip, middle and mature regions, respectively.
(D) The abaxial side of wild-type cotyledon used as a positive control.
(E) The abaxial side of *tmm* cotyledon also used as a positive control.
GUS stained GMCs and GCs (arrows) seen with 40x objective.

Discussion

Stomatal patterning in the dicot stem is different from the monocot leaf, despite organizational similarity

To understand the stomatal development and patterning in dicot stems, I hypothesized that asymmetric divisions forming meristemoids occur relative to the stem axis, so as to create one-cell spacing between adjacent stomata. The polarized asymmetric cell divisions that occur in monocot leaves are responsible for one-cell spacing (Hernandez *et al*., 1999). Quantification of orientation of meristemoids meristemoids in wild-type *Arabidopsis* stems showed that they do not form relative to the apical-basal axis of the stem (Fig. 5B). In this comparison study the lateral meristemoids are not taken into consideration because they are not formed in maize leaves and also they did not follow the apical-basal axis of the stem. Based on this result, the hypothesis that dicot stems orient the asymmetric divisions that form the meristemoids relative to the stem axis like in monocot leaves has been rejected. Here I conclude that although Arabidopsis stems are similar to monocot leaves in some features (Table 1), they probably do not follow the same stomatal patterning mechanism. This indicates that Arabidopsis stems use a different mechanism for spacing stomata. My assumption is that, spacing between adjacent stem stomata could be regulated by limiting the acquisition of MMC fate in protodermal cells or by lateral inhibition between stomatal precursors. This provided evidence to demonstrate that morphology of the organ might not always have a role in patterning.

Similarities in stomatal development between the Arabidopsis stem and leaf

Asymmetric divisions of meristemoids appear to be similar in leaves and stems of Arabidopsis (Fig. 6B). Although the formative asymmetric divisions in stems are randomly oriented as they are in leaves, amplifying divisions are frequently directed inward and follow a spiral pattern Fig. 6K). This indicates that initial stages of stomatal development are similar in Arabidopsis leaf and stem (Table 1).

In dicot leaves stomatal patterning and spacing is achieved by random generation of stomatal initials at early stages of leaf development with spacing maintained by lateral inhibition after meristemoids are formed. Amplification divisions lead to the production of pavement cells that surround future stomata in leaf (Larkin *et al*., 1997). Thus local spacing of individual stomata from close neighbor stomata is achieved. This local spacing hypothesis cannot be applied to all the stomata, as the number of amplification divisions is not sufficient to completely surround the stomata.

Formation of inward spiral asymmetric divisions in meristemoids is a stage in the development of stomata in leaves that allows local spacing between adjacent stomata. Appearance of this stage in stems suggests that dicot stems might be following the similar stomatal development and patterning mechanism. Future studies are required to prove this.

Longitudinally oriented symmetric divisions occur in GMCs of stems

In contrast to the leaf, the long axis of all mature stomata was parallel to the long axis of the organ. Quantitative analysis of the orientation of symmetric cell divisions in GMCs was performed to determine if this division occurred with stereotypical

orientation, or whether it was altered afterward during organ elongation growth. My results showed that symmetric divisions in GMCs always occur parallel to the stem long axis. Hence I accept the hypothesis that stomatal pores in mature region are longitudinally oriented because of the orientation of symmetric division in GMCs. I reject the hypothesis that stomatal pores align to the long axis of stem because of the cell expansion that changed the cell arrangement to align the longitudinal axis of the pore after the symmetric division occurred.

Meristemoids are formed in tmm stems

Mature stems of *tmm* plants have no stomata. To better understand this defect, I examined stems during the early stages of development. Visual observations showed that asymmetric divisions occur in *tmm* stems. This prompted a more thorough analysis of *tmm* plants using a plasma membrane GFP marker to reproducibly illustrate the outlines of cells (Fig. 6). Meristemoid-like cells formed from asymmetric divisions of the stem epidermis are small in size and triangular in shape, as are leaf meristemoids. Based on these cytological characteristics, these cells are likely to be meristemoids that do not ultimately form stomata.

Meristemoids did not differentiate to the GMC stage in tmm stems

To determine if the smaller cells formed in asymmetric divisions have the molecular identity of meristemoids I used three cell-type specific reporters, *TMMpro::GFP* (Fig. 7), the GUS enhancer trap line ET1967 (Fig. 8), and *GRL2pro::GUS* (Fig. 9, 10). The *TMM* promoter driving GFP expression was used as a marker for

young cells of the stomatal lineage, including SLGCs (Stomatal Lineage Ground Cells) and meristemoids. In *tmm* mutant stems, this marker showed small triangular cells with intense GFP expression that were most likely meristemoids. The enhancer trap line ET1967 was used as a second, independent stomatal lineage marker. This reporter showed darkly stained meristemoids in the tip region and diffused expression in all other epidermal cells. The expression pattern of both transgenic lines supports my observation that cells of stomatal lineage identity, including meristemoids, are likely formed in *tmm* stems. Expression of *GRL2pro::GUS*, a meristemoid specific marker, at high levels in the small triangular cells also supports the idea that meristemoids are formed in *tmm* stems.

To determine whether stomatal precursors in *tmm* stems differentiated into GMCs before they failed to differentiate as stomata, I used the GMC and GC specific molecular marker *KAT1pro::GUS* (Fig. 11). Wild-type stem tissue used as a positive control showed faint expression in GMCs and darker expression in mature stomata. Absence of even faint expression in all regions of *tmm* stems showed that meristemoids did not proceed to the GMC stage. Hence, I conclude that the meristemoids are formed in *tmm* stems but they do not differentiate to the GMC stage. Therefore, I conclude that TMM is required for differentiation of meristemoids into stomata in stems but not for the initiation of stomatal lineage cells.

TMM regulates entry into the stomatal pathway in stems and leaves

TMM appears to regulate the number of meristemoids produced in stems. Quantification of the number of meristemoids showed that there are significantly more

meristemoids in *tmm* stem tips than in the same region on wild-type stems (Fig. 5A). Yang and Sack (1995) showed that there are more meristemoids in the *tmm* cotyledons compared to wild-type cotyledons, so TMM is a negative regulator. My research suggests that TMM is a negative regulator of meristemoid initiation in stems. This finding suggests that TMM regulates entry into the stomatal pathway in wild-type stems (Fig. 6) as it does in leaves. This contrasts with the idea that TMM is a positive regulator of stomatal precursor cell differentiation, since mature stomata are absent from *tmm* stems. In other words, TMM is required for formation of meristemoids as well as development/differentiation of meristemoids to stomata in stems. This shows that in spite of noticeable similarities between leaves and stems, there are some differences in stomatal development in different organs, possibly due to a difference in the protein interactors or the underlying mechanisms involved.

Asymmetric divisions are randomly oriented in wild-type stems but are oriented relative to the stem axis in tmm

To determine if the *tmm* mutation alters the orientation or polarity of stem meristemoid formation the orientation of meristemoids in *tmm* stem tips was quantified. In wild-type stems there are an almost equal number of apical and basal meristemoids (Fig. 5B), however, most of the meristemoids are apically oriented in *tmm* stems. It is possible that in the absence of functional TMM, meristemoids use a default pattern of division that results in apical asymmetric divisions. Perhaps in wild-type stems with functional TMM, meristemoids are not following this default orientation. But in the mature wild-type stem there are non-randomly spaced stomata and there are almost no

clusters or pairs of stomata. This could be because few cells at the stem tips assume the stomatal lineage fate despite an absence of signaling to control spacing. This result, which may hint at the presence of a default orientation for stem meristemoids, may provide limited evidence that TMM is involved in orientation of asymmetric division as well.

Future study

There are additional questions apparent from my work that needs further examination. Observation of *tmm* stems showed that at the base of the stem there are a few stomata (Fig. 12D) and occasionally there were pairs of stomata (Fig 12E). The rare occurrence of stomata and some in pairs in *tmm* stems is consistent with my discovery that meristemoids are more numerous in *tmm* stems. It is not known why some meristemoids are able to form stomata but most are not.

Almost all the GMCs form symmetric division parallel to the long axis of the stem that results in formation of longitudinal stomatal pore (Fig. 12A). It would be interesting to know the underlying mechanism for uniformity in the orientation of stomata in stems.

In mature region of Arabidopsis stem I observed that occasionally some meristemoids arrested and did not differentiate into stomata. It would be useful to know if this was because cell-cell communication was occurring to regulate the number of stomata and inhibiting differentiation in some meristemoids.

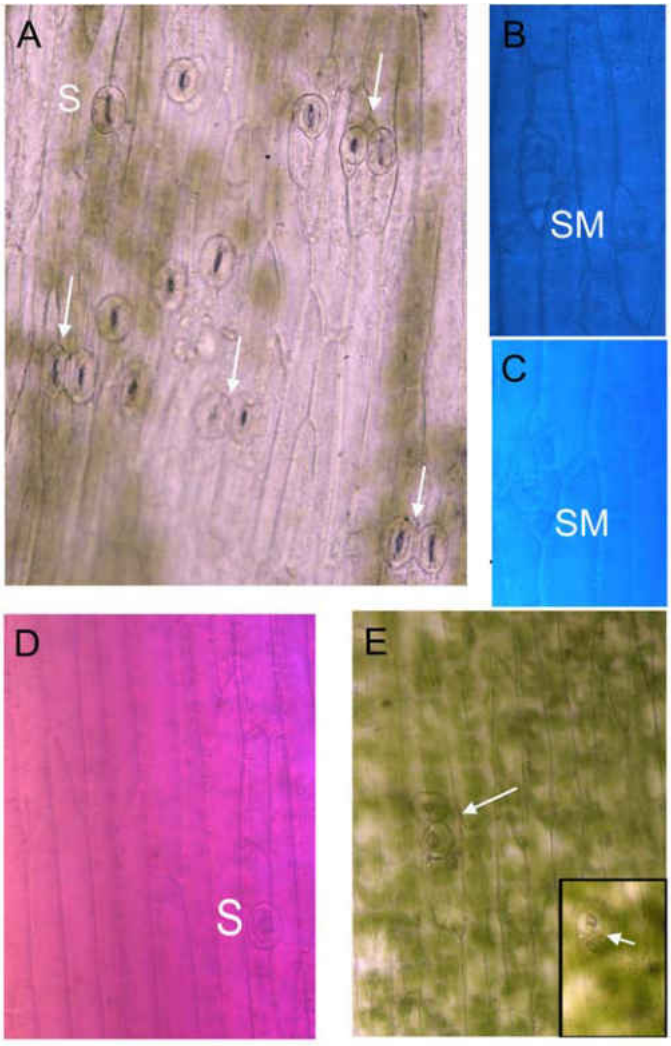


Figure 12. Uncommon features in wild-type and *tmm* stems

(A) Wild-type stem epidermis from 30 day old plant showing stomatal pairs (arrows) formed at the mature end of stem.
(B, C) Wild-type stem epidermis showing satellite meristemoids (SM). Stems stained with toluidine blue.
(D) *tmm* stem from mature region showing rare occurrence of stomatal (S) formation. Stem stained with safranin.
(E) *tmm* stem epidermis from mature region showing rare formation of stomatal pairs (arrows). Inset showing the stomatal pair (arrow) containing one defective shaped stomata in the pair.
Light microscope images with 40X magnification.

CHAPTER 2:
GENE EXPRESSION PROFILE OF STOMATAL STEM CELLS

Introduction

The stem in *Arabidopsis* is an excellent system to examine plant stem cells because it shows a gradient of different stages of stomatal development from tip to base. In the wild-type (30 day old) stem tip (2mm from apex) there are meristemoids (plant stem cells) that renew themselves and finally differentiate into stomata. In Chapter 1, I showed that meristemoids are formed in *tmm* stems (Fig. 6E), but that these fail to differentiate into stomata. The contrast between normal and mutant stem tissues allowed us to use the Arabidopsis stem tip (Fig. 4) to generate contrasting gene expression profiles to identify genes that are important in meristemoid and stomatal development.

Materials and methods

RNA extraction and hybridization

The *Arabidopsis* inflorescence stem is a cylindrical organ that exhibits a gradient of cells at different stages of differentiation. The youngest cells are found at the tip in a zone of cell proliferation, while the middle region contains a mix of dividing and differentiating cells and the basal region contains mature, fully-expanded cells. The tip of the stem in a region containing only stomatal precursor cells (but not stomata, Fig. 2C) was selected for the experiment. Plants were grown in a Percival AR-66L growth chamber with 16 hours of light of approximately 150µmol/m^2 per sec intensity, 70% humidity, at 20-22°C. Plants were manually watered once every 2-3 days as needed

and fertilized 1X week with Peter's 20-20-20 (1.27g/L). Six seeds were planted in each of 21pots in a tray, each pot was 7 x 7.5 cm wide, 6 cm tall, filled with autoclaved and prewet Promix BX soil (Hummert International). One month after sowing the plants had an approximately 7.5 cm inflorescence stem. Each stem was individually removed from the plant for dissection of the tip region. First, all flowers and buds were separated from the stem with #5 tweezers at the base of the pedicels near the junction with the stem. The inflorescence meristem and very young buds were also cut off. Then a 2 mm length of stem (Fig. 5) was cut and immediately dropped in 1.5 mL microcentrifuge tubes containing liquid nitrogen (LN_2) floating in a LN_2 bath. After collecting 15- 20 stem tips, each tube was stored in an -80°C freezer. This process was repeated until 150 tips were collected from Columbia *gl1* wild-type (Col) and *tmm-1* mutant plants (also in the Col background) so as to obtain approximately 200 mg of tissue. Before RNA extraction tips from separate tubes were pooled to yield two samples of ~100mg each, and then ground to a fine powder using a mortar and pestle previously chilled to -80°C. Total RNA was immediately extracted from ground tissue using the Qiagen Plant RNA miniprep kit according to manufacturer's instructions and the specifications provided by the Nottingham Arabidopsis Stock Center (NASC) Microarray Facility for concentration, storage and shipping of RNA samples. Two samples for each condition (wild-type or *tmm*-1) representing biological replicates were mailed to NASC for probe synthesis and hybridization to the Affymetrix ATH1 genechip. NASC RNA quality control included analysis of degradation assessed from rRNA peaks using an electropherogram (Appendix C). NASC conducted the chip hybridizations and chip scanning, and

provided raw data about intensity values and minimally processed data to us in electronic format.

Quality assurance

Possible errors associated with this high-throughput technology are encountered at several levels such as tissue sampling, RNA extraction, probe synthesis, printing of chip, hybridization, scanning and gene expression analysis. Hence care is taken at every step to recognize any possible errors. While collecting the sample tissue, care was taken to treat all the samples as described in the methods. Tissue was immersed and stored in liquid nitrogen to avoid degradation of RNA by RNAses present inside the tissue. After the RNA was extracted and mailed to NASC it was tested for quality by capillary electrophoresis to produce an electropherogram that gave sharp peaks for rRNA bands (Appendix C). The presence of sharp peaks indicated that RNA received by them was of good quality.

Affymetrix chip details

The Arabidopsis ATH1-121501 genome chip contains 22,500 oligonucleotide probe sets representing more than 24,000 *Arabidopsis* genes as well as control features. Data used to design this array is based on the information obtained from IASP (International Arabidopsis Sequencing Project) in December 2000. On this genechip some similar genes were represented by non-unique probe sets hence there are fewer probe sets (22,500) than represented genes (24,000). This genome chip also does not contain probe sets for all genes in the Arabidopsis genome (~ 26,200 genes).

Oligonucleotides in ATH1-121501 gene chip are 25-mer (probe length) long. There are 11 probe pairs per sequence. The control sequences used on this array were- *E.coli* genes *bioB, bioC*, *bioD*, Phage P1 *cre* gene, *B.subtilis* gene *lysA*, common (or maintenance) *Arabidopsis* genes such as actin, ubiquitin and GAPDH.

Gene expression analysis

Differentially expressed genes by Genespring analysis

Standard Affymetrix software Genespring GX 7.3.1 (Agilent Technologies, Palo Alto, CA) was used for normalization and statistical analysis of changes in gene expression. Raw data from CEL (Cell Intensity values) files were normalized using Robust Multichip Averaging (RMA) and then subjected to a conventional t-test statistic. RMA normalization (Irizarry *et al*., 2003) involves three steps; background adjustment, quantile normalization and summarization. This generates a normalized and summarized file that has been split into two groups based on the treatment (Col wild-type and *tmm-1*).

The data was grouped by tissue type (wild-type/Col & *tmm-1*) and variances were assumed to be equal for the parametric test. A Benjamini and Hochberg correction was selected, and only genes with p-value<0.05 were considered. This restriction selected ~1,147 differentially expressed genes, of which about 56 (5.0%) would be expected to pass the restriction by chance. The list was then filtered for genes that change by more than 1.5 fold, resulting in a final list of 260 genes that are differentially regulated between wild-type and *tmm-1* stem tissues (Appendix A).

These 260 differentially expressed genes were categorized into different groups based on function to see which categories are affected by TMM activity. Graphs showing proportion of genes falling into key gene ontology groups were generated. Genes from interesting categories that also had the lowest p-value were selected for further analysis.

An Affymetrix algorithm that was designed to assess whether transcripts for any particular gene are present in the sample was utilized. Each gene is assigned either a P (for present), M (marginal) or A (absent) based on the expression value comparison between that gene and 11 matching probes and 11 mis-match probes. If there is significant difference in the expression value and the high number corresponds to matching probes then it indicates that the gene is present, otherwise it is not considered to be expressed above background noise in the hybridization. The numbers of genes present in different samples were different but approximately there were 11,048 genes present in all samples (Appendix D).

Analysis by RACE

The microarray data obtained was also analyzed using a Bayesian statistical method using the web-based tool Remote Analysis Computation for gene Expression (RACE) (Psarros *et al*., 2005). This tool also carries out Robust Multichip Averaging prior to analysis (Irizarry *et al*, 2003). All the genes with Bayesian p-value <0.05 and fold change >1.5 were considered differentially regulated. The final list contained 352 differentially regulated genes (Appendix B).

Categorization by Gene Ontology

Graphs were used to visualize the functional categories of genes that are differentially expressed. Gene ontology (GO) is a convention for categorizing genes into functional categories based on biological information available (www.geneontology.org). GO provides three ways of categorizing genes; based on biological processes they are involved with, the molecular function of the putative protein, or cellular component the protein is likely found in. The biological processes are divided and organized into different hierarchical levels based on the latest research evidence available. Sometimes one gene can be categorized under two different networks where these networks tend to entangle.

Bioinformatics analysis

The web-based tool DAVID (database for annotation, visualization and integrated discovery) was used to determine if any classes of genes were overrepresented in the list of differentially regulated genes. DAVID has three advantages as a gene functional classification tool: grouping large list of genes into separate classes based on their functional similarity, searching for other non-represented genes from the genome that are functionally related to some of the interesting representatives from the gene list, and two-dimensional visualization of genes and their annotations in each functional cluster or group. The 260 genes obtained from Genespring analysis were analyzed using the DAVID tool. Classification stringency was set to “medium” to allow more functional groups to appear in the list and avoid many genes appearing in an unclustered group.

Kappa similarity was set to 4, the threshold was set to 0.35, and classification multiple linkage threshold was set to 0.5.

For comparison, I also used the web-based tool ATTED II to produce a list of genes co-regulated with *TMM* (*At1g80080*). ATTED-II utilizes microarray data available in public databases to identify co-expressed genes using a Pearson correlation statistic (http://www.atted.bio.titech.ac.jp/). Genes identified as co-expressed across in multiple experiments were then compared to genes identified by t-test.

AtGenExpress visualization tool is a web-based tool that utilizes microarray gene expression data available in public databases. This tool was used to view expression pattern of the genes of interest (Schmid *et al*., 2005).

Protein characterization

There are links to various protein sequence analysis tools collected at the website ExPASY (Expert Protein Analysis System) that were used in my study for characterizing the proteins found in my study. ProP v.1.0b ProPeptide Cleavage Site Prediction tool (Duckert *et al*., 2004) was used to find potential recognition sites for subtilisin proteases in At1g34245. TargetP 1.1 server prediction tool (Nielsen *et al*., 1997) was used to predict signal peptides cleavage sites. SMART (Simple Modular Architecture Research Tool) (Schultz *et al*., 1998) was used to predict known domains in the proteins. PSORT version 6.4 (Nakai and Kanehisa, 1991) was used for prediction of protein subcellular localization.

Comparison of differentially regulated genes with other gene expression profiles to reveal potential stomatal regulators

In order to focus my study on genes that were more generally involved in either stomatal development or in mechanisms of cell proliferation, I compared my gene list to those of two other research groups. Microarray analysis performed was on leaf samples to identify the genes involved in regulating cell proliferation during organ growth (Beemster *et al*., 2005). Loss of the MAPKKK activity of *YODA* resulted in excess stomata on the leaf while excessive *YODA* activity resulted in the formation of all pavement cells with the loss of stomatal cells (Bergmann *et al*., 2004). They capitalized on this condition to conduct a microarray gene expression profiling experiment designed to reveal genes important to all stages of stomatal development. Publicly available microarray experimental data from Bergmann *et al* (2004) and Beemster *et al* (2005) were compared with my genelist to identify and narrow the list to the stomatal development specific genes. Excel software was used to compare the geneIDs of large lists and to determine the genes common between three microarray experiments. The Microsoft Excel formula used for finding overlaps was [=IF(ISERROR(MATCH(A7,$E$2:$E$2067,0)),"",A7)]. Obtained numbers were depicted as Venn diagrams showing the absolute numbers of genes in overlapping sets.

Phenotypic screening

To prioritize the mutants to examine, differentially regulated genes were sorted by significance of the change in expression level. Those with obvious housekeeping functions were omitted, and 65 genes that had T-DNA insertional mutants available

were ordered from the Arabidopsis Biological Resource Center (ABRC). In Appendix B these genes are shown in white, orange and blue highlighted rows, except yellow rows that did not have insertions or were not ordered for screening. In order to assess function of genes differentially regulated in my experiment, T-DNA insertional mutants in differentially regulated genes were obtained and examined for stomatal or other phenotypes.

Seeds from segregating stocks were planted 16 seeds per 4" pot, while seeds from homozygous stocks were planted at 9 seeds per pot. One cotyledon from each plant was taken and put on slide abaxial side up. Ten-day old cotyledons from each seedling were observed for defects in stomatal patterning, stomatal number, and cotyledon shape using an Olympus BX60 compound microscope. As plants grew, other defects in morphology were noted.

Results

Fraction of genes expressed in the treatments

The genelist obtained was filtered on present and absent flags to determine the fraction of genes expressed in replicates of each treatment (Appendix D). It showed that 52% of total genes from Arabidopsis genome were expressed in the wild-type (Col) stem tips and a little less than 51% of genes are expressed in the *tmm* sample. This shows that half of the genes from the genome are expressed in stem tip tissue. It is interesting that approximately the same number of genes is expressed in both tissue types with very little difference. This difference in the fraction/number of genes expressed between wild-type stem tips and *tmm* stem tips is largely due to mutation in

tmm. Among those genes expressed in two treatments I were interested in those that were differentially expressed. Differentially expressed genes in these treatments were those that were differentially regulated by *TMM*, or those that play a role in meristemoid cell fate or regulation.

Downregulated and upregulated genes in tmm stem tip tissue

From the 172 genes downregulated in *tmm*, an ANOVA was applied to show that 57 changes are likely to be statistically significant. 151 genes were upregulated in *tmm* and ANOVA analysis showed that 3 among them were statistically significant. In order to avoid overlooking genes of potential significance, all 260 genes obtained by conventional t-test were considered to be differentially regulated.

From the 352 differentially expressed genelist obtained by Bayesian statistics, 266 genes were downregulated in *tmm* and 86 genes were upregulated in *tmm*.

Table 2. Upregulated and downregulated genes in wild-type tissue and *tmm* mutant tissue

Regulation	**Conventional ANOVA**		**Bayesian statistics**
	1.5 fold	**ANOVA on 1.5 fold**	**1.5 fold & p-value<0.05**
Down in *tmm*	172	57	266
Up in *tmm*	151	3	86

Assessment of microarray expression profiling outcome

First, I determined whether the stem tissue harvesting technique was precise enough to eliminate guard mother cells and guard cells from samples. This was essential to avoid identifying genes that were differentially regulated simply because of the absence of guard mother cells and stomata from *tmm-1* samples. To verify that sample collected did not have any stomatal lineage cells from advanced stages like

guard mother cell or guard cells the genelist was examined by filtering on present and marginal flags and searched for specific genes known to be expressed in guard mother cells or guard cells. The guard cell specific *KAT1* gene is not expressed in wild and *tmm* samples, indicating low (or no) contamination with middle or mature regions of the stem.

Similarly, I also established that the stem sample tissue had cells from early stages of stomatal development because of the presence of transcripts from genes expressed in meristemoid stages, such as *TMM, YODA,* and *SPEECHLESS*. Interestingly, *TMM* transcripts were more abundant in *tmm-1* mutant stems than in wild-type stems. *SPEECHLESS* is supposed to be expressed in all the cell types but it is present only in one of the *tmm* and Col replicate samples, and absent from the other two. This may reflect a very low level of expression of this gene. In my experiment, samples contained genes from early developmental stages and negligible contamination showed that this was a successful attempt to harvest appropriate tissue.

Source of error

There were 4 samples in total in the microarray experiment with 2 replicates for each sample (1*tmm*1, 2Col, 3*tmm*1, 4Col). Among the four replicates used there was variation in the intensity of hybridization signal in one *tmm*-1 RNA sample (Fig. 13). It could have been during the cDNA synthesis hence the picture of this sample's genechip after hybridization is dull compared to other genechips. This difference in variation in the samples made the t-test a less appropriate statistical method for identification of differentially regulated genes. This source of error could have been avoided by using

three or more replicates for each sample but was not economically feasible because each hybridization costs ~$1200. In spite of this potential issue with the experimental sample I was still able to correctly identify several genes with a role in stomatal development, demonstrating that this method was an effective technique for gene discovery.

It is unlikely that remaining pedicel base tissue biases the microarray data. Geisler *et al* (1998) studied stomatal development in *tmm* pedicels and reported that there were more clusters of stomata at its apical end, fewer clusters in the middle and no stomata at the base. My observations of GUS stained pedicel bases showed that the meristemoids do form at the base of pedicels but these precursors also arrest and do not form stomata. This probably indicates that meristemoids at the pedicel bases are similar to the meristemoids of the stem, so their presence probably would not significantly alter the expression profile.

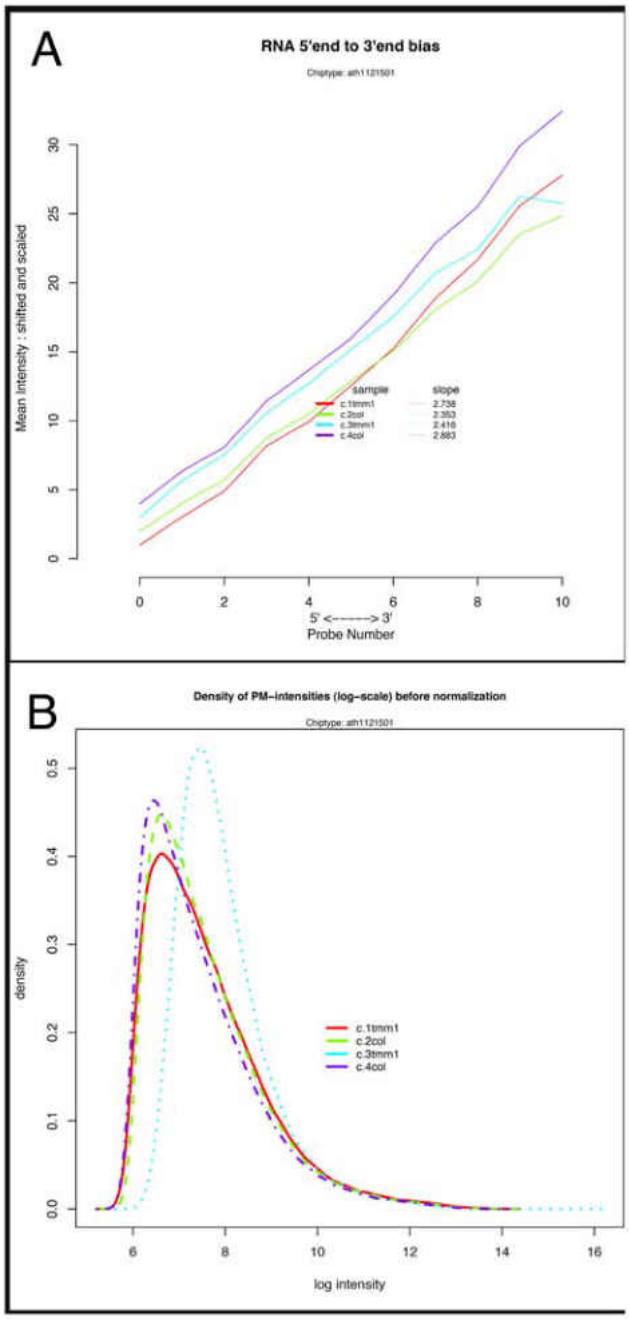


Figure 13. Analysis of microarray replicate samples by the web-based tool RACE

A. Mean intensity of the probe-labeled samples, showing that there is variation in one of the two *tmm-1* samples
B. Spot density versus log intensity

Classification of genes based on Gene Ontology

To gain a better understanding of the biological significance of the genes, Gene Ontology (GO) annotation was used to categorize the genes based on the biological process in which they participate. This helped to identify the general biological themes present in the differentially expressed genes identified. Gene Ontology classifies genes based on biological processes (Fig 14) into 6 categories, they are listed here with number of genes in brackets: physiological processes (116), cellular processes (101), regulation of biological process (21), response to stimulus (19), development (6) and growth (1). Categorization of 260 differentially expressed genes based on biological processes showed that large fraction (90%) of these genes is involved in cellular and physiological processes. 10% is involved in growth, development and response to stimulus. The aim of this experiment was to further characterize the differentially regulated genes and recognize biologically significant themes in the geneset. Inspection of the "development" genes (Fig. 15A) revealed that *TMM* that appeared (Table 3). Also in the "development" list was *At1g79700*, which is described as unknown protein similar to AP2 domain transcription factors. Because this could be a transcription factor and might be involved in cell signaling, I chose this for further study.

Of 260 differentially expressed 101 fell into the "cellular process" category. Further categorization of 101 genes showed that there are 95 genes that belong to cellular physiological processes, one gene falls into the cell differentiation category, 12 genes to cell communication and 20 genes to regulation of cellular processes. Among

these the cell communication category was chosen for further investigation because these are good candidates for involvement in stomatal patterning. Among the 12 genes involved in the cell communication category there were 11 genes that are categorized as involved in signal transduction (Fig. 15B) and 1 gene involved in response to extracellular stimulus. *TMM* was found in the list of signal transduction genes (Table 4).

This gene categorization allowed finding the genes with relevant function and helped us to select the genes for further study.

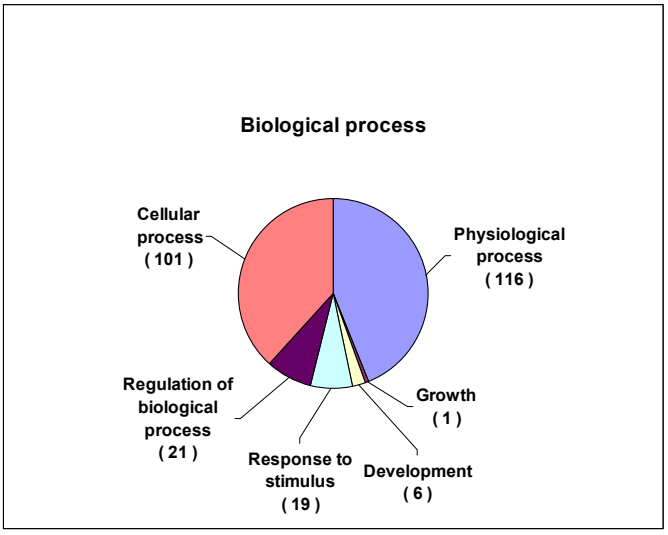


Figure 14. Differentially expressed genes categorized into different "biological processes" categories defined by Gene Ontology

These genes are derived from 260 differentially expressed genes. Number of genes are specified for each category.

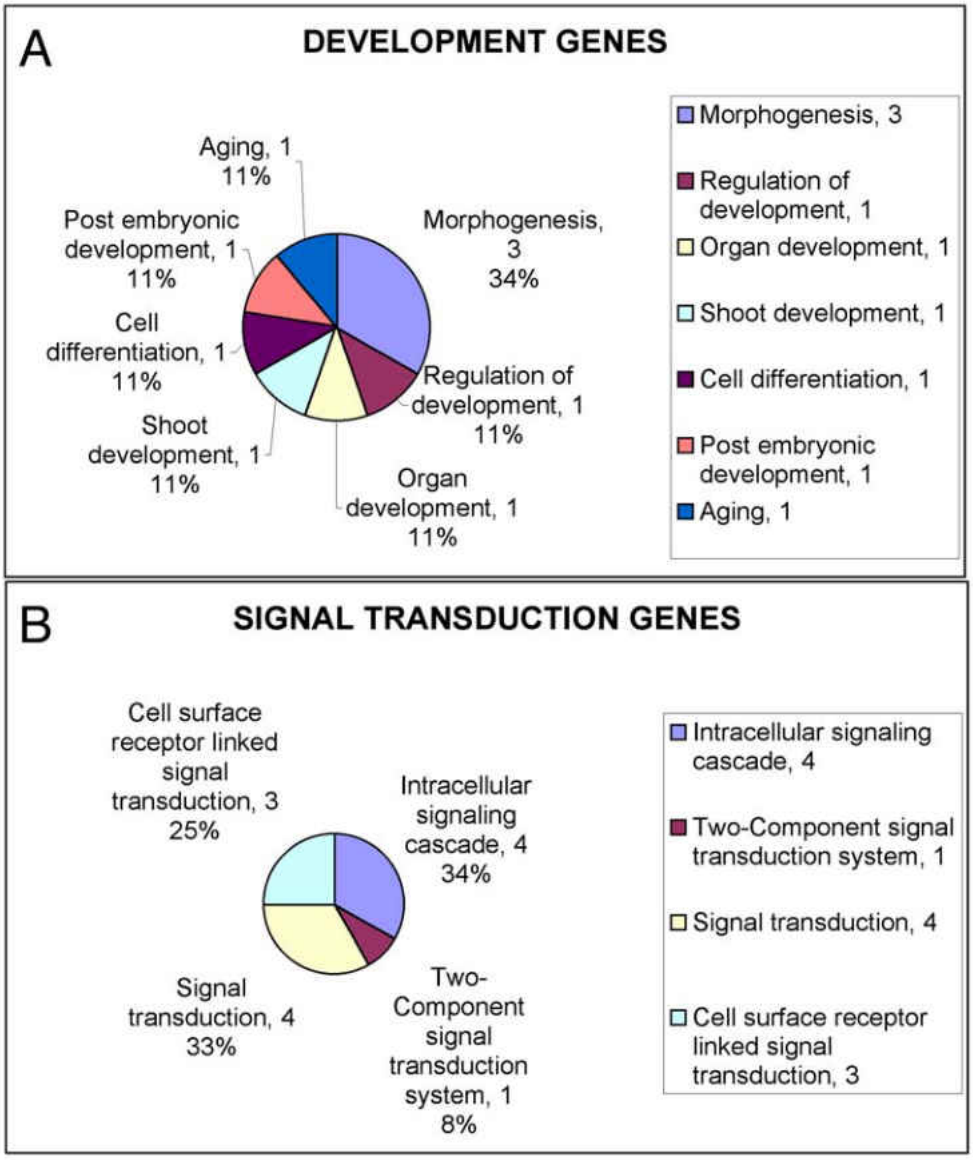


Figure 15. Development and signal transduction categories showing further classification of genes

These genes are derived from 260 differentially expressed genes. For each pie chart the number of genes and fraction of genes in the category are labeled.

Table 3. Genes involved in signal transduction. Bold font indicates known stomatal regulators *(TMM)*.

Genbank	Description
At5g58300	receptor-like protein kinase
At5g47220	Ethylene responsive element binding factor 2 (ATERF2) (sp|O80338); supported by full-length cDNA: Ceres: 3012.
At5g02430	putative protein rab11 binding protein, Bos taurus, EMBL:AF117897
At4g36180	Putative receptor protein kinase Cf-2.1 leucine rich repeat protein, Solanum pimpinellifolium, PATX: G1184075
At3g08510	phosphoinositide specific phospholipase (AtPLC2) identical to phosphoinositide specific phospholipase (AtPLC2) GI:857374 [Arabidopsis thaliana]; supported by cDNA: gi_13430587_gb_AF360206.1_AF360206
At3g13590	hypothetical protein predicted by genemark.hmm
At1g63830	unknown protein similar to putative protein GB:CAA20468 [Arabidopsis thaliana]
At1g80080	receptor protein kinase, putative similar to receptor protein kinase GI:1389566 from [Arabidopsis thaliana]
At1g73070	disease resistance protein, putative similar to disease resistance protein GI:3894383 from [Lycopersicon esculentum]
At2g31085	CLE4 CLAVATA3/ESR-Related 4 (CLE4); supported by full-length cDNA: Ceres: 270513.
At2g02680	hypothetical protein predicted by genscan

Table 4. Genes involved in development. Bold font indicates *TMM* and a putative transcription factor (*At1g79700*) which is further described by this research.

GeneName	Description	GO Biological Process
Aging		
At3g49620	Putative protein SRG1	aging; traceable author statement; cellular response to starvation; traceable author statement
Cell differentiation		
At5g11790	Putative protein SF21,	cell differentiation; inferred from sequence similarity; cell differentiation; inferred from electronic annotation
Development and morphogenesis		
At5g63840	Glucosidase II alpha subunit	unidimensional cell growth ; inferred from mutant phenotype ; cellulose biosynthesis ; inferred from mutant phenotype
At1g79700	Unknown protein similar to putative AP2 domain transcription factor	organ morphogenesis ; inferred from sequence similarity; regulation of transcription, DNA-dependent ; inferred from sequence similarity
At1g80080	Receptor protein kinase, putative similar to receptor protein kinase	asymmetric cell division; inferred from mutant phenotype; stomatal complex morphogenesis; inferred from mutant phenotype; signal transduction; inferred from curator
Organ Development		
At1g79700	Unknown protein similar to putative AP2 domain transcription factor	organ morphogenesis; inferred from sequence similarity; regulation of transcription, DNA-dependent; inferred from sequence similarity
Post embryonic development		
At5g03840	Terminal flower1 (TFL1)	response to sucrose stimulus; inferred from mutant phenotype; negative regulation of flower development; inferred from mutant phenotype

Table 4. Continued

GeneName	Description	GO Biological Process
Regulation of development		
At5g03840	Terminal flower1 (TFL1)	response to sucrose stimulus; inferred from mutant phenotype; negative regulation of flower development; inferred from mutant phenotype
Shoot development		
At1g80080	Receptor protein kinase, putative similar to receptor protein kinase	asymmetric cell division; inferred from mutant phenotype; stomatal complex morphogenesis; inferred from mutant phenotype; signal transduction; inferred from curator

Comparison with previously published gene expression profiling data reveals overlapping genes with potential roles in cell proliferation or stomatal development

We compared the stem dataset derived from the t-test statistic with leaf cell proliferation genes dataset (Beemster *et al,* 2005) and mutated *yoda* leaf sample dataset (Bergmann *et al, 2004)* to narrow the list of potential stomatal regulators. While most of the genes are present in only one of the datasets some were common in two datasets. There were no genes common among these three experiments (Fig. 16).

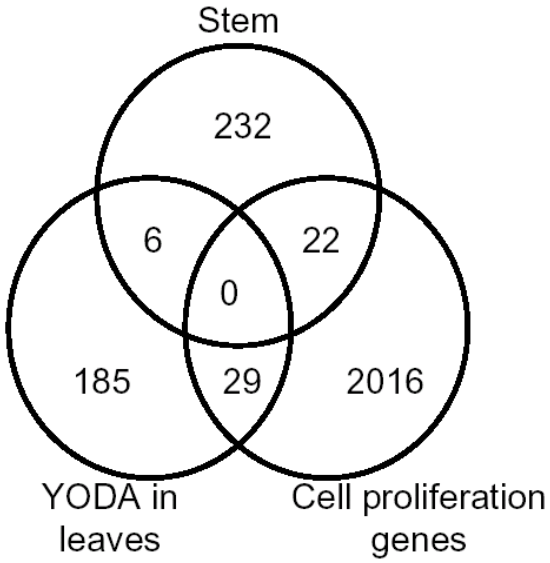


Figure 16. Venn diagram showing common genes between stem, YODA, and leaf cell proliferation microarray expression experiments

I used stem dataset containing 260 differentially regulated genes and cell proliferation genes dataset (Beemster *et al*, 2005) containing 2067 genes specifically expressed in proliferating cells, to find genes specific to early stages of stomatal development. Beemster *et al.,* (2005) performed microarray analysis on leaf samples to identify the genes involved in regulating cell proliferation during organ growth. Their

sample tissue contained entire leaf tissue that includes epidermis (stomatal lineage cells, stomata, pavement cells and meristemoids), mesophyll and vascular tissue. Our stem tissue contained all the layers from the stem tip but the difference is that there were no advanced stage stomatal lineage cells. Comparing these two datasets I expected to find genes with a role in early stages of development. I found 22 genes common in these two datasets (Fig. 16) including *TMM* (*At1g80080*), a known stomatal regulator (Table 5).

Similarly, I used the stem dataset containing 260 differentially regulated genes and YODA dataset (Bergmann *et al*, 2004) containing 220 differentially regulated genes. Our aim was to find genes common to early stages of stomatal development. Bergmann *et al* (2004) observed that loss of functional *YODA* resulted in excess stomata on leaf and excessive *YODA* activity resulted in the formation of all pavement cells with the loss of stomatal cells. They capitalized on this condition to conduct a microarray gene expression profiling experiment designed to reveal genes important to all stages of stomatal development. Because their samples contained cells at all stages of stomatal development (meristemoids, GMCs, GCs) they also identified genes important in processes such as GC differentiation. In contrast, the sample contained only stomatal precursors from the stomatal lineage cells that helped us generate an expression profile that contains stomatal initiation and early development genes. There were 6 genes common in the stem and YODA datasets (Fig. 16). Interestingly, *At1g79700* was one of these common genes (Table 5) I found that its mutant has a

stomatal phenotype in leaves, reinforcing its common importance to stomatal development.

Table 5. Overlap of genes between stem, YODA and leaf cell proliferation microarray expression experiments

Stems vs Cell proliferation genes	**YODA vs Cell proliferation genes**	**Stems vs Yoda**
At5g49330	At5g54970	At5g20110
At5g20630	At5g18430	At2g29440
At5g04200	At5g12900	At1g79700
At5g03300	At5g01870	At1g76410
At4g36180	At4g21850	At1g17700
At4g33810	At4g03010	At1g10585
At4g29030	At4g01700	
At4g01270	At3g55500	
At3g56810	At3g23840	
At3g55450	At3g16670	
At3g26200	At3g16660	
At3g23810	At3g05730	
At3g11520	At3g02550	
At2g40550	At2g47240	
At2g30010	At2g39690	
At2g22170	At2g36690	
At2g16660	At2g36490	
At1g80080	At2g21140	
At1g74030	At2g20875	
At1g34245	At2g17880	
At1g29980	At2g16630	
At1g16390	At2g04570	
	At1g65400	
	At1g44760	
	At1g33811	
	At1g24260	
	At1g12845	
	At1g10060	
	At1g04110	

In cell proliferation genes dataset and YODA dataset there was a known stomatal regulator gene, which is *At2g20875* (Hara *et al*., 2007). This suggests that comparing these datasets was useful in focusing attention on candidate genes that have real roles in stomatal development.

DAVID bioinformatics analysis revealed absence of overrepresentation of any gene category

In order to statistically assess whether any functional category is overrepresented in the dataset, I used the DAVID web-based software tool. Overrepresentation of particular classes of genes might provide general insight into pathways or processes regulated by TMM signaling. The 260 genes obtained from Genespring analysis fell into 8 different functional clusters based on their annotation (Appendix E), and there is no gene functional group that is statistically over-represented.

Although no groups were significantly overrepresented, deeper analysis of group1 genes showed that there are three genes in this group that are highly related to each other with similarity score of 0.5-0.75, they are At4G26890 (MAPKKK16), At4G38830 (receptor like protein kinase RLK3) and At3G45440. (Receptor like protein kinase) that belongs to kinase family of proteins. We are more interested in genes involved in signal transduction, like kinases, because my aim to find new candidates involved in TMM signaling that regulates stomatal development. These genes will be the subject of future investigation.

Genes co-expressed with TMM

A set of genes that are expressed at the same point of time determine cell identity by specifying a unique molecular signature. For this reason co-expressed genes may be involved in the same or related cellular processes or functions. Using ATTED II (Obayashi *et al*., 2007) web based tool to search for genes that are co-expressed with *TOO MANY MOUTHS* (*At1g80080*) generated a list of co-regulated genes (Table 6). The top ten genes (Table 6) with correlation co-efficient in the range of 1-0.64 were selected for a network diagram (Fig. 17) showing their putative relationship based on degree of coregulation. One of these genes, *At1g34245*, is also differentially expressed in the microarray data. ATTED analysis shows that there is high correlation in the expression of *At1g34245* with TMM, shown by correlation coefficient of 0.78. My study also demonstrated that plants mutant for this gene have a stomatal phenotype. Another gene coregulated with *TMM* is *At5g60880* (Correlation co-efficient 0.64). This gene was found to differentially regulated in the gene list generated through the Bayesian statistical method. Plants mutant for this gene showed pairs of stomata indicating that it has a role in stomatal patterning.

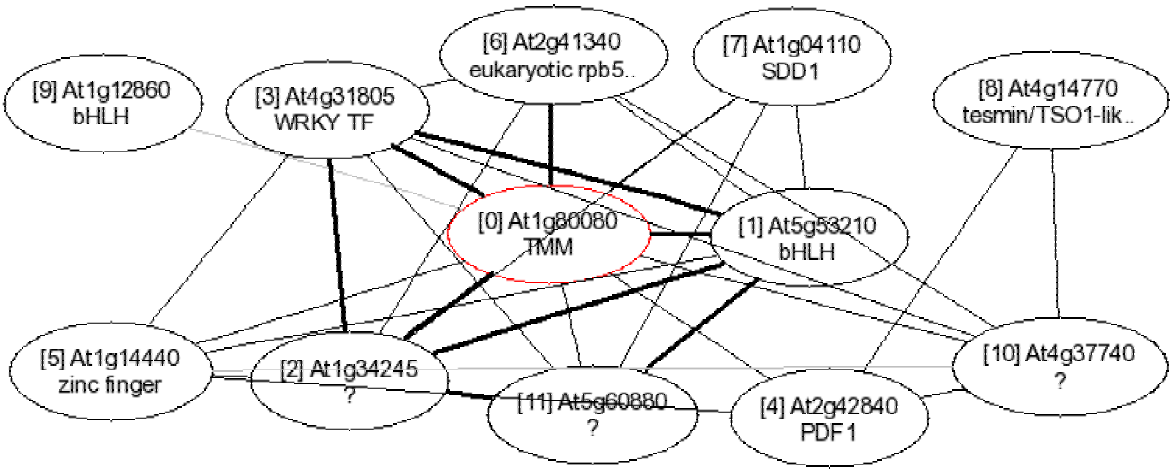


Figure 17. Co-expressed gene network centered on *At1g80080* (*TMM*)

ATTED II was used to find co-expressed genes in publicly available microarray datasets. ATTED II uses a Pearson correlation statistic to list genes based correlation with a query gene (*TMM* in this case). Alias names is presented under the gene locus name, or "?" indicates unknown function. Bold lines represent averaged rank of 1-5; normal lines 5-30; weak lines 30-50; no line 50 above.

Table 6. The top 10 genes co-expressed with *TMM* (*At1g80080*)

cor	locus	function
1.00	At1g80080	leucine-rich repeat family protein
0.79	At5g53210	basic helix-loop-helix (bHLH) family protein
0.78	At1g34245	expressed protein
0.76	At4g31805	WRKY family transcription factor
0.71	At2g42840	protodermal factor 1 (PDF1)
0.70	At1g14440	zinc finger homeobox family protein / ZF-HD homeobox family protein
0.65	At2g41340	eukaryotic rpb5 RNA polymerase subunit family protein
0.65	At1g04110	SDD1 (STOMATAL DENSITY AND DISTRIBUTION)
0.65	At4g14770	tesmin/TSO1-like CXC domain-containing protein
0.64	At1g12860	basic helix-loop-helix (bHLH) family protein / F-box family protein
0.64	At4g37740	expressed protein
0.64	At5g60880	expressed protein

To illustrate the expression patterns of TMM and At1g34245 (Fig. 18A) in various organs of Arabidopsis, the AtGenExpress visualization tool was used. This tool obtained data from many microarray experiments and provide expression patterns of the requested genes. ATTED II analysis showed that At1g34245 is co regulated with TMM and it was also found in my genelist so I selected this gene for further analysis with AtGenExpress. It showed that these two genes are co-expressed in all the organs of the plant in all the experiments. It showed that these two genes have same expression levels (close matching in graphs, Fig. 18A) in stem, leaf and floral organs of the Arabidopsis plant in all the experiments and these organs have stomata in their epidermis. Whereas in root and seeds they were shown to be co-regulated but not as closely as in the inflorescence and lateral shoot organs. Co-expression of TMM and At1g34245 in all the green organs of the plant and its role in regulation of stomatal patterning suggests that they could be involved in same signaling pathway.

At1g79700 was found in common genes among YODA dataset and the stem genelist and also it was found to differentially regulated in the stem genelist. Hence I chose this gene for further analysis. To illustrate the expression patterns of *TMM* and *At1g79700* (Fig. 18C) in various organs of Arabidopsis again AtGenExpress tool was used. The graph obtained showed that these two genes have reverse expression patterns in all the organs of the plant (Fig. 18C). This supports the observation that it is downregulated in *tmm* stems in my gene expression experiment.

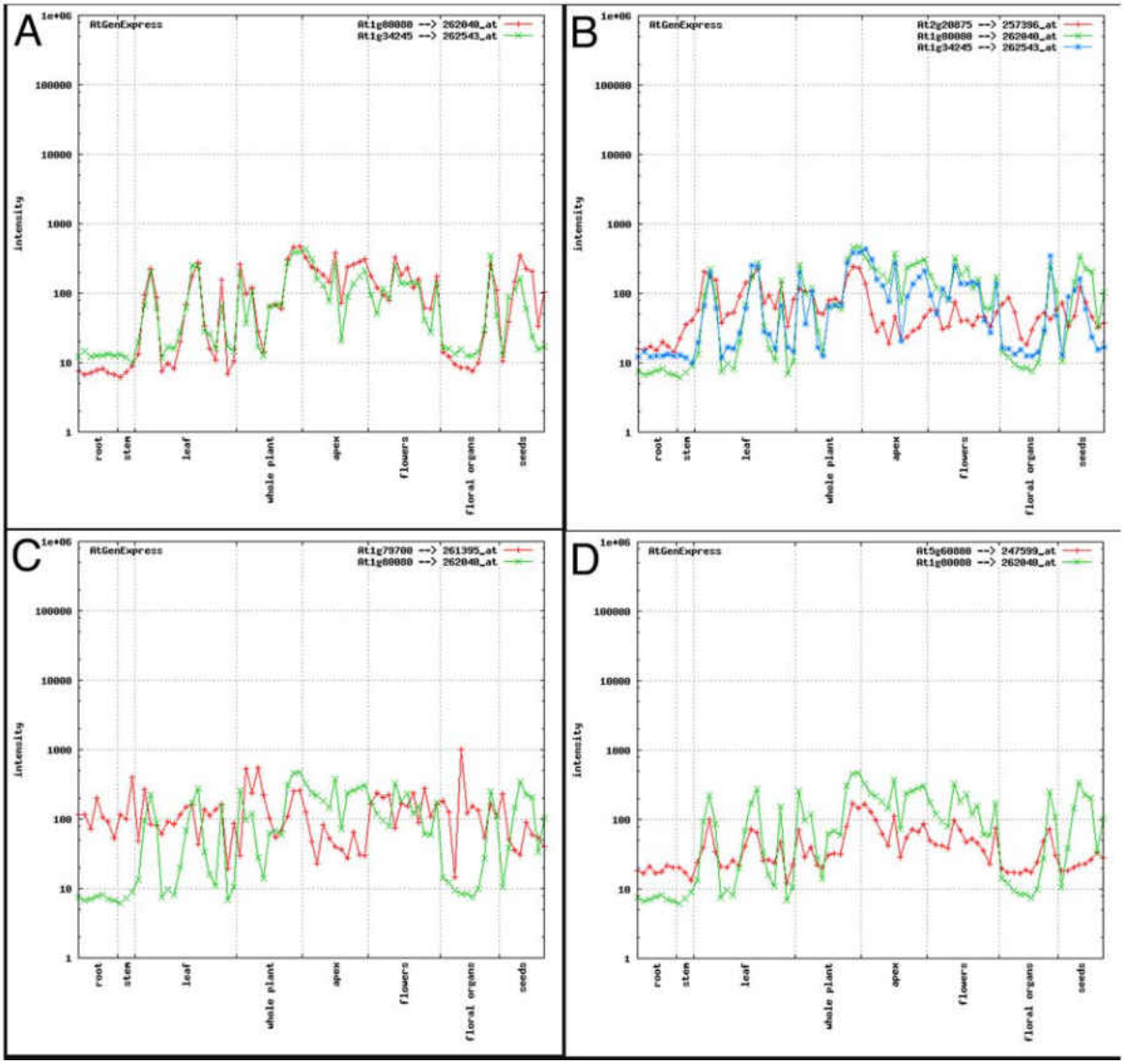


Figure 18. Global gene expression profiles in developmental microarray profiles (AtGenExpress)

(A) Expression of *At1g80080* (*TMM*) and *At1g34245* is strongly coregulated.
(B) Expression of *At1g80080*, *At1g34245* and *At2g20875* is strongly coregulated
(C) Expression of *At1g80080* and *At1g79700* is inversely correlated within some tissues.
(D) Expression of *At1g80080* and *At5g60880* is moderately coregulated.

Five new stomatal regulators were discovered

The ultimate purpose of this gene expression profiling experiment was to uncover new genes that function as stomatal regulators. To assess whether the differentially regulated genes had any biological significance, sixty five genes from the RACE list were selected for phenotypic analysis. The SALK insertion mutants of these candidate genes (Alonso *et al*., 2003) were planted and after 10 days their cotyledons observed under microscope for stomatal phenotypes. From a total of 65 genes screened (Appendix B), 5 showed defects in stomatal patterning. These were At1g34245 (unknown protein, SALK_047918 and SALK_102777), At1g79700 (similar to AP2 domain transcription factor, SALK_046920), At5g67480 (unknown protein, SALK_045370C), At3g20810 (Jumonji domain transcription factor, SAIL_811_H12), At2g05540 (putative glycine rich protein, SAIL_255_A01) and At5g60880 (unknown protein, SALK_086397). In addition, the gene *At2g20875*, which is closely related to *A1g34245*, and *At1g01060*, which is related to *At1g79700* were also ordered from the stock center to examine possible functional redundancy.

Phenotype analysis of two SALK insertion mutant lines of *At1g34245* (SALK_102777 and SALK_047918) showed pairs and triplets of stomata in mature cotyledons. Their young cotyledons showed (that there were more meristemoids compared to wild-type) that many cells resemble meristemoids indicating that more cells acquired meristemoid fate due to defect in this gene. Future studies are required to confirm this phenotype/defect.

Phenotype analysis of SALK_046920, a mutant line of *At1g79700* showed occasional pairs of stomata. Phenotype analysis of 5-7 day old young cotyledons of SALK_086397, a mutant line of At5g60880 showed pairs and triplets of stomata. Their meristemoids were undergoing frequent divisions and resembled a caterpillar. Future studies are required to confirm this phenotype/defect. Peripheral analysis of other mutants showed pairs and occasional triplets of stomata.

Two of these genes, *At1g34245* and *At1g79700* are described in detail below.

At1g34245 encodes a small peptide

The *At1g34245* gene is 4601 bp in length, with 3 exons and 2 introns (Fig. 19A). The full-length mRNA is 885 nucleotides and encodes an unknown protein of 120 amino acids (Fig. 19B-D) with a calculated molecular mass of 13.3 kDa and isoelectric point of 8.96. The putative protein encodes a signal peptide of 25 amino acids with cleavage site between 25th and 26th amino acids (Fig. 19B), and no other recognizable domains (SMART). TargetP (TargetP 1.1, Nakai and Kanehisa, 1991) predicts that it is a secretory protein localized outside cells. There are 8 cysteines found in the mature protein sequence (Fig. 19C-D), and the overall structure is reminiscent of other *Arabidopsis* proteins that are secreted peptide ligands for signaling pathways (Ryan *et al*., 2002).

Two closely related genes of *At1g34245* were found using a BLAST search, they are *At1g20875* and *At1g71868*. Mature protein sequence of these genes showed they are most similar to At1g34245 at the C-terminal end (Fig. 19D). This shows that the C-terminal end of this protein is conserved through evolution and probably has a role in its

function. Inspection of the sequence alignment shows that all three related genes have 8 cysteines in perfect alignment (Fig. 19D). *At1g71868* does not appear to encode a protein, so it could be a pseudogene. *At2g20875* encodes a protein of 11.4 kDa with and isoelectric point of 9.08. It was also predicted to be a small secretory protein and suggested as putative ligand of TMM by Hara *et al* (2007).

Both mutant lines of *At1g34245*, SALK_047918 (64 bp) and SALK_102777 (32 bp), are T-DNA insertions in the third exon (Fig. 19A). Since these insertions fall in the C-terminal region of the coded protein and this region is presumably important for its function, I predict that it might strongly affect the function of this protein although it is not possible to say that it could be a null mutation without further experimentation.

When mutant plants were examined, pairs of stomata and rarely triplets (Fig. 19E) were commonly observed in the mature epidermis of 10-day old cotyledons. In young cotyledons there were a large number of cells surrounding stomata that were dividing asymmetrically to form new meristemoids. This may be because more cells are acquiring the stomatal lineage fate, and stomata form in contact as a result of overabundant stomatal initiation. Both mutant alleles showed a similar phenotype.

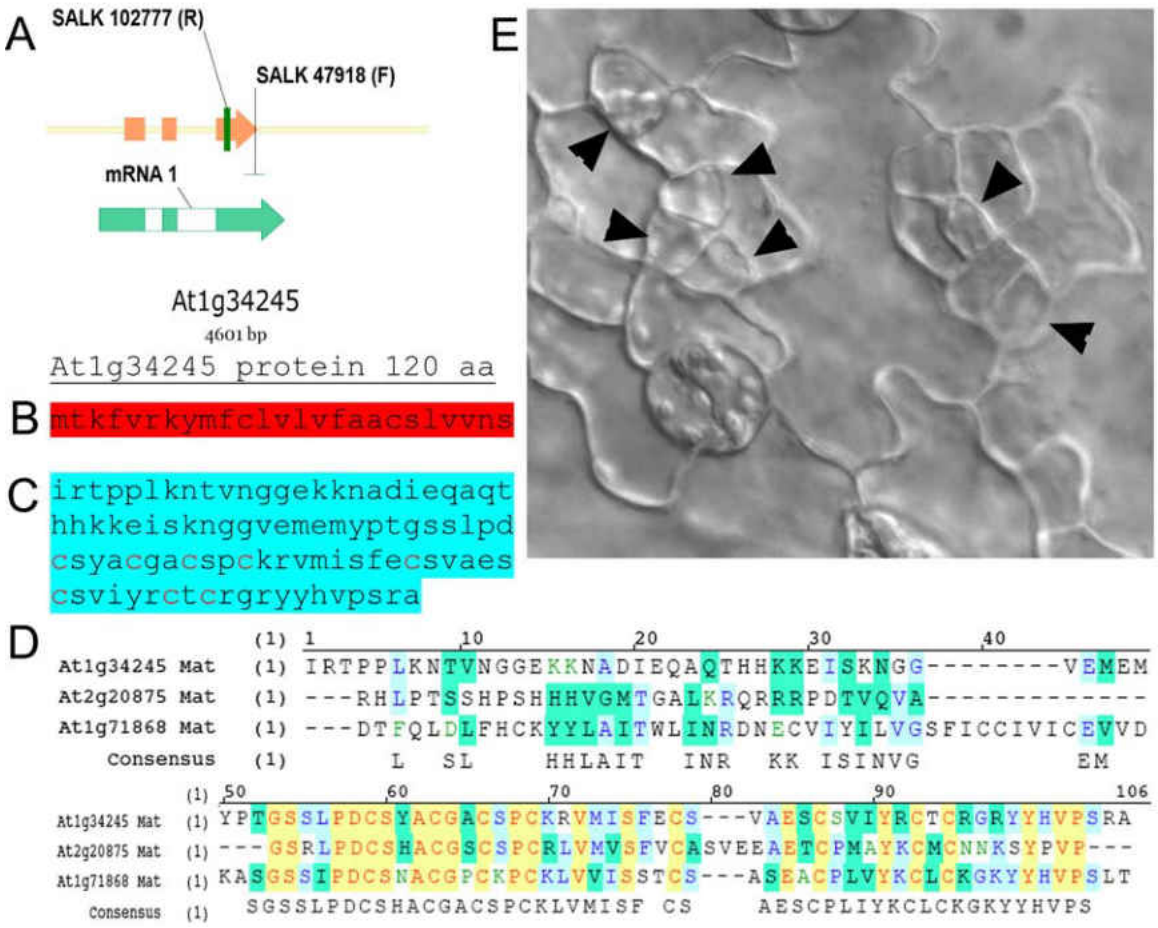


Figure 19. Features of *At1g34245* gene, protein and mutant phenotype

(A) *At1g34245* has 3 exons (orange bars) and 2 introns (light yellow bars). SALK 102777 and SALK 0447918 are both insertional mutations in the third exon. The mRNA shown by green arrow with arrowhead pointing the direction of transcription.
(B) The amino acid sequence of the protein has 120 amino acids with a signal peptide of 25 amino acids at the N-terminus (red highlight).
(C) The predicted mature small peptide amino acid sequence is highlighted in blue.
(D) Mature protein sequence alignment of related genes of *At1g34245* that are *At2g20875* and *At1g71868*. The sequence alignment showing more conservation among the sequences towards C-terminus where there are also cysteines present.
(E) Abaxial side of the young cotyledon of SALK 102777 insertional mutant showing clusters of meristemoids (arrowhead). Light microscope picture under 40X magnification.

At1g79700 encodes an AP2-domain transcription factor

At1g79700 encodes 8 exons and 7 introns (Fig. 20A). Genbank contains expressed sequence tag (EST) data for 2 mRNAs of varying lengths, suggesting that this mRNA is alternatively spliced. The first mRNA is 912 nucleotides containing the open reading frame that encodes a protein of 303 amino acids with a calculated molecular mass of 34.2 kDa and isoelectric point of 7.06. The second mRNA containing 942 nucleotides with an open reading frame that encodes a protein of 313 amino acids with a calculated mass of 35.4 kDa and isoelectric point of 7.5.

None of the predicted proteins that could be encoded by *At1g79700* seems to have a signal peptide. This protein may be targeted to the nucleus based on the PSORT algorithm likelihood score of 0.6 for nuclear localization. This gene is likely to encode a transcription factor because it contains two plant-specific AP2 domains (Fig. 20B) that are known to bind DNA (SMART) and also its potential nuclear localization hints at this. The extra 10 amino acids encoded by mRNA 2 lie in the second AP2 domain of the putative protein (Fig. 20B). Because AP2 domains are functionally important in transcription factors, I predict these two proteins with difference in one AP2 domain could have different interactors and may have different biological functions.

The T-DNA mutation SALK_046920 (insertion in exon 2) was selected for phenotypic characterization. Because this insertion would interrupt translation of the second AP2 domain that is important to the function of transcription factors, I predict that this mutation would most probably result in a null mutation. Again, it is hard to

predict correctly if the mutation is null until all the functional regions of proteins are known.

The abaxial side of 10-day-old mature cotyledons from this mutant was observed under the microscope to have a few pairs of stomata (Fig. 20C). Young cotyledons were not observed so I do not know about early development. Future studies are required.

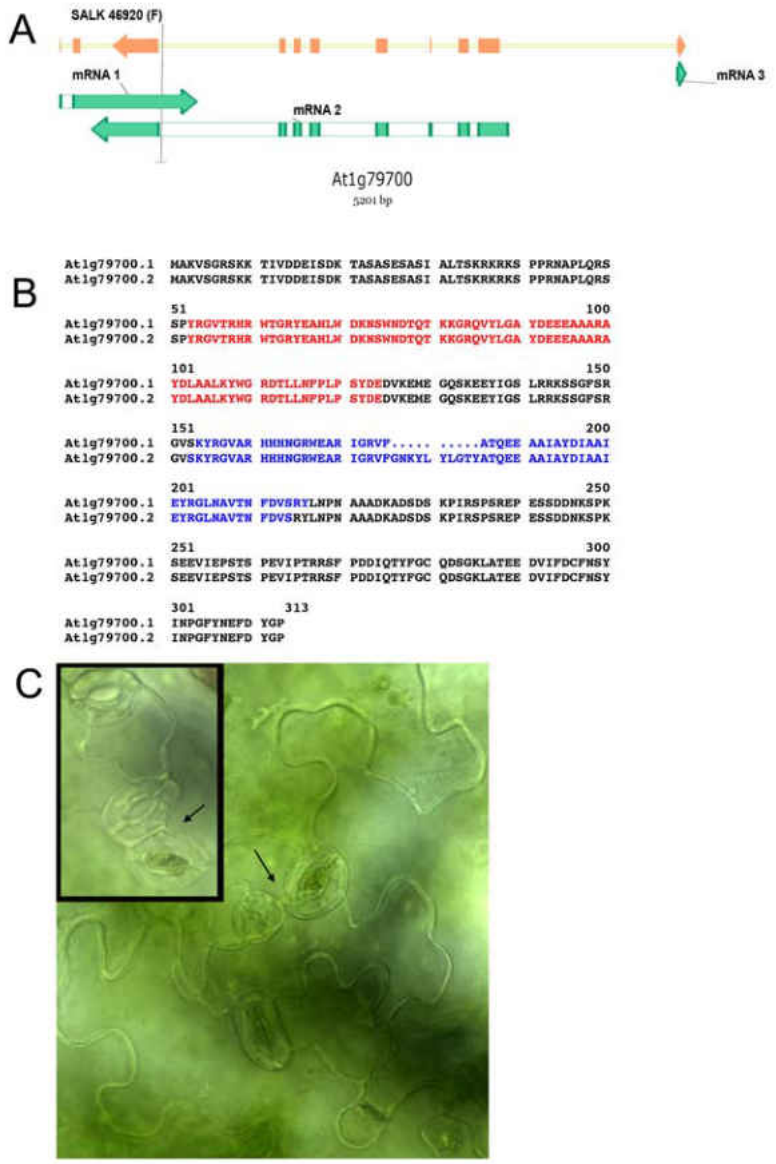


Figure 20. Features of *At1g79700* gene, protein and mutant phenotype

(A) *At1g79700* contains 8 exons (orange bars) and 7 introns (light yellow bars). SALK 46920 is the insertion mutation at the 3' end of the last intron near the junction of the 8th exon. The mRNA shown by green arrow with arrowhead pointing the direction of transcription.
(B) The first mRNA encodes a protein of 303 amino acids, while the second mRNA encodes a protein of 313 amino acids. The first AP2 domain is indicated by letters in red and the second AP2 domain is indicated by letters in blue.
(C) The abaxial side of the mature cotyledon of insertional mutant SALK 46920 showing two stomata in contact (arrow) probably formed by separate meristemoids. Inset showing the pair of stomata (arrow). Light microscope picture under 40X magnification.

Discussion

In order to identify novel genes involved in controlling early stages of stomatal development I used gene expression profiling. To generate this expression profile I compared wild-type and *tmm* stem tissues undergoing stomatal development. Stem tissue chosen contained meristemoids that are dividing and would differentiate later to stomata in wild-type plants, but in *tmm* mutant plants these cells differentiate into pavement cells instead. When these tissues were compared, I found 260 differentially expressed genes using a t-test statistical analysis and 352 genes using a Bayesian statistical analysis. This is a fairly small percentage of the total number of genes in Arabidopsis (1.46% of the total 24000 genes on chip), which probably reflects the extremely targeted approach I employed. Categorization of these differentially expressed genes based on the gene ontology category of “biological processes” showed that 10% of 260 differentially expressed genes are involved in growth, development and response to stimulus. The remaining fraction of genes (90%) includes those involved in cellular and physiological processes.

Because sample tissue contains actively dividing cells I expected to find more genes associated with cell division and metabolism to differentially regulate in these tissues. In contrast, my analysis using the DAVID web-based tool showed that neither of these categories was significantly overrepresented (Appendix E). In fact, no particular functional category was overrepresented in the gene set. This indicates that there is no difference in the major cellular processes in the samples, and that the few

genes that are differentially regulated represent a broad cross-section of functional classes.

Another approach to identifying genes significant to regulating cell division or meristemoid behavior was to compare genes identified in my experiment to those previously identified using other approaches to gene expression profiling. For this I compared the stem dataset with leaf cell proliferation dataset (Beemster *et al*, 2005) and YODA (Bergmann *et al,* 2004) experimental dataset. As a result, I found that there are 22 genes common between my stem dataset and cell proliferation dataset. Interestingly, *At1g34245* was found in the common genes between these datasets. My phenotypic analysis revealed that this gene plays a role in early stomatal development. There were 6 genes common in YODA and the stem datasets, including *At1g79700*. This gene was also found to have stomatal phenotype in the mutant analysis. This approach helped us to take an informed approach to prioritizing the genes to focus my attention on candidate genes.

To find if there were any known stomatal development and patterning genes appearing in my gene list I looked for some of the candidates as described further. For example some genes are known to have a role in the initiation of meristemoids such as *TMM, YODA* (Bergmann *et al*., 2004) and *SPEECHLESS* (MacAlister *et al*., 2007). The gene lists obtained from traditional t-test and Bayesian analysis showed known stomatal regulator genes are expressed *TOO MANY MOUTHS, YODA, SPEECHLESS, ERECTA* and *ERECTA LIKE 1* (Shpak *et al*., 2005). These results prove that the sample tissue had cells from early stages of stomatal development. In addition it shows

that the *TMM* gene is defective but the transcript is expressed. *MUTE* (Pillitteri *et al*., 2007) plays a role in differentiation of meristemoids into GMCs. Stomatal lineage cells from advanced stages of differentiation, such as the guard mother cell or guard cell, express the *KAT1* gene. Absence of *MUTE* and *KAT1* from the samples indicates that none of the samples have guard mother cells or guard cells.

My research illustrates that *TMM* is upregulated in *tmm* mutant samples. It was showed by northern blot analysis that TMM is expressed in wild-type, *tmm-1* and *tmm-2* rosette leaf tissue, although there is presumably no functional protein produced (Nadeau and Sack, 2002). One possibility for increase in *TMM* transcripts in the *tmm* mutant sample is that it is self-regulatory and hence when it is defective it cannot negatively regulate its own expression. The second interesting possibility is that there are more meristemoids (as shown in Chapter 1) in *tmm* stem tissue hence there is higher level of *tmm* expression.

The aim of these experiments was to identify new genes involved in stomatal patterning. To determine if any of these genes had such a role, candidates were chosen from the gene list for phenotypic analysis of mutants. One insertional mutant line for each of 65 different genes was grown and examined, and six showed defects in some aspect of stomatal development (*At1g34245, At1g79700, At5g67480, At3g20810, At2g05540 and At5g60880*). *At1g34245*, *At5g60880* and *At5g67480* encode unknown proteins. *At1g79700* encodes a protein similar to AP2 domain transcription factors, while *At3g20810* encodes a Jumonji domain transcription factor. *At2g05540* encodes a putative glycine rich protein of unknown function. In conclusion, I was able to confirm

the validity of this approach to gene discovery by identifying six new stomatal regulators. Below I discuss my analysis of *At1g34245* and *At1g79700* because they seem to have a role in early stages of stomatal development.

At1g3425 is a putative small ligand

The predicted protein encoded by *At1g34245* shows features consistent with those of secreted peptide ligands involved in signal transduction. One of the common features of extracellular ligands is their small size, as well as sequence conservation with other putative peptides in the C-terminal region of the protein. Many extracellular small peptide ligands studied in plants (Table 7) are composed of two to few hundred amino acids and are secreted outside the cell so that they can interact with the receptor and aid in relaying the signal through a membrane receptor. For example, the 18 amino acid long tomato Systemin was the first polypeptide hormone discovered in plants. It was shown to be involved in synthesis and accumulation of proteinase inhibitors as a defense response to insect attacks or wounds (Pearce *et al*., 1991). Tobacco systemin is also 18 amino acids long (Pearce *et al*., 2001). Matsubayashi & Sakagami (Matsubayashi and Sakagami, 1996) discovered phytosulfokines (PSK-α and PSK-β), the smallest known plant peptide hormones (3-5 amino acids) in asparagus suspension cultured cells. They are involved in regulating cell proliferation. Phytosulfokines were also discovered in other plants such as *Oryza sativa* (Matsubayashi *et al*., 1997), *Zinnia* (Matsubayashi *et al*., 1999) and many other plants. Rice PSK-α precursor is 89 amino acids (Yang *et al*., 1999b) that includes N-terminal signal sequence of 22 amino acids. Its C-terminus has the peptide sequence that contains the 5 amino acid peptide. PSK-α

precursor has the flanking aspartic acid residues, that were suggested as the sites for endoproteolytic processing in animals (Harris, 1989). It was suggested that it is post-translationally processed to form the functional small peptide. The C-terminus of At1g34245 is ~ 50 amino acids long as predicted by the conserved region when it is aligned with its related genes (Fig. 19D). It is probable that he 50 amino acid C-terminus of At1g34245 could be a precursor. It may be processed by proteases to generate a smaller processed functional small peptide.

Because At1g34245 appears to be a secreted peptide involved in stomatal patterning, it could be a target for the STOMATAL DENSITY and DISTRIBUTION-1 (SDD1) subtilisin protease. SDD1 protease was suggested to be involved in processing the extracellular signal that interacts with TMM receptor complex (Berger and Altmann, 2000; von Groll *et al*., 2002). In order to find if At1g34245 might undergo proteolytic processing like other hormone precursors I analyzed the presence of protease recognition sites with predictive software. The study revealed no protease recognition sites indicating that it may not be processed by SDD1. Peptide hormone precursors in animals are known to be processed by the subtilisin proteinases at a dibasic pair of amino acids, usually a pair of arginines (Harris, 1989). Homologs of subtilisins (Berger and Altmann, 2000) were found in plants, but there is no evidence showing these are involved in precursor processing. Not much is known about the precursor hormone processing in plants so I cannot conclude that At1g34245 would not undergo any processing.

ATTED II analysis showed that *At1g34245* is co-regulated with *TMM* suggesting that this putative ligand is co-expressed with a potential receptor, which would be expected if they act together in a signaling pathway. AtGenExpress analysis showed that *TMM* and *At1g34245* are strongly co-regulated in all green organs of the plant (Fig. 17).

Phenotypic analysis of the *At1g34245* mutants (SALK_047918 & SALK_102777) showed that in young cotyledons there are more meristemoids, indicating that more cells are entering into stomatal lineage (Fig. 19E). Mature cotyledons showed some pairs of stomata, a violation of the normal spacing pattern, but there were no large clusters of stomata. This shows that loss of functional small putative ligand does not disturb the one-cell spacing rule as frequently as *tmm* but some pairs of stomata may form due to overabundance of cells acquiring the stomatal lineage. Both the molecular identity and its mutant phenotype similar to *tmm* suggest that this protein might act as a positional signal, perhaps detected by the TMM receptor complex.

I predict that At1g34245 may interact with TMM by acquiring the globular structure of cysteine knot proteins through formation of several disulphide bonds. This structure may aid in the interaction with a specific receptor. Observation of At1g34245 protein sequence shows 10 cysteines (Fig. 19B-C). Such cysteine rich sequences are characteristic of other small peptide ligands such as SCR/SP11 (S-locus cysteine rich proteins/S-locus protein 11) that interact with protein binding sequences of the receptors (Schopfer *et al*., 1999). Eight conserved cysteines have been found in small peptides of SCR/Sp11 family members that were suggested to play a role in acquiring specific

conformation required for their activity. Plant cyclotides such as Kalata B1 (Saether *et al*., 1995) and Cycloviolacin O1 (Craik *et al*., 1999) are small peptides that were suggested to be involved in defense response. These plant cyclotides were studied in detail for their conserved cysteine residues that would give similar overall conformation structure (cysteine knots) by formation of several disulfide bonds between their cysteines. Further analysis is required to conclude if At1g34245 small peptide ligand forms any cysteine knots and thus interacts with TMM.

Based on these biochemical features and its role in stomatal patterning, I suggest that At1g34245 could be putative small peptide ligand of TMM.

Table 7. Small peptide ligands and their properties. Size of protein given in number of amino acids

Small Polypeptide name	**Size**	**Function**
Tomato Systemin	18	Defense response
Tobacco Systemins I & II	18	Defense response
RALF	49	Regulates stress, growth and development
DEVIL	51	Development of leaves, siliques, etc.,
CLAVATA3	79	Regulation of meristem growth
Phytosulfokines	3-5	Cellular de-differentiation and proliferation

Similarities in At1g34245 and At2g20875 suggest these are putative TMM ligands

At1g34245 has a closely related gene in Arabidopsis (*At2g20875*) with a recently discovered function that helped to understand the potential function of these genes. We identified *At2g20875* and *At1g71868* in a BLAST search using *At1g34245*. *At1g71868*

is not known to encode an mRNA based on EST data, so it is probably a pseudogene. *At2g20875* has been recently discovered by Hara *et al*. (2007) to be involved in stomatal patterning. Hara *et al* (2007) took a high-throughput approach to examine the function of many putative small peptides in Arabidopsis by selecting 153 genes from the genome that code for proteins of 150 or fewer amino acids that are predicted to be secreted by the PSORT program. The selected genes were overexpressed in transgenic plants using a constitutive Cauliflower Mosaic Virus (CaMV) promoter and searched for phenotypes. They discovered *EPF1* (*EPIDERMAL PATTERNING FACTOR 1* = *At2g20875*) had very few stomata on cotyledons when overexpressed in wild-type plants. They also demonstrated that overexpression of EPF1 protein in *tmm* mutant plants did not reduce the number of stomata, indicating that *tmm* is epistatic to the effect of EPF1 overexpression. This is consistent with what would be expected if EPF1 were a ligand for the TMM receptor complex. They showed by *insitu* RNA hybridization that EPF1 is expressed in stomatal precursors, the probable source of stomatal patterning signals. My observation of the mutant (SALK_137549) of *At2g20875* showed some clusters of as many as 2-4 stomata but they are not as severe as *tmm*. Similarly, mutants of *At1g34245* also showed pairs and rarely triplets of stomata, but phenotype was not as severe as *tmm*.

The sequences of these two genes are well conserved at the C-terminal end and are much less conserved at the N-terminal end (Fig. 19D). In the C-terminal end all cysteines are conserved. Conservation of this region suggests that it is significant for its function. Some small peptide families of small ligands like CLAVATA3, RALF and

PHYTOSULFOKINE and emphasized that the C-terminal sequences are conserved in small peptides and required for their activity (Ryan *et al*., 2002). The C-terminal sequences are conserved in the RALF (RAPID ALKALINIZATION FACTOR) small peptides family and N-terminal sequences are not very conserved (Pearce *et al*., 2001). This suggests that C-terminal sequence in At1g34245 is important perhaps for interaction with TMM or another receptor.

AtGenExpress analysis of expression of TMM, *At1g34245* and *At2g20875* in multiple experiments showed that they coregulated in green organs of plant (Fig. 18B). This suggested that they are co-expressed with *TMM*. The similarities between *At1g34245* and *At2g20875*, such as small size, presence of a signal peptide, conserved cysteines and the similar phenotype of the mutants with subtle difference, suggest that both could perform a similar function in plants. *At1g34245* and *At2g20875* are from same organism, have sequence and phenotypic similarities, these features suggest that they could be paralogs.

At1g79700 is predicted to be transcription factor

At1g79700 is another new gene identified as having a role in stomatal patterning by this approach. At1g79700 is predicted to be a transcription factor because it contains two characteristic AP2 DNA-binding domains. According to my phenotypic analysis of its mutant there is no observable stomatal patterning defect apart from few pairs. It is downregulated in the *tmm* sample. AtGenExpress shows that these two genes have an inverse expression pattern in multiple developmental experiments (Fig.

18C), in addition to ours. That is, if TMM is downregulated in the leaf, stem and floral organs, then *At1g79700* is generally upregulated.

At1g79700 was shown to be member of AP2/ERF family (APETALA-2 and ethylene-responsive element binding proteins) of transcription factors (TF). This family is one of the largest known TF families and is unique to plants. To understand the role of *At1g79700* in plant development I searched for similar regulators that have been studied previously. Two well-studied genes encoding AP2-domain transcription factors are *AINTEGUMENTA* (*ANT*) (Elliott *et al*., 1996) and *APETALA-2 (AP2)* (Jofuku *et al*., 1994). APETALA-2 transcription factors were discovered to play roles various steps in flower and seed development. ANT contains two AP2 domains that are >50% identical to the two AP2 domains of APETALA-2 protein. ANT is expressed in primordia of stem procambia, cotyledon and leaves but not in roots. ANT is believed to control the indeterminacy of cells that form lateral shoot organs.

Nole-Wilson *et al* showed relationship among 15 Arabidopsis genes containing two AP2 domains using a Neighbor Joining phylogenetic tree (Nole-Wilson *et al*., 2005). This tree has 3 subgroups; they are the AP2-like group, AINTEGUMNETA-like group and a third group composed of 4 genes. *At1g79700* falls in this third group and is related to other AP2 domain containing genes due to similarity only in the AP2 domain sequences. Based on this information and my own data (Fig. 20C), I predict that At1g79700 is a putative transcription factor that is involved in stomatal development.

At1g79700 transcripts are more abundant in wild-type stems. The stomatal precursors in wild-type stems undergo asymmetric divisions before differentiating into

stomata and the presence of At1g79700 in these cells may suggest that, like other ANT-like transcription factors, it could be regulating cell division competency. It is possible that more asymmetric divisions occur in wild-type meristemoids that express At1g79700. There is an evidence that ectopic overexpression of *ANT* using 35S::ANT results in an increase in the number of cells, which causes enhancement in the size of lateral shoot organs (Mizukami and Fischer, 2000). They also showed that loss of functional ANT reduces organ size by inhibiting cell division causing a reduction in cell number that lead ultimately to smaller leaves and floral organs. Based on these similarities I propose that *At1g79700* has a function similar to *ANT*, but limited to stomatal stem cells. Expression of *At1g79700* in wild-type meristemoids could reflect a role for At1g79700 in regulating the capacity of meristemoids to divide asymmetrically. Normally, the asymmetric divisions in meristemoids are required for separating adjacent stomata by atleast one-cell according to the one-cell spacing rule. I hypothesize that the mutant cotyledons of *At1g79700* showed pairs because loss of functional At1g79700 protein reduced meristemoid capacity to undergo asymmetric divisions. Thus all meristemoids could not obtain their local spacing and hence it resulted in stomatal pairs.

Future study

Characterization of new genes

The new stomatal regulators discovered by microarray analysis showed that they are differentially regulated by *TMM*. To find out if these genes interact with *TMM* to

relay the signal and to study their role in stomatal patterning additional investigation will be required.

In general, in a signaling pathway the ligands interact with receptors and activate them to pass the cue to the secondary messengers in the cytoplasm. So the ligand is considered as the upstream element in a signaling pathway. The activated secondary messengers then activate the transcription factors that translocate to the nucleus to activate the expression of specific genes.

To find if At1g34245 is upstream to TMM in the signaling pathway overexpress the gene in *tmm* plants and observe the phenotype. If there is *tmm* phenotype then it means that the gene may be genetically upstream in the signaling pathway. On the other hand if the phenotype is same as overexpression of the gene then it indicates that the gene is downstream of *tmm* or in another pathway altogether.

Create a double mutant of *At1g34245* and *tmm,* if the phenotype of double mutant is same as phenotype of *tmm* this would be consistent with At1g34245 being upstream of TMM. Otherwise, if the double mutant phenotype is additive (more clusters than in *tmm* and the gene mutant phenotype) then it shows that At1g34245 and TMM function independently.

To find which region in *At1g34245* is required for it to be functional, serial deletion experiments could be performed. To determine if cysteines in At1g34245 have a role in acquiring globular structure that allow interaction with receptor protein, amino acid substitution experiments could be performed.

To find if At1g79700 is downstream or downstream to TMM, a double mutant of *At1g79700* and TMM could be made. If the phenotype of the double mutant is the same as *At1g79700* mutant then it shows that At1g79700 is downstream of TMM.

APPENDIX A:
GENELIST (260 GENES) OBTAINED FROM GENESPRING ANALYSIS

Appendix A. Genelist (260 genes) obtained from Genespring analysis

Genbank	P-value	Description
At3g45440	0.05	receptor like protein kinase
At5g02430	0.05	putative protein rab11 binding protein
At1g15740	0.05	unknown protein
At3g29970	0.05	unknown protein
At3g30320	0.05	hypothetical protein
At5g05780	0.05	26S proteasome regulatory subunit S12 (MOV34 protein)
At2g14250	0.05	unknown protein
At3g49620	0.05	putative protein SRG1 protein
At1g29980	0.05	unknown protein
At2g30510	0.05	unknown protein
At4g32050	0.05	putative protein norbin
At5g08170	0.05	putative protein hypothetical bacterial proteins
At1g51680	0.05	4-coumarate:CoA ligase 1 identical to 4-coumarate:CoA ligase
At5g49330	0.05	putative protein contains similarity to myb-related transcription factor
At4g26890	0.05	putative NPK1-related protein kinase
At4g33870	0.05	putative peroxidase peroxidase ATP12a
At3g13750	0.05	galactosidase
At2g39910	0.05	hypothetical protein
At3g56360	0.05	putative protein
At1g19540	0.05	2-hydroxyisoflavone reductase
At1g64260	0.05	hypothetical protein similar to hypothetical protein
At3g49150	0.05	putative protein various predicted proteins
At3g63270	0.05	putative protein
At5g60190	0.049	putative protein
At1g02450	0.049	unknown protein
At1g28480	0.049	glutaredoxin, putative similar to glutaredoxin
At3g59670	0.049	putative protein hypothetical protein
At1g40550	0.048	hypothetical protein
At3g05770	0.048	unknown protein
At5g23960	0.048	(+)-delta-cadinene synthase (d-cadinene synthase)
At3g15350	0.048	unknown protein _
At1g18940	0.048	hypothetical protein
At3g22231	0.048	Expressed protein
At2g42890	0.048	putative RNA-binding protein
At1g34500	0.048	hypothetical protein
At3g02020	0.048	putative aspartate kinase similar to aspartate kinase
At5g57910	0.047	putative protein similar to unknown protein
At1g02720	0.047	hypothetical protein similar to putative glycosyl transferase
At2g40890	0.046	putative cytochrome P450

Genbank	P-value	Description
At1g79700	0.046	unknown protein similar to putative AP2 domain transcription factor
At3g55450	0.046	serine/threonine-specific protein kinase -like NAK
At3g18220	0.046	diacylglycerol pyrophosphate phosphatase
At5g64330	0.046	non-phototropic hypocotyl 3
At1g60710	0.046	auxin-induced protein, putative similar to auxin-induced atb2
At3g24150	0.046	hypothetical protein
At2g34490	0.046	putative cytochrome P450
At1g50620	0.046	hypothetical protein
At1g18690	0.046	alpha galactosyltransferase
At5g21170	0.046	AKIN beta1
At5g66360	0.046	dimethyladenosine transferase-like protein
At1g02390	0.046	unknown protein
At3g06980	0.046	putative DEAD/DEAH box helicase
At1g22770	0.046	putative gigantea protein
At5g01200	0.046	putative protein Myb-related transcriptional activator mybSt1
At4g31150	0.046	putative protein predicted protein
At4g37550	0.046	formamidase - like protein formamidase
At4g15900	0.045	PRL1 protein
At5g57350	0.045	plasma membrane ATPase 3 (proton pump)
At3g63180	0.045	hypothetical protein
At1g80830	0.045	metal ion transporter
At1g12780	0.045	uridine diphosphate glucose epimerase
At5g55600	0.045	unknown protein
At4g16515	0.044	Expressed protein
At1g80440	0.044	unknown protein contains two Kelch motifs
At2g43310	0.044	hypothetical protein
At3g15770	0.044	hypothetical protein
At1g12730	0.044	hypothetical protein
At1g59870	0.044	ABC transporter
At1g16390	0.044	putative transport protein may be a member of sugar transporter family
At4g24340	0.044	putative protein storage protein - Populus deltoides
At2g16660	0.044	nodulin-like protein
At1g17700	0.044	hypothetical protein
At1g09700	0.043	hypothetical protein
At4g38840	0.043	auxin-induced protein
At3g01450	0.043	unknown protein
At3g23810	0.043	S-adenosyl-L-homocysteinas
At5g52040	0.042	arginine/serine-rich splicing factor RSP41 homolog
At5g20110	0.042	dynein light chain - like protein dynein light chain LC6
At5g44990	0.042	putative protein strong
At2g47310	0.042	putative FCA-related protein

Genbank	P-value	Description
At1g15670	0.042	unknown protein
At1g63830	0.042	unknown protein
At5g20630	0.042	germin-like protein
At3g08510	0.042	phosphoinositide specific phospholipase (AtPLC2)
At5g63580	0.042	flavonol synthase
At5g54060	0.042	flavonol 3-O-glucosyltransferase
At1g77760	0.042	nitrate reductase 1 (NR1)
At1g45170	0.042	hypothetical protein contains similarity to vacuolating cytotoxin (vacA)
At4g12690	0.042	putative protein
At1g11080	0.042	Serine carboxypeptidase isolog
At1g43160	0.042	AP2 domain containing protein
At2g22560	0.042	hypothetical protein
At5g66140	0.042	20S proteasome subunit PAD2
At2g15680	0.042	putative calmodulin-like protein
At1g54700	0.041	hypothetical protein
At5g67320	0.041	putative protein
At1g11360	0.041	unknown protein
At1g02070	0.04	hypothetical protein
At2g29370	0.04	putative tropinone reductase
At5g42100	0.04	beta-1,3-glucanase-like protein
At2g13360	0.04	alanine-glyoxylate aminotransferase
At1g21680	0.04	unknown protein
At5g44620	0.04	flavonoid 3,5-hydroxylase-like; cytochrome P450
At2g25250	0.039	unknown protein
At2g44500	0.039	similar to axi 1 protein from Nicotiana tabacum
At2g17580	0.039	putative poly(A) polymerase
At3g01310	0.039	unknown protein
At5g03840	0.039	Terminal flower1 (TFL1)
At3g13590	0.039	hypothetical protein
At1g06990	0.039	hypothetical protein
At5g11790	0.039	putative protein SF21 protein
At1g13450	0.039	DNA binding protein GT-1
At4g24940	0.039	ubiquitin activating enzyme
At5g16530	0.039	putative protein contains similarity to auxin transport protein
At4g15530	0.039	pyruvate,orthophosphate dikinase
At1g74030	0.039	putative enolase
At1g79440	0.039	succinic semialdehyde dehydrogenase
At1g10585	0.039	Expressed protein
At5g12340	0.039	putative protein
At3g23280	0.039	unknown protein
At3g16370	0.039	putative APG protein

Genbank	P-value	Description
At1g71030	0.039	putative transcription factor
At1g76410	0.039	putative RING zinc finger protein
At4g34650	0.039	predicted protein
At2g31085	0.038	CLE4 CLAVATA3/ESR-Related 4 (CLE4)
At2g29440	0.038	putative glutathione S-transferase
At3g43630	0.038	nodulin -like protein nodulin-21
At2g28085	0.038	Expressed protein
At5g24560	0.038	phloem-specific lectin-like protein
At3g48650	0.038	hypothetical protein
At4g31300	0.038	multicatalytic endopeptidase complex, proteasome precursor
At1g28710	0.038	unknown protein
At3g48530	0.038	putative protein probable transcription regulator protein
At2g31970	0.038	putative RAD50 DNA repair protein
At4g19840	0.038	lectin like protein lectin phloem protein PP2
At1g23190	0.038	putative phosphoglucomutase
At1g05310	0.038	putative pectin methylesterase
At2g02680	0.038	hypothetical protein
At3g47340	0.037	glutamine-dependent asparagine synthetase
At3g22540	0.037	hypothetical protein
At2g21090	0.037	unknown protein
At3g15880	0.037	putative WD-repeat protein
At3g56810	0.037	hypothetical protein
At1g33790	0.037	myrosinase binding protein
At5g04200	0.037	latex-abundant protein
At5g03300	0.037	putative protein
At2g31810	0.037	putative acetolactate synthase
At5g66650	0.037	putative protein
At2g03980	0.036	putative GDSL-motif lipase/hydrolase
At4g16630	0.036	RNA helicase like protein
At2g18690	0.036	unknown protein
At5g23380	0.036	putative protein
At1g80920	0.036	J8-like protein similar to DnaJ homologue J8
At2g30010	0.035	hypothetical protein
At3g60320	0.035	bZIP protein
At1g09140	0.035	putative SF2/ASF splicing modulator, Srp30
At2g22170	0.035	unknown protein
At5g05540	0.035	putative protein
At4g33810	0.035	beta-xylan endohydrolase -like protein
At4g00300	0.035	contains weak similarity to S. cerevisiae BOB1 protein
At5g18110	0.035	eukaryotic cap-binding protein
At4g39800	0.035	myo-inositol-1-phosphate synthase
At3g29590	0.035	Anthocyanin 5-aromatic acyltransferase

Genbank	P-value	Description
At5g43390	0.035	putative protein
At5g06460	0.035	ubiquitin activating enzyme 2 (UBA2)
At4g05470	0.035	F-box protein family
At1g73070	0.035	disease resistance protein
At2g37650	0.035	putative SCARECROW gene regulator
At2g24850	0.035	putative tyrosine aminotransferase
At3g57840	0.035	hypothetical protein
At1g54030	0.035	myrosinase-associated protein
At1g52890	0.035	NAM-like protein -NAM (no apical meristem)
At4g39090	0.035	cysteine proteinase RD19A identical to thiol protease
At3g47680	0.035	hypothetical protein
At1g76470	0.035	putative cinnamoyl-CoA reductase
At5g65660	0.035	unknown protein
At1g50410	0.035	DNA-binding protein
At4g38490	0.034	Expressed protein
At1g56300	0.034	DnaJ protein
At1g14910	0.033	unknown protein
At1g52150	0.033	HD-Zip protein
At4g30190	0.033	H+-transporting ATPase type 2
At5g54120	0.033	unknown protein
At1g49310	0.033	hypothetical protein
At5g22920	0.033	PGPD14 protein
At5g02810	0.033	putative protein
At2g25500	0.033	hypothetical protein
At1g05910	0.033	tat-binding protein
At1g35510	0.033	growth regulator
At1g01820	0.033	unknown protein
At4g32340	0.033	putative protein predicted proteins
At2g33830	0.033	putative auxin-regulated protein
At1g29670	0.033	lipase/hydrolase
At5g47780	0.033	putative protein
At5g61910	0.033	putative protein B2 protein
At1g26580	0.033	hypothetical protein
At2g30420	0.033	myb-like protein isolog
At2g20420	0.033	succinyl-CoA ligase beta subunit
At2g40550	0.033	hypothetical protein
At4g29030	0.033	glycine-rich protein
At3g55770	0.033	transcription factor L2
At3g10970	0.033	unknown protein
At5g67290	0.033	putative protein
At3g47670	0.033	putative protein pectinesterase
At2g41220	0.033	ferredoxin-dependent glutamate synthase (GLU2)

Genbank	P-value	Description
At4g03400	0.033	putative GH3-like protein
At5g63840	0.033	glucosidase II alpha subunit
At1g79710	0.033	hypothetical protein
At1g32060	0.033	phosphoribulokinase precursor
At5g18200	0.033	galactose-1-phosphate uridyl transferase-like protein
At3g10020	0.033	unknown protein
At2g05940	0.033	putative protein kinase
At5g18420	0.033	putative protein non-consensus GC donor splice site at exon 1
At1g44960	0.033	unknown protein
At4g05050	0.033	unknown protein
At3g17240	0.033	lipoamide dehydrogenase precursor
At3g08840	0.033	unknown protein
At4g13020	0.033	serine/threonine-specific protein kinase MHK
At2g33710	0.033	putative AP2 domain transcription factor
At1g34245	**0.033**	**Expressed protein**
At3g57390	0.033	MADS transcription factor
At4g34710	0.033	arginine decarboxylase SPE2
At4g38830	0.033	receptor-like protein kinase
At2g45850	0.033	putative AT-hook DNA-binding protein
At1g70780	0.033	unknown protein
At1g64040	0.033	phosphoprotein phosphatase 1
At4g17110	0.033	hypothetical protein
At3g26200	0.033	cytochrome P450
At1g70290	0.033	trehalose-6-phosphate synthase
At1g80080	**0.033**	**receptor protein kinase, TMM**
At2g26400	0.033	unknown protein
At5g15600	0.033	nitrilase associated protein-like NAP16kDa
At3g11520	0.033	putative cyclin
At4g22370	0.033	hypothetical protein
At4g22110	0.033	alcohol dehydrogenase
At3g51610	0.033	putative protein
At3g61580	0.033	delta-8 sphingolipid desaturase
At5g61590	0.033	ethylene responsive element binding factor
At5g47220	0.033	ethylene responsive element binding factor 2 (ATERF2)
At3g20430	0.033	unknown protein
At1g67840	0.031	F12A21.3 unknown protein
At1g19680	0.029	unknown protein
At4g01270	0.026	putative RING zinc finger protein
At5g05870	0.026	glucuronosyl transferase-like protein
At3g13445	0.026	transcription initiation factor TFIID-1
At5g47330	0.026	palmitoyl-protein thioesterase precursor
At4g36180	0.026	putative receptor protein kinase Cf-2.1 leucine rich repeat protein

Genbank	P-value	Description
At4g13730	0.026	putative protein GTPase activating protein
At5g44650	0.026	unknown protein
At2g36310	0.026	hypothetical protein
At1g62290	0.026	aspartic protease
At5g03350	0.026	putative protein
At2g02760	0.026	E2, ubiquitin-conjugating enzyme 2 (UBC2)
At5g58300	0.026	receptor-like protein kinase
At5g52640	0.023	heat-shock protein
At1g70370	0.023	aromatic rich glycoprotein
At2g38310	0.022	unknown protein
At1g22550	0.022	peptide transporter
At2g43220	0.022	hypothetical protein
At5g12100	0.015	putative protein

APPENDIX B:
GENELIST (352 GENES) OBTAINED FROM RACE ANALYSIS

Appendix B. Genelist (352 genes) obtained from RACE analysis

orange = up in yda, down in DN-YDA
blue = down in yda, up in DN-YDA
yellow = no insert available or seed not ordered

AGI	B	Gene Title
AT2G33830	2.88	dormancy/auxin associated family protein
AT1G34245	**2.47**	**expressed protein**
AT1G80080	**2.33**	**leucine-rich repeat family protein**
AT2G26400	2.17	acireductone dioxygenase (ARD/ARD') family protein
AT4G35770	1.99	senescence-associated protein (SEN1)
AT1G56300	1.98	DNAJ heat shock N-terminal domain-containing protein
AT3G55970	1.88	oxidoreductase, 2OG-Fe(II) oxygenase family protein
AT2G24850	1.78	aminotransferase, putative
AT3G22231	1.74	expressed protein
AT1G10070	1.65	branched-chain amino acid aminotransferase 2 / branched-chain amino acid transaminase 2 (BCAT2)
AT2G15960	1.64	expressed protein
AT3G47340	1.59	asparagine synthetase 1 (glutamine-hydrolyzing)
AT5G61590	1.55	AP2 domain-containing transcription factor family protein
AT3G10020	1.49	expressed protein
AT2G16660	1.39	nodulin family protein
AT1G79700	**1.33**	**ovule development protein, putative**
AT5G42900	1.3	expressed protein
AT4G34970	1.26	actin-depolymerizing factor, putative
AT5G18600	1.24	glutaredoxin family protein
AT4G32340	1.19	expressed protein
AT1G27760	1.16	interferon-related developmental regulator family protein
AT1G10060	1.14	branched-chain amino acid aminotransferase 1
AT1G22550	1.04	proton-dependent oligopeptide transport (POT) family protein
AT1G12780	1.03	UDP-glucose 4-epimerase
AT2G13360	0.97	serine-glyoxylate aminotransferase-related
AT4G15530	0.96	pyruvate phosphate dikinase family protein
AT3G48360	0.92	speckle-type POZ protein-related
AT1G80440	0.91	kelch repeat-containing F-box family protein
AT1G70290	0.88	trehalose-6-phosphate synthase, putative
AT2G21660	0.88	glycine-rich RNA-binding protein (GRP7)
AT2G43510	0.87	trypsin inhibitor, putative
AT4G12690	0.82	expressed protein
AT3G13750	0.81	beta-galactosidase, putative / lactase, putative
AT4G28040	0.7	nodulin MtN21 family protein

AGI	B	GeneTitle
AT5G24470	0.64	pseudo-response regulator 5 (APRR5)
AT3G45140	0.64	lipoxygenase (LOX2)
AT5G23380	0.63	expressed protein
AT5G67480	0.59	TAZ zinc finger family protein / BTB/POZ domain-containing protein
AT3G16530	0.59	legume lectin family protein
AT5G60880	**0.58**	**expressed protein**
AT3G49620	0.55	2-oxoacid-dependent oxidase, putative (DIN11)
AT3G60420	0.54	expressed protein
AT3G20810	0.49	transcription factor jumonji (jmjC) domain-containing protein
AT1G71030	0.47	myb family transcription factor
AT1G15380	0.44	lactoylglutathione lyase family protein / glyoxalase I family protein
AT3G13450	0.42	2-oxoisovalerate dehydrogenase / 3-methyl-2-oxobutanoate dehydrogenase / branched-chain alpha-keto acid dehydrogenase E1 beta subunit (DIN4)
AT3G52840	0.41	beta-galactosidase, putative / lactase, putative
AT5G03350	0.41	legume lectin family protein
AT3G22240	0.37	expressed protein
AT5G65660	0.37	hydroxyproline-rich glycoprotein family protein
AT2G05540	0.36	glycine-rich protein
AT5G06690	0.35	thioredoxin family protein
AT1G19540	0.3	isoflavone reductase, putative
AT3G26200	0.29	cytochrome P450 71B22, putative (CYP71B22)
AT4G26260	0.16	expressed protein
AT2G18690	0.15	expressed protein
AT2G32150	0.12	haloacid dehalogenase-like hydrolase family protein
AT5G07580	0.12	ethylene-responsive element-binding family protein
AT4G25170	0.11	expressed protein
AT1G11080	0.11	serine carboxypeptidase S10 family protein
AT3G25770 AT3G25760	0.1	allene oxide cyclase, putative / early-responsive to dehydration protein
AT2G43590	0.07	chitinase, putative
AT4G37870	0.05	phosphoenolpyruvate carboxykinase (ATP), putative / PEP carboxykinase, putative / PEPCK, putative
AT1G28330	0.03	dormancy-associated protein, putative (DRM1)
AT5G21170	0.01	5'-AMP-activated protein kinase beta-2 subunit, putative
AT3G22750	-0	protein kinase, putative
AT1G02450	-0	NPR1/NIM1-interacting protein 1 (NIMIN-1)
AT1G76410	-0.1	zinc finger (C3HC4-type RING finger) family protein
AT1G20620	-0.1	catalase 3 (SEN2)
AT2G20670	-0.1	expressed protein
AT3G05880	-0.2	hydrophobic protein (RCI2A) / low temperature and salt responsive protein (LTI6A)

AGI	B	GeneTitle
AT1G22770	-0.2	gigantea protein (GI)
AT5G22300	-0.3	nitrilase 4 (NIT4)
AT5G57655	-0.3	xylose isomerase family protein
AT2G29460	-0.3	glutathione S-transferase, putative
AT5G20250	-0.3	raffinose synthase family protein / seed imbibition protein
AT1G25275	-0.3	expressed protein
AT3G15630	-0.3	expressed protein
AT2G05520	-0.4	glycine-rich protein (GRP)
AT1G76470	-0.4	cinnamoyl-CoA reductase family
AT2G39030	-0.4	GCN5-related N-acetyltransferase (GNAT) family protein
AT1G19960	-0.4	expressed protein
AT4G38840	-0.4	auxin-responsive protein, putative
AT1G43160	-0.4	AP2 domain-containing protein RAP2.6 (RAP2.6)
AT3G59350	-0.5	serine/threonine protein kinase, putative
AT3G26830	-0.5	cytochrome P450 71B15, putative (CYP71B15)
AT2G03980	-0.5	GDSL-motif lipase/hydrolase family protein
AT2G15890	-0.5	expressed protein
AT5G17440	-0.5	LUC7 N_terminus domain-containing protein
AT5G64240	-0.6	latex-abundant family protein (AMC3) / caspase family protein
AT3G62950	-0.6	glutaredoxin family protein
AT4G19840	-0.6	lectin-related
AT4G23600	-0.6	coronatine-responsive tyrosine aminotransferase / tyrosine transaminase
AT4G03400	-0.6	auxin-responsive GH3 family protein
AT4G27440	-0.6	protochlorophyllide reductase B, chloroplast / PCR B / NADPH-protochlorophyllide oxidoreductase B (PORB)
AT1G21680	-0.6	expressed protein
AT2G31010	-0.6	protein kinase family protein
AT1G32960	-0.6	subtilase family protein
AT1G33590	-0.7	disease resistance protein-related / LRR protein-related
AT4G36930	-0.7	basic helix-loop-helix (bHLH) protein SPATULA (SPT)
AT2G25730	-0.7	expressed protein
AT1G08630	-0.7	L-allo-threonine aldolase-related
AT4G34950	-0.8	nodulin family protein
AT5G05540	-0.8	exonuclease family protein
AT1G15670	-0.8	kelch repeat-containing F-box family protein
AT1G33790	-0.8	jacalin lectin family protein
AT1G79440	-0.8	succinate-semialdehyde dehydrogenase (SSADH1)
AT2G30600	-0.8	BTB/POZ domain-containing protein
AT1G25400	-0.9	expressed protein
AT1G54740	-0.9	expressed protein
AT2G26430	-0.9	ania-6a type cyclin (RCY1)
AT5G10180	-0.9	sulfate transporter

AGI	B	GeneTitle
AT1G03610	-0.9	expressed protein
AT5G05750	-0.9	DNAJ heat shock N-terminal domain-containing protein
AT4G02520 AT2G02930	-0.9	glutathione S-transferase, putative /// glutathione S-transferase, putative
AT1G35612	-0.9	expressed protein
AT2G38400	-0.9	alanine--glyoxylate aminotransferase, putative / beta-alanine-pyruvate aminotransferase, putative / AGT, putative
AT4G39950	-0.9	cytochrome P450 79B2, putative (CYP79B2)
AT5G18140	-1	DNAJ heat shock N-terminal domain-containing protein
AT5G17300	-1	myb family transcription factor
AT3G18530 AT3G01450	-1	expressed protein /// expressed protein
AT3G20340	-1	expressed protein
AT3G29970	-1	germination protein-related
AT2G22450	-1	riboflavin biosynthesis protein, putative
AT1G53885	-1	senescence-associated protein-related
AT5G61520	-1	hexose transporter, putative
AT3G06500	-1	beta-fructofuranosidase, putative / invertase
AT5G44400	-1	FAD-binding domain-containing protein
AT4G01870	-1	tolB protein-related
AT1G28050	-1.1	zinc finger (B-box type) family protein
AT5G57350	-1.1	ATPase 3, plasma membrane-type / proton pump 3
AT4G03510	-1.1	zinc finger (C3HC4-type RING finger) family protein (RMA1)
AT3G47950	-1.1	ATPase, plasma membrane-type
AT1G13670	-1.1	expressed protein
AT5G49730 AT5G49740	-1.1	ferric reductase-like transmembrane component family protein
AT2G34600	-1.1	expressed protein
AT2G30520	-1.1	signal transducer of phototropic response (RPT2)
AT5G54960	-1.1	pyruvate decarboxylase, putative
AT2G18700	-1.1	glycosyl transferase family 20 protein / trehalose-phosphatase family protein
AT1G76990	-1.1	ACT domain containing protein
AT3G28300 AT3G28290	-1.1	integrin-related protein 14a /// integrin-related protein 14a
AT1G18020 AT1G17990	-1.1	12-oxophytodienoate reductase, putative /// 12-oxophytodienoate reductase
AT1G56220	-1.1	dormancy/auxin associated family protein
AT3G26780	-1.1	phosphoglycerate/bisphosphoglycerate mutase family protein
AT2G31810	-1.2	acetolactate synthase small subunit, putative
AT1G44350	-1.2	IAA-amino acid hydrolase 6, putative (ILL6) / IAA-Ala hydrolase, putative
AT3G59950 AT3G59940	-1.2	autophagy 4b (APG4b) /// kelch repeat-containing F-box family protein
AT2G46650	-1.3	cytochrome b5, putative

AGI	B	GeneTitle
AT4G18340	-1.3	glycosyl hydrolase family 17 protein
AT2G22990	-1.3	sinapoylglucose:malate sinapoyltransferase (SNG1)
AT5G24800	-1.3	bZIP transcription factor family protein
AT1G75800	-1.3	pathogenesis-related thaumatin family protein
AT3G14990	-1.4	4-methyl-5(b-hydroxyethyl)-thiazole monophosphate biosynthesis protein, putative
AT1G70700	-1.4	expressed protein
AT2G02710	-1.4	PAC motif-containing protein
AT4G29270	-1.4	acid phosphatase class B family protein
AT5G24490	-1.4	30S ribosomal protein, putative
AT5G19120	-1.4	expressed protein
AT5G09220	-1.5	amino acid permease 2 (AAP2)
AT4G11360	-1.5	zinc finger (C3HC4-type RING finger) family protein (RHA1b)
AT3G09410	-1.5	pectinacetylesterase family protein
AT1G48100	-1.5	glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein
AT5G11070	-1.5	expressed protein
AT2G22980	-1.5	serine carboxypeptidase S10 family protein
AT5G48250	-1.6	zinc finger (B-box type) family protein
AT1G03090	-1.6	methylcrotonyl-CoA carboxylase alpha chain, mitochondrial / 3-methylcrotonyl-CoA carboxylase 1 (MCCA)
AT5G61380	-1.6	ABI3-interacting protein 1 (AIP1)
AT5G62360	-1.6	invertase/pectin methylesterase inhibitor family protein
AT3G53800	-1.6	armadillo/beta-catenin repeat family protein
AT1G10660	-1.7	expressed protein
AT1G50420	-1.7	scarecrow-like transcription factor 3 (SCL3)
AT4G14270	-1.7	expressed protein
AT2G40750	-1.7	WRKY family transcription factor
AT5G50450	-1.8	zinc finger (MYND type) family protein
AT3G26960	-1.8	expressed protein
AT4G24050	-1.8	short-chain dehydrogenase/reductase (SDR) family protein
AT4G27800	-1.8	protein phosphatase 2C PPH1 / PP2C PPH1 (PPH1)
AT1G02930 AT1G02920	-1.8	glutathione S-transferase, putative /// glutathione S-transferase, putative
AT5G27920	-1.8	F-box family protein
AT5G40450	-1.8	expressed protein
AT3G21260	-1.8	glycolipid transfer protein-related
AT4G37580	-1.8	N-acetyltransferase, putative / hookless1 (HLS1)
AT4G22710 AT4G22690	-1.8	cytochrome P450 family protein /// cytochrome P450 family protein
AT1G26800	-1.9	zinc finger (C3HC4-type RING finger) family protein
AT2G43570	-1.9	chitinase, putative
AT1G02640	-1.9	glycosyl hydrolase family 3 protein

AGI	B	GeneTitle
AT5G56870	-1.9	beta-galactosidase, putative / lactase, putative
AT1G63180	-1.9	UDP-glucose 4-epimerase, putative / UDP-galactose 4-epimerase
AT5G07010	-1.9	sulfotransferase family protein
AT4G16130	-1.9	GHMP kinase family protein
AT2G16365	-1.9	F-box family protein
AT4G21870	-2	26.5 kDa class P-related heat shock protein (HSP26.5-P)
AT2G07711	-2	pseudogene, similar to NADH dehydrogenase subunit 5
AT4G38860	-2	auxin-responsive protein, putative
AT5G01240	-2	amino acid permease, putative
AT2G30540	-2	glutaredoxin family protein
AT3G28600	-2.1	AAA-type ATPase family protein
AT3G53460	-2.1	29 kDa ribonucleoprotein, chloroplast / RNA-binding protein cp 29
AT5G53160	-2.1	expressed protein
AT4G35240	-2.1	expressed protein
AT5G45670	-2.1	GDSL-motif lipase/hydrolase family protein
AT2G28305	-2.1	expressed protein
AT1G60140	-2.1	glycosyl transferase family 20 protein / trehalose-phosphatase family protein
AT3G16857	-2.1	two-component responsive regulator family protein / response regulator family protein
AT4G37470	-2.1	hydrolase, alpha/beta fold family protein
AT5G62720	-2.1	integral membrane HPP family protein
AT5G14620	-2.2	cytosine methyltransferase (DRM2)
AT1G68530	-2.2	very-long-chain fatty acid condensing enzyme (CUT1)
AT1G06460	-2.2	31.2 kDa small heat shock family protein / hsp20 family protein
AT5G46690	-2.2	basic helix-loop-helix (bHLH) family protein
AT1G70410	-2.2	carbonic anhydrase, putative / carbonate dehydratase, putative
AT1G52000	-2.2	jacalin lectin family protein
AT1G11380	-2.2	expressed protein
AT5G05440	-2.2	expressed protein
AT2G43010	-2.2	phytochrome-interacting factor 4 (PIF4) / basic helix-loop-helix protein 9 (bHLH9)
AT4G25520	-2.3	transcriptional co-regulator family protein
AT3G14660 AT3G14650	-2.3	cytochrome P450, putative /// cytochrome P450, putative
AT2G01100	-2.3	expressed protein
AT4G28190	-2.3	expressed protein
AT1G02060	-2.3	pentatricopeptide (PPR) repeat-containing protein
AT4G04040	-2.3	pyrophosphate--fructose-6-phosphate 1-phosphotransferase beta subunit
AT5G13330	-2.3	AP2 domain-containing transcription factor family protein
AT1G30135	-2.4	expressed protein
AT2G30770	-2.4	cytochrome P450 71A13, putative (CYP71A13)

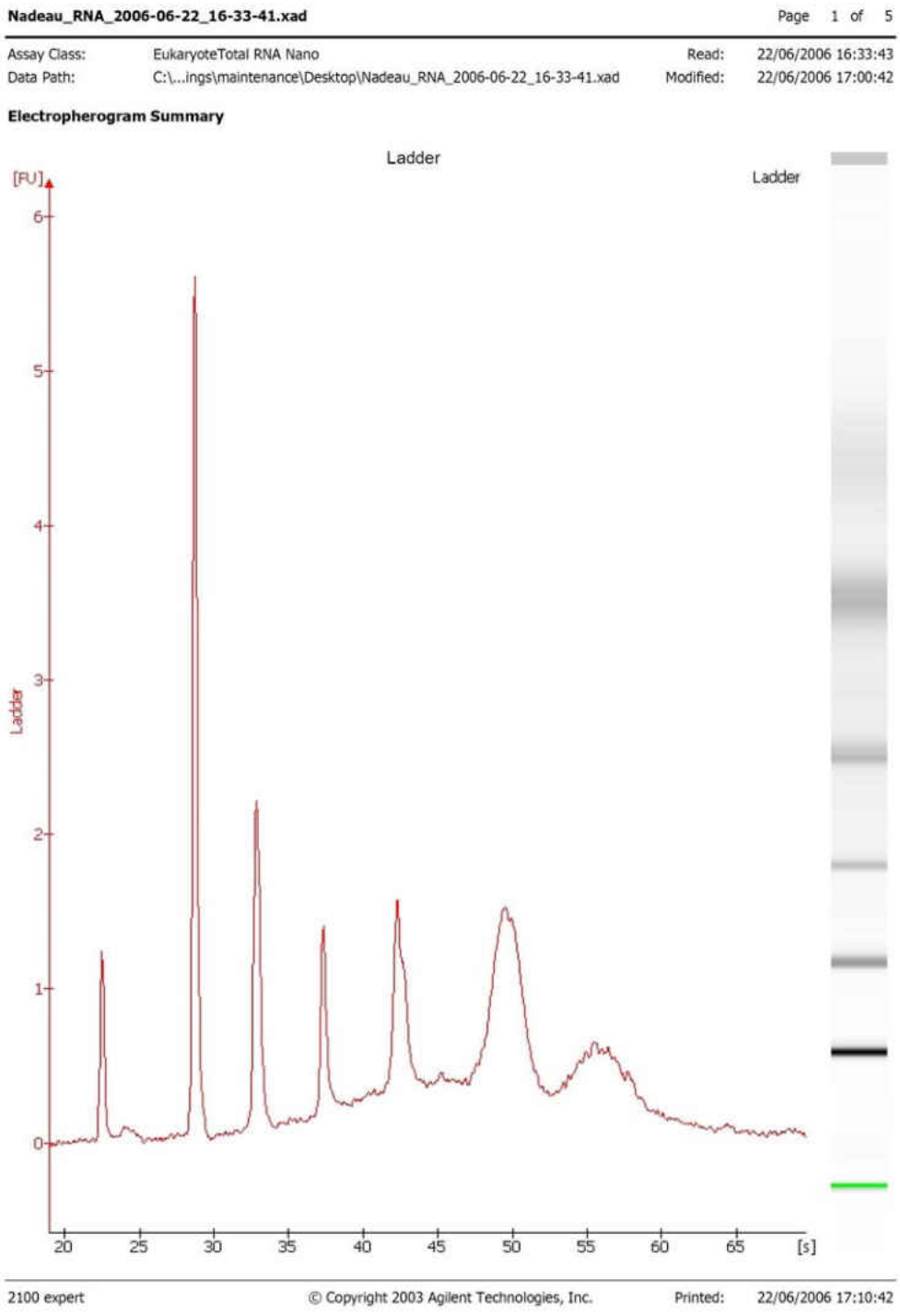
AGI	B	GeneTitle
AT5G02150	-2.4	expressed protein
AT3G49260	-2.4	calmodulin-binding family protein
AT1G75180 AT5G04790	-2.4	expressed protein /// hypothetical protein
AT5G02160	-2.5	expressed protein
AT1G15180	-2.5	MATE efflux family protein
AT4G16330	-2.5	oxidoreductase, 2OG-Fe(II) oxygenase family protein
AT1G26450	-2.5	beta-1,3-glucanase-related
AT1G45976	-2.5	expressed protein
AT5G07870	-2.5	transferase family protein
AT1G13740	-2.5	expressed protein
AT5G08350	-2.6	GRAM domain-containing protein / ABA-responsive protein-related
AT2G22000	-2.6	expressed protein
AT4G29950	-2.6	microtubule-associated protein
AT3G60360	-2.6	expressed protein
AT1G61800	-2.7	glucose-6-phosphate/phosphate translocator, putative
AT1G22710	-2.7	sucrose transporter / sucrose-proton symporter (SUC2)
AT2G42790	-2.7	citrate synthase, glyoxysomal, putative
AT1G77060	-2.7	mutase family protein
AT4G28703	-2.7	expressed protein
AT5G49980	-2.7	transport inhibitor response protein, putative
AT1G27650	-2.7	U2 snRNP auxiliary factor small subunit, putative
AT5G19260	-2.7	expressed protein
AT5G22340	-2.7	expressed protein
AT4G03205	-2.8	coproporphyrinogen III oxidase, putative / coproporphyrinogenase
AT2G47930	-2.8	hydroxyproline-rich glycoprotein family protein
AT3G14560	-2.8	expressed protein
AT3G04290	-2.8	GDSL-motif lipase/hydrolase family protein
AT1G77000	-2.8	F-box family protein
AT5G47720	-2.8	acetyl-CoA C-acyltransferase, putative / 3-ketoacyl-CoA thiolase, putative
AT1G11350	-2.8	S-locus lectin protein kinase family protein
AT1G75180 AT5G04790	-2.9	expressed protein /// hypothetical protein
AT5G18525 AT5G18530	-2.9	WD-40 repeat family protein /// beige/BEACH domain-containing protein
AT5G67440	-2.9	phototropic-responsive NPH3 family protein
AT5G40910	-2.9	disease resistance protein (TIR-NBS-LRR class), putative
AT1G01120	-2.9	fatty acid elongase 3-ketoacyl-CoA synthase 1 (KCS1)
AT3G51270	-2.9	RIO1 family protein
AT4G23870	-3	expressed protein
AT1G17145	-3	expressed protein
AT3G15620	-3	6-4 photolyase (UVR3)

AGI	B	GeneTitle
AT1G33050	-3	expressed protein
AT3G14570	-3	glycosyl transferase family 48 protein
AT2G03890	-3	phosphatidylinositol 3- and 4-kinase family protein
AT2G43190	-3	ribonuclease P family protein
AT3G47160	-3	expressed protein
AT1G50600	-3	scarecrow-like transcription factor 5 (SCL5)
AT4G30660	-3	hydrophobic protein, putative / low temperature and salt responsive protein, putative
AT5G55100	-3.1	SWAP (Suppressor-of-White-APricot)/surp domain-containing protein
AT4G30650	-3.1	hydrophobic protein, putative / low temperature and salt responsive protein, putative
AT1G25230	-3.1	purple acid phosphatase family protein
AT2G37480	-3.1	expressed protein
AT1G04300	-3.2	meprin and TRAF homology domain-containing protein / MATH domain-containing protein
AT3G58000	-3.2	VQ motif-containing protein
AT2G41560	-3.2	calcium-transporting ATPase 4, plasma membrane-type
AT1G10450 AT1G10250	-3.2	paired amphipathic helix repeat-containing protein /// expressed protein
AT2G19970	-3.3	pathogenesis-related protein, putative
AT5G14270	-3.3	DNA-binding bromodomain-containing protein
AT1G55370	-3.3	expressed protein
AT3G10450	-3.3	serine carboxypeptidase S10 family protein
AT3G57330	-3.3	calcium-transporting ATPase, plasma membrane-type
AT1G75540	-3.4	zinc finger (B-box type) family protein
AT1G53035	-3.4	expressed protein
AT5G41080	-3.4	glycerophosphoryl diester phosphodiesterase family protein
AT2G22420	-3.5	peroxidase 17 (PER17) (P17)
AT1G56680	-3.5	glycoside hydrolase family 19 protein
AT1G44800	-3.5	nodulin MtN21 family protein
AT4G11900	-3.5	S-locus lectin protein kinase family protein
AT1G05890 AT1G63760	-3.5	zinc finger protein-related /// pseudogene, ARI protein
AT1G05410	-3.6	expressed protein
AT3G49840	-3.6	proline-rich family protein
AT1G65500	-3.6	expressed protein
AT3G25610	-3.7	haloacid dehalogenase-like hydrolase family protein
AT5G64880	-3.7	expressed protein
AT4G07456	-3.7	gypsy-like retrotransposon family (Athila)
AT5G06190	-3.8	expressed protein
AT3G05120	-3.9	expressed protein
AT3G51500	-3.9	expressed protein
AT3G29370	-3.9	expressed protein

AGI	B	GeneTitle
AT3G07010	-3.9	pectate lyase family protein
AT4G05150	-3.9	octicosapeptide/Phox/Bem1p (PB1) domain-containing protein
AT1G29500	-3.9	auxin-responsive protein, putative
AT5G23350 AT5G23360	-3.9	GRAM domain-containing protein / ABA-responsive protein-related
AT5G49400	-4	zinc knuckle (CCHC-type) family protein
AT2G23580	-4	hydrolase, alpha/beta fold family protein
AT2G03470	-4	myb family transcription factor / ELM2 domain-containing protein
AT4G19390	-4	expressed protein
AT5G66090	-4	expressed protein
AT5G57785	-4	expressed protein
AT5G23340	-4	expressed protein
AT3G26750	-4.1	expressed protein
AT2G46400	-4.1	WRKY family transcription factor
AT1G16950	-4.1	expressed protein
AT1G11220	-4.2	expressed protein
AT5G10290	-4.2	leucine-rich repeat family protein / protein kinase family protein
AT4G01350 AT5G59920	-4.3	DC1 domain-containing protein /// DC1 domain-containing protein
AT4G29000	-4.3	tesmin/TSO1-like CXC domain-containing protein
AT1G53160	-4.4	squamosa promoter-binding protein-like 4 (SPL4)
AT5G16970 AT5G16980 AT5G16990 AT5G17000	-4.4	NADP-dependent oxidoreductase, putative (P1) /// NADP-dependent oxidoreductase, putative /// NADP-dependent oxidoreductase, putative /// NADP-dependent oxidoreductase, putative
AT5G60400	-4.4	expressed protein
AT1G53260	-4.4	hypothetical protein
AT3G11970	-4.5	gypsy-like retrotransposon family
AT3G26240 AT3G26250	-4.6	DC1 domain-containing protein /// DC1 domain-containing protein
AT3G43860	-4.6	glycosyl hydrolase family 9 protein
AT3G49770	-4.6	hypothetical protein
AT1G09350	-4.6	galactinol synthase, putative
AT5G25530	-4.6	DNAJ heat shock protein, putative
AT3G58870	-4.7	expressed protein
AT2G14460	-4.7	expressed protein
AT5G54300	-4.7	expressed protein
AT1G77330	-4.8	1-aminocyclopropane-1-carboxylate oxidase, putative / ACC oxidase, putative
AT2G37000	-4.8	TCP family transcription factor, putative
AT5G58340	-4.8	expressed protein
AT5G34870	-4.8	zinc knuckle (CCHC-type) family protein
AT1G40230	-4.8	hypothetical protein
AT2G27550	-4.9	centroradialis protein, putative (CEN)

AGI	B	GeneTitle
NA	-4.9	NA
AT4G25070	-4.9	expressed protein
AT2G23110	-5	expressed protein
AT4G02140	-5	expressed protein
AT2G32460	-5	myb family transcription factor (MYB101)
AT1G79770	-5	expressed protein
AT1G61590	-5	protein kinase, putative
AT1G64360	-5.1	expressed protein
AT3G03540 AT3G03530	-5.1	phosphoesterase family protein /// phosphoesterase family protein
AT2G06230	-5.1	hypothetical protein
AT3G01390	-5.1	vacuolar ATP synthase subunit G 1 (VATG1) / V-ATPase G subunit 1 (VAG1) / vacuolar proton pump G subunit 1 (VMA10)
AT1G65620	-5.1	LOB domain protein 6 / lateral organ boundaries domain protein 6 (LBD6) / asymmetric leaves2 (AS2)
AT3G42720 AT2G06720 AT1G39110 AT4G08013 AT5G28165 AT5G29058	-5.2	CACTA-like transposase family (Ptta/En/Spm) /// CACTA-like transposase family (Ptta/En/Spm) /// CACTA-like transposase family (Ptta/En/Spm) /// hypothetical protein /// CACTA-like transposase family (Tnp2/En/Spm) /// pseudogene, hypothetical protein
AT2G42870	-5.3	expressed protein

APPENDIX C:
ELECTRPPHEROGRAM SHOWING QUALITY OF RNA



Appendix C. Electropherogram showing RNA quality of sample tissue

APPENDIX D:
FILTER ON FLAGS SHOWING NUMBER OF GENES PRESENT

Appendix D: Filter on flags to determine genes with present and absent flags

Tissue sample type	Present	Absent
All samples	14015	11,118
1-1 *tmm*-1	13,028	9,241
1-3 *tmm*-1	11,587	10,599
1-2 col	12,968	9,312
1-4 col	13,050	9,291
Tissue type col	13,550	9,864
Tissue type *tmm*	13,292	10,886

APPENDIX E:
FUNCTIONAL CLASSIFICATION OF GENES WITH DAVID BIOINFORMATICS TOOL

Appendix E. Gene Functional Classification containing 8 clusters with enrichment score for each group. Higher enrichment score means that category is more significant

Gene group 1 Enrichment Score: 4.54	
Affy_ID	**Function**
266037_at	hypothetical protein at2g05940
252964_at	receptor-like protein kinase-like protein
253937_at	putative npk1-related protein kinase
254747_at	serine/threonine-protein kinase mhk
251789_at	serine/threonine-specific protein kinase-like
254084_at	sumo activating enzyme
255257_at	polyubiquitin (ubq3)
255257_at	ubiquitin / ribosomal protein cep52
247080_at	proteasome component 6c
260475_at	serine carboxypeptidase s10 family protein
261862_at	putative dna-binding protein
247856_at	at5g58299/at5g58299
255257_at	polyubiquitin
247670_at	arabidopsis thaliana genomic dna, chromosome 5, bac clone:f15l12
253078_at	leucine-rich repeat family protein
256221_at	at1g56300
257926_at	at3g23280
255257_at	ubiquitin / ribosomal protein s27a.2
255257_at	polyubiquitin 2 alt_names:ubiquitin-like protein 7
255939_at	hypothetical protein at1g12730
267484_at	ubiquitin-conjugating enzyme e2-17 kda 2
261140_at	f6f9.27 protein
249862_at	at5g22920/mrn17_15
252575_at	receptor like protein kinase
255228_at	f-box family protein (fbl21)
255257_at	hypothetical protein at4g02890
248332_at	at5g52640/f6n7_13
255257_at	protein t19e23.13 [imported]
255613_at	putative ring zinc finger protein
248812_at	palmitoyl-protein thioesterase-like
245697_at	metacaspase 9 precurser
264741_at	putative aspartic protease

Gene group 1 Enrichment Score: 4.54	
Affy_ID	**Function**
255257_at	protein f28c11.5 [imported]
255257_at	ubiquitin / ribosomal protein s27a.1 alt_names:protein t17j13.210; ubiquitin extension protein 1
250729_at	ubiquitin activating enzyme 2
253297_at	beta-xylan endohydrolase-like protein
245528_at	protein at4g15530
259982_at	ring-h2 finger protein atl8
Gene group 2 Enrichment Score: 3.59	
260928_at	hypothetical protein at1g02720/t14p4_8
257056_at	glycosylation enzyme-like protein
248721_at	hypothetical protein at5g47780
252863_at	myo-inositol-1-phosphate synthase 1
258990_at	hypothetical protein at3g08840
247298_at	glucosidase ii alpha subunit
Gene group 3 Enrichment Score: 3.5	
259375_at	at3g16370/mya6_18
259788_at	at1g29670/f15d2_22
263482_at	putative gdsl-motif lipase/hydrolase
263156_at	at1g54030/f15i1_11
Gene group 4 Enrichment Score: 3.32	
245101_at	cytochrome p450 98a3
256186_at	at4cl1
247358_at	flavonol synthase, putative
259975_at	putative cinnamoyl-coa reductase; 27707-26257

Gene group 5 Enrichment Score: 3.1	
246985_at	similarity to unknown protein
246144_at	at5g20110
267034_at	hypothetical protein
246028_at	akin beta1
265494_at	putative calmodulin-like protein
249153_s_at	emb|cab86628.1
246346_at	hypothetical protein
254730_at	hypothetical protein
252299_at	f-box family protein
261828_at	universal stress protein (usp) family protein
245201_at	protein at1g67840
251704_at	hypothetical protein
267393_at	protein at2g44500
254354_at	hypothetical protein at4g22370
262025_at	putative growth regulator
266805_at	at2g30010/f23f1.7
249007_at	arabidopsis thaliana genomic dna, chromosome 5, tac clone:k15c23
263613_at	hypothetical protein at2g25250

Gene group 6 Enrichment Score: 3.09	
251132_at	myb transcription factor
267495_at	hypothetical protein at2g30420
265060_at	f5f19.21 protein
264415_at	rap2.6
259751_at	putative transcription factor
256961_at	tata-binding factor 1
248794_at	ethylene-responsive transcription factor 2
251623_at	mads-box protein agl18
261395_at	ovule development protein, putative
248596_at	similarity to myb-related transcription factor
267451_at	putative ap2 domain transcription factor
245358_at	pp1/pp2a phosphatases pleiotropic regulator prl1
247540_at	putative ethylene responsive element binding factor
260904_at	hypothetical protein at1g02450/t6a9_28
263210_at	david_3583214
261857_at	hypothetical protein f17j6.14
267159_at	putative scarecrow gene regulator
259412_at	dna binding protein gt-1
Gene group 7 Enrichment Score: 1.41	
258554_at	putative dead/deah box helicase
262899_at	pdr8 abc transporter
256440_at	haloacid dehalogenase-like hydrolase family protein
245201_at	protein at1g67840
Gene group 8 Enrichment Score: 0.93	
249012_at	flavonoid 3',5'-hydroxylase-like; cytochrome p450
257636_at	cytochrome p450 71b22
245101_at	cytochrome p450 98a3
266996_at	at2g34490/t31e10.17

APPENDIX F:
GLOSSARY

Glossary

- Stomatal regulators: Genes involved in stomatal development
- Asymmetric divisions: Cell divisions that lead to formation of one small and one bigger cell. The resulting daughter cells assume different cell fate.
- Amplification divisions: Series of asymmetric cell divisions occurring in meristemoids that add new cells to the epidermis, each time renewing the meristemoid
- Stomatal lineage cells: Meristemoid mother cell, meristemoid, guard mother cell and guard cell
- Stomatal cell lineage marker: Genes expressed in stomatal lineage cells
- GUS: a β-glucuronidase enzyme used a reporter to visualize gene expression
- *TMM*: there are two alleles of *TMM* they are *tmm-1* and *tmm-2*. Throughout this paper only *tmm-1* has been referred to.
- Meristemoid: A plant stem cell with limited self-renewal capacity. Ultimately differentiates to guard mother cell in wild-type tissue.
- Guard Mother cell: A precursor cell that divides symmetrically to produce two guard cells of stomata.

REFERENCES

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., and Ecker, J.R. (2003). Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana*. Science **301,** 653-657.

Anderson, J., A.,, Huprikar, S., S.,, Kochian, L., V.,, Lucas, W., J.,, and Gaber, R., F., . (1992). Functional expression of a probable Arabidopsis thaliana potassium channel in Saccharomyces cerevisiae. Proc Natl Acad Sci **89(9),** 3736–3740.

Beemster, G.T.S., De Veylder, L., Vercruysse, S., West, G., Rombaut, D., Van Hummelen, P., Galichet, A., Gruissem, W., Inze, D., and Vuylsteke, M. (2005). Genome-Wide Analysis of Gene Expression Profiles Associated with Cell Cycle Transitions in Growing Organs of Arabidopsis. Plant Physiology **138,** 734-743.

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

Charlton, W.A. (1990). Differentiation in the leaf epidermis of *Chlorophytum cosmosum* Baker. Ann. Bot. **66,** 567-578.

Craik, D.J., Daly, N.L., Bond, T., and Waine, C. (1999). Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. Journal of Molecular Biology **294,** 1327-1336.

Cutler, S.R., Ehrhardt, D.W., Griffitts, J.S., and Somerville, C. (2000). Random GFP::cDNA fusions enable visualization of subcellular structures in cells of *Arabidopsis* at high frequency. Proceedings of the National Academy of Sciences of the United States of America **97,** 3718-3723.

Duckert, P., Brunak, S., and Blom, N. (2004). Prediction of proprotein convertase cleavage sites. Protein Eng Des Sel **17,** 107-112.

Elliott, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W., Gerentes, D., Perez, P., and Smyth, D.R. (1996). AINTEGUMENTA, an APETALA2-like Gene of Arabidopsis with Pleiotropic Roles in Ovule Development and Floral Organ Growth. Plant Cell **8,** 155-168.

Geisler, M., Yang, M., and Sack, F.D. (1998). Divergent regulation of stomatal initiation and patterning in organ and suborgan regions of the *Arabidopsis* mutants *too many mouths* and *four lips*. Planta **205,** 522-530.

Geisler, M.D., Nadeau, J.A., 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.

Harris, R.B. (1989). Processing of pro-hormone precursor proteins. Archives of Biochemistry and Biophysics **275,** 315-333.

Hernandez, M.L., Passas, H.J., and Smith, L.G. (1999). Clonal Analysis of Epidermal Patterning during Maize Leaf Development. Developmental Biology **216,** 646-658.

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

Jofuku, K.D., den Boer, B.G., Van Montagu, M., and Okamuro, J.K. (1994). Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. Plant Cell **6,** 1211–1225.

Kim, J.H., Choi, D., and Kende, H. (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. The Plant Journal **36,** 94-104.

Larkin, J.C., Marks, M.D., Nadeau, J., and Sack, F. (1997). Epidermal cell fate and patterning in leaves. Plant Cell **9,** 1109-1120.

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.

Matsubayashi, Y., and Sakagami, Y. (1996). Phytosulfokine, Sulfated Peptides that Induce the Proliferation of Single Mesophyll Cells of Asparagus officinalis L. Proceedings of the National Academy of Sciences of the United States of America **93,** 7623-7627.

Matsubayashi, Y., Takagi, L., and Sakagami, Y. (1997). Phytosulfokine-, a sulfated pentapeptide, stimulates the proliferation of rice cells by means of specific high- and low-affinity binding sites. Proc. Natl. Acad. Sci. **94,** 13357-13362.

Matsubayashi, Y., Takagi, L., Omura, N., Morita, A., and Sakagami, Y. (1999). The Endogenous Sulfated Pentapeptide Phytosulfokine-α Stimulates Tracheary Element Differentiation of Isolated Mesophyll Cells of Zinnia. Plant Physiol **120,** 1043–1048.

Mizukami, Y., and Fischer, R.L. (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. Proceedings of the National Academy of Science U S A **97,** 942-947.

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

Nakai, K., and Kanehisa, M. (1991). Expert system for predicting protein localization sites in gram-negative bacteria. Proteins **1,** 95-110.

Nakamura, R.L., McKendree., W.L., Jr., Hirsch, R.E., Sedbrook, J.C., Gaber, R.F., and Sussmank, M.R. (1995). Expression of an *Arabidopsis* potassium channel gene in guard cells. Plant Physiology **109,** 371-374.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering, **10,** 1-6.

Nole-Wilson, S., Tranby, T.L., and Krizek, B.A. (2005). AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. Plant Molecular Biology **V57,** 613-628.

Obayashi, T., Kinoshita, K., Nakai, K., Shibaoka, M., Hayashi, S., Saeki, M., Shibata, D., Saito, K., and Ohta, H. (2007). ATTED-II: a database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. Nucleic Acids Research **35,** D863-869.

Pearce, G., Strydom, D., Johnson, S., and Ryan, C.A. (1991). A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. Science **253,** 895-898.

Pearce, G., Moura, D.S., Stratmann, J., and Ryan, C.A., Jr. (2001). RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. Proceedings of the National Academy of Science U S A **98,** 12843-12847.

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.

Psarros, M., Heber, S., Sick, M., Thoppae, G., Harshman, K., and Sick, B. (2005). RACE: Remote Analysis Computation for gene Expression data. Nucleic Acids Research **33,** W638-W643.

Ryan, C., A. ,, Pearce, G., Scheer, J., and Moura, S., D. (2002). Polypeptide Hormones. Plant Cell **14(Supplement),** s251–s264.

Saether, O., Craik, D.J., Campbel1, I.D., Sletten, K., Juul, J., and Norman, D., G. . (1995). Elucidation of the Primary and Three-Dimensional Structure of the Uterotonic Polypeptide Kalata B 1. Biochemistry **34,** 4147-4158.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., and Lohmann, J. (2005). A gene expression map of Arabidopsis development. Nature Genetics **37,** 501-506.

Schopfer, C.R., Nasrallah, M.E., and Nasrallah, J.B. (1999). The Male Determinant of Self-Incompatibility in Brassica. Science **286,** 1697-1700.

Schultz, J., Milpetz, F., Bork, P., and Ponting, C. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA **95,** 5857-5864.

Serna, L., and Fenoll, C. (2000). Stomatal development in Arabidopsis: How to make a functional pattern. Trends in Plant Science **5,** 458-460.

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.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J.D., Dean, C., Ma, H., and Martienssen, R. (1995). Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements Genes & Development **9,** 1797-1810.

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.

Yang, H., Matsubayashi, Y., Nakamura, K., and Sakagam, Y. (1999b). Oryza sativa PSK gene encodes a precursor of phytosulfokine-, a sulfated peptide growth factor found in plants National Academic Science USA **96,** 13560-13565.

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.

Zhao, L., and Sack, F.D. (1999). Ultrastructure of stomatal development in *Arabidopsis* (Brassicaceae) leaves. American Journal of Botany **86,** 929-939.