ASSESSING THE CONTRIBUTION OF HAN FAMILY GATA FACTORS TO ARABIDOPSIS EMBRYOGENESIS

by

MATTHEW FRANCIS VOLNY

(Under the Direction of Wolfgang Lukowitz)

ABSTRACT

The diverse architecture of many plant species belies their similar developmental origin; each was derived from a single fertilized egg cell. Subsequent divisions of this cell and its daughters, produces a simplified version of the adult plant, comprised of an apical and basal meristem interconnected by a central vascular cylinder. But how is the spatial arrangement of the meristems decided upon, and what comprises the molecular network regulating this process? In this dissertation I detail my efforts towards elucidating the molecular mechanisms regulating embryonic polarity in *Arabidopsis thaliana*. Previously, it was shown that the GATA transcription factor HANABA TARANU is required for positioning the inductive boundary at which the root initiates; *han* mutants display an apical shift in gene expression domains normally delimited by the proembryo boundary, yet often recover to form complete seedlings. I show that their recovery is due to the action of two closely related genes, HAN-LIKE 1 (HANL1) and HANL2. Loss of all three HAN genes produces embryos that are initially similar to *han* single mutants, but over time deteriorate and arrest as blimp-shaped structures lacking a discernible shoot or root apical meristem, and vascular precursors. Molecular markers of embryonic polarity, such as WUS and PLT1, remain unchanged in these embryos, yet many regulators of

organogenesis or meristem maintenance, such as FIL, TMO5, and CLV3 are never expressed. Strikingly, genes normally confined to the basal side of the proembryo boundary, such as WOX5 and WOX8, which are normally transcribed in the precursors of the root quiescent center and distal root, expand into sub-surface and surface cell layers. I interpret this effect as partial radialization of the apical/basal pattern. HAN family genes are characterized by a short, deeply conserved amphiphilic stretch at their N-terminus, called the HAN domain. Through a number of approaches, I show that this domain is required for normal activity, and functions to mediate interactions between HAN and transcriptional co-repressors of the TPL family. These results suggest that HAN genes, which are transcribed in the proembryo, may function to directly antagonize the expression of basal or suspensor transcripts in this domain.

INDEX WORDS: Embryogenesis, Patterning, Development, *Arabidopsis thaliana*,

ASSESSING THE CONTRIBUTION OF HAN FAMILY GATA FACTORS TO ARABIDOPSIS EMBRYOGENESIS

by

MATTHEW FRANCIS VOLNY

B.S., State University of New York at Oswego, 2008

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

ASSESSING THE CONTRIBUTION OF HAN FAMILY GATA FACTORS TO ARABIDOPSIS EMBRYOGENESIS

by

MATTHEW FRANCIS VOLNY

Major Professor:	Wolfgang Lukowitz
Committee:	Kelly Dawe
	Xiaoyu Zhang
	Richard Meagher
	Douglas Menke

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2015

DEDICATION

For Kristin and Roo.

ACKNOWLEDGEMENTS

I would like to acknowledge my committee - Dr. Kelly Dawe, Dr. Xiaoyu Zhang, Dr. Rich Meagher, and Dr. Doug Menke, and I would especially like to thank my advisor Dr. Wolfgang Lukowitz for his invaluable help.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

The plant body is comprised of two distinct organ systems. The aboveground portion includes the stems, leaves, and flowers, while the belowground portion is made of an extended network of roots. These elaborate structures are formed post-embryonically in a reiterative process occurring throughout the life of the plant (Clowes, 1961; Dolan et al., 1993). The origin of the shoot and root can be traced to two primary meristems set at either pole of the seedling, an arrangement that is first realized during embryogenesis (Natesh and Rau, 1984). Beginning with a fertilized egg cell, or zygote, subsequent divisions produce the first stem cells and tissue precursors, and by the end of embryogenesis, the formation of a simplified version of the adult plant (Esau, 1977). The mature embryo exhibits a stereotypical organization, two primary meristems interconnected by a vascular cylinder, and a radial pattern of concentric tissue layers.

The process of embryogenesis is best studied in *Arabidopsis thaliana*, where cell divisions follow a regular, predictable pattern (Mansfield and Briarty, 1991; Fig. 1.1). After fertilization, the single-celled zygote elongates and divides asymmetrically producing a small apical daughter cell and a larger basal daughter cell. The apical cell undergoes a single transverse division followed by two longitudinal divisions, forming a spherical proembryo of 8-cells, while the basal cell and its descendants divide transversely, producing the hypophysis and a stalk-like suspensor. At this stage, four distinct regions can be recognized by anatomy: an upper tier, which will give rise to the shoot apical meristem and most of the cotyledons; a lower

tier, which will generate the hypocotyl, basal portion of the cotyledons, and proximal root; the hypophysis, which will produce the root apical meristem; and the suspensor, which connects the embryo proper to the maternal tissue (Jürgens and Mayer, 1994; Fig. 1.1).

Clonal analyses have found that the contribution of individual cells or groups of cells to the seedling body plan is highly predictable (Scheres et al., 1994; Kidner et al., 2000; Saulsberry et al., 2002). Such a stereotypical pattern of divisions might imply that cell fates are regulated by lineage; but cell ablation studies in the root meristem provide strong evidence that cell fate is determined by position instead (Van den Berg et al., 1995). These results seem to indicate that cell communication plays a prominent role in the patterning process. The hormone auxin is a pervasive long-range signal, and is required for organizing the apical-basal axis while promoting organogenesis (Friml et al., 2003; Cheng et al., 2007; De Smet et al., 2010). In spite of this, it remains unclear whether auxin impacts very early fate decisions during embryogenesis. Candidates for the establishment of polarity for example, may include members of the WUSCHEL-RELATED HOMEOBOX (WOX) family of transcription factors, which become expressed in precisely delimited and partly overlapping domains with the first few rounds of cell division after fertilization (Haecker et al., 2004; Fig. 1.1). In flower development, members of the MADS box transcription factor family become similarly expressed in partly overlapping domains to specify organ identity by combinatorial interactions (Egea-Cortines et al., 1999; Honma and Goto, 2001); it is tempting to speculate that an analogous combinatorial mechanism may be utilized in the early embryo.

My aim for this paper is to outline our current understanding of the molecular and genetic mechanisms establishing embryonic polarity in plants. Auxin is the most important signal for development after the globular stage, and for this reason I begin by describing auxin perception

and signal transduction in plants. Then, taking a backwards approach, I describe how the shoot and the root meristems are organized, when they are initiated in the embryo, and how shoot or root fates are enforced; before moving to discuss embryonic polarity, the process of how the apical-basal axis is set up and the first spatial domains established. I conclude this review by discussing the possible role of the GATA transcription factor HANABA TARANU (HAN) in early embryogenesis, the focus of my own research. HAN regulation is required to position the early expression domains of several genes involved in regulating the root apical meristem and distal root cap, and thus provides an excellent entry point into further elucidating early patterning processes.

Auxin transport, perception and signal transduction

Auxin is a systemic signal integrating plant growth responses like apical dominance and lateral branching (Thimann and Skoog, 1933; Hayward et al., 2009; De Smet et al., 2010), and controlling root and vascular cylinder formation (Mattsson et al., 1999; Cheng et al., 2007). The best understood mechanism for auxin biosynthesis requires the conversion of tryptophan to indole (Reviewed in Mano and Nemoto, 2012), and auxin production is thought to occur in all plant organs (Ljung et al., 2001; Cheng et al., 2006, 2007; Stepanova et al. 2008; Tao et al., 2008). The effect of auxin on developmental processes requires the creation of a graded differential of the hormone, with responses often strongest in localized areas of auxin accumulation, called auxin maxima (Sabatini et al., 1999; Friml et al., 2002; Dubrovsky et al., 2008). Auxin is protonated and uncharged at the low pH of the extracellular matrix and can freely enter the cell, but is becomes charged and trapped in the cytosol; thus, auxin transport requires active efflux, providing individual cells with an opportunity to make decisions about

preferential direction of export (Rubery and Sheldrake, 1974; Raven, 1975). Mathematical models for transport in the shoot (De Reuille et al., 2006) and the root meristems (Grieneisen et al., 2007) suggest that transport rather than localized auxin production likely drives the creation of auxin maxima.

Directional (or polar) auxin transport requires members of the PIN (named for PIN-FORMED 1) family of efflux transporters (Křeček et al., 2009). These are transmembrane proteins, and their subcellular localization determines the direction of auxin transport (Friml et al., 2004; Petrásek et al. 2006; Wiśniewska et al. 2006). PIN proteins are not stably inserted into the plasma membrane, but instead dynamically cycle between the plasma membrane and intracellular compartments (Feraru and Friml, 2008). The endocytosis of PIN proteins is clathrin-dependent (Dhonukshe et al., 2007), and their recycling from endosomes to the plasma membrane requires the guanidine nucleotide exchange factor GNOM (GN) (Mayer et al., 1993; Steinmann et al., 1999; Geldner et al., 2003). A prevalent theory states that polar auxin transport is under positive feedback control, enabling "canalization" of flux by self-organization (Sachs, 1981, 1989). In this model, cells exposed to auxin increase their transport capacity and tend to form connected conduits or 'canals' for the hormone through plant tissues; these conduits eventually differentiate into the vasculature (Sachs, 2003). Support for this idea comes from studies showing that auxin directly induces the transcription of many PIN transcripts (Vieten et al. 2005), represses PIN internalization (Paciorek et al. 2005), and can influence the subcellular polarity of PIN proteins (Sauer et al. 2006), resulting in altered auxin flux.

PIN localization is also controlled by feedback-independent mechanisms involving the serine-threonine kinase PINOID (PID), and the PROTEIN PHOSPHATASE 2A (PP2A) (Bennett et al., 1995; Benjamins et al., 2001). These factors control the phosphorylation status of

PIN proteins in embryos and roots; phosphorylated PINs are often targeted to the apical plasma membrane, whereas non-phosphorylated PIN proteins are targeted to the basal plasma membrane (Friml et al., 2004; Michniewicz et al., 2007). Polar auxin transport has been observed in many vascular plant species (Schrader et al., 2003; Balzan et al., 2014) and even certain mosses (Galván-Ampudia and Offringa, 2007; Fujita et al., 2008; Fujita and Hasebe, 2009; Lau et al., 2009). Additionally, there is evidence that the green alga *Chara coralline* can carry out polar auxin transport (Boot et al., 2012). Thus, the emerging picture is that the directional transport of auxin may be an ancient characteristic; predating the divergence of land plants from a green algal-like ancestor.

Auxin accumulation is interpreted at the level of individual cells by a pathway that directly influences transcription. Auxin interacts with members of the TRANSPORT INHIBITOR RESISTANT/AUXIN F-BOX (TIR/AFB) family of proteins, a small subclade of six genes in the *Arabidopsis* F-box gene family, which are a subunit of the SKP1–CULLIN1–F-BOX $(SCF)^{TIR1/AFB}$ ubiquitin ligase complex (Dharmasiri et al., 2005a, b; Kepinski et al., 2005; Tan et al., 2007). The binding of auxin to TIR/AFP proteins enhances their ability to interact with members of the AUX/IAA family of corepressors and targets them for ubiquitin-mediated degradation (Gray et al., 2001; Kepinski and Leyser, 2005; Tan et al., 2007; Maraschin et al., 2009). Aux/IAA proteins bind transcriptional regulators called AUXIN RESPONSE FACTORS (ARFs); which when freed by the auxin-dependent degradation of Aux/IAAs induce transcriptional changes (Ulmasov et al. 1999; Tiwari et al. 2003). ARF proteins have a B3-type DNA-binding domain at their N-terminus, facilitating their specific recognition of TGTCTC auxin response elements (AuxRE) in the promoters of auxin response genes (Ulmasov et al., 1999).

Arabidopsis contains 29 Aux/IAA proteins and 23 ARF proteins (Lokerse and Weijers, 2009). These factors are found in all clades of land plants (Remington et al., 2004; Finet et al., 2012); their diverse expression patterns (Rademacher et al., 2011, 2012), and the presence of variable structural elements (Lokerse and Weijers, 2009) suggests they have conserved as well as unique or at least quantitatively distinct functions. Aux/IAAs typically contain a repression domain containing an LxLxL EAR motif (Tiwari et al., 2004; Szemenyei et al., 2008), a domain containing a 'degron' sequence which facilities auxin-mediated binding to AFBs (Tan et al., 2007), and a region that is thought to facilitate interactions with other Aux/IAA and/or ARF proteins (Lokerse and Weijers, 2009); ARFs contain a domain that mediates interaction with Aux/IAA and/or other ARF proteins, a DNA binding domain, and a region that is believed to enhance DNA-binding affinity (Ulmasov et al., 1999; Boer et al., 2014). There is also structural diversity within these gene families. Certain members of the Aux/IAA family lack the degron sequence (Remington et al., 2004; Sato and Yamamoto, 2008), indicating that they cannot act in a canonical auxin-signaling pathway; and ARF proteins display a degree of variability in the central region of the protein, which may impact their effect on the transcriptional machinery (Tiwari et al., 2003). In protoplast assays, ARFs with a glutamine-rich middle region were found to activate a synthetic AuxRe reporter gene, while those with less pronounced Q enrichment repressed transcription (Tiwari et al., 2003); although these observations make for a useful rule of thumb, assessing the activity of ARF proteins in detail is likely to yield a more complex picture (Okushima et al., 2005; Lokerse et al., 2009).

Initiation and organization of the root apical meristem – auxin gradients and a radial pattern of tissues

Apical-to-basal auxin flux over the proembryo is initiated at the early globular stage. PIN7 assembles at the basal side of suspensor cells and PIN1 localizes to the basal membrane of cells in the lower tier of the proembryo (Friml et al., 2003). This arrangement of PIN proteins results in the formation of an auxin maximum in the hypophysis and sub hypophyseal cell (Friml et al., 2003; Fig. 1.2), and this signals the establishment of an apical-basal axis. Accordingly, embryos that are unable to synthesize, transport, or respond to auxin often block root formation and differentiation of the vascular cylinder (Reviewed in Möller and Weijers, 2009). Mutants lacking multiple PIN transporters display abnormal divisions in the basal embryo, are often rootless, and occasionally form ball-shaped seedlings without anatomical features of polarity (Friml et al., 2003). Similar phenotypes are observed for embryos treated with chemical inhibitors of polar auxin transport (Liu et al., 1993, Hadfi et al., 1998), or in mutants where auxin flux is redirected by reversing the subcellular localization of PIN1 from the basal to the apical side of proembryo cells (Friml et al., 2004; Michniewicz et al., 2007). Mutations in GNOM (GN), which inhibit the dynamic recycling of PIN proteins, typically result in rootless seedlings with short fused cotyledons and occasionally in spherical embryos lacking an interconnected vasculature (Mayer et al., 1993; Shevell et al. 1994, Busch et al. 1996). These *gn* embryos do express PIN1, yet PIN1 localization is apparently random and not aligned with the embryonic axis (Steinmann et al., 1999), indicating that axis formation at the globular stage is disrupted.

The best characterized ARF/IAA pair in embryonic development is MONOPTEROS (MP, also called ARF5) and BODENLOS (BDL, also called IAA12). Loss of MP, or dominant, auxin-resistant variants of BDL cause a failure to specify the hypophysis or vascular precursors

of the hypocotyl, and produce rootless seedlings with fused cotyledons (Berleth and Jürgens, 1993; Hardtke and Berleth, 1998; Hamann et al., 1999). MP is required for the auxin induced transcription of ATHB8, a member of the homoedomain-leucine zipper family III (HD-ZIP III) family (Baima et al., 1995; Mattsson et al., 2003) involved in specifying vascular fate. The role of ATHB8 in vascularization of leaves is well established: ATHB8 expression in the leaf provasculature is substantially lost in *mp* mutants (Donner et al., 2009); DNA-binding studies have found that MP directly interacts with the promoter of ATHB8 in seedlings; and mutational analyses confirms that ATHB8 functions in leaf vein pattern formation downstream of MP (Donner et al., 2009). The MP and ATHB8 expression domains overlap in central provascular cells at the heart stage, yet it is unclear whether ATHB8 plays a similar role promoting vascularization of the embryonic hypocotyl in response to MP (Hardtke and Berleth, 1998; Prigge et al., 2005).

In addition, MP directly targets two basic helix–loop–helix (bHLH) transcription factors, TARGET OF MONOPTEROS 5 (TMO5) and TMO7 at the early globular stage (Schlereth et al., 2010). TMO5 and its closest homolog TMO5-LIKE 1 (T5L1) are both expressed in the central cylinder where they are redundantly required for the first periclinal division of vascular precursors, and during post-embryonic development, growth of the vascular bundle (De Rybel et al., 2013). TMO7 is initially expressed like TMO5, in the cells of lower tier, but TMO7 protein travels, presumably through plasmodesmata, from the lower tier of the proembryo to the hypophysis, where it is required for an asymmetric division producing the lens-shaped progenitor cell of the root quiescent center (QC) (Schlereth et al., 2010). The quiescent center is thought to organize the surrounding cells into initials of the root apical meristem, such that the division of the hypophysis at the late globular stage is commonly interpreted as initiation of the root. As the

loss of TMO7 does not result in rootless seedlings, and TMO7 is only partially able to suppress the rootless phenotype of a weak *mp* allele, it is likely that other signals are required for root formation (Schlereth et al., 2010). A systematic survey of Aux/IAA and ARF family expression patterns identified IAA10 and ARF9 proteins as likely candidates for mediating this auxin response in the hypophysis and suspensor (Rademacher et al., 2011, 2012), suggesting the possibility that this process may involve distinct auxin responses in the cells of the suspensor and proembryo.

Auxin maxima can be visualized using the synthetic DR5 promoter consisting of nine concatenated auxin response elements (AuxREs) similar to a consensus found in an auxin-responsive promoter of soybean (Ulmasov et al., 1997), and driving the expression of β-glucuronidase (Ulmasov et al., 1997), luciferase (Jones et al., 2009), or fluorescent protein (Friml et al., 2003). DR5 expression is dose-responsive between 10–1000 nM IAA (Nakamura et al. 2003), and is a reliable marker for many processes known to require auxin (Ljung et al., 2001). As DR5 expression is dependent on the auxin signaling components present in a tissue or cell of interest, its sensitivity and specificity is likely to vary. Perhaps this variability accounts for discrepancies between computational models of auxin accumulation patterns and DR5 expression in the root tip (Grieneisen et al., 2007) and in root hair cells (Jones et al., 2009). Recent studies assessing the strength of ARF binding sites using a protein microarray found that the AuxREs in DR5 is not a high-affinity binding site (Boer et al., 2014). Replacing the nine AuxREs (TGTCTC) of DR5 with higher affinity sequences (TGTCGG) created a reporter that was more sensitive to the exogenous application of auxin, and broadened its expression domain during later stages of embryogenesis and in the root (Liao et al., 2015). Interestingly, this more sensitive auxin reporter, called DR5v2 was expressed very similar to DR5 in the early embryo.

The ability of auxin to induce cell fates in a concentration-dependent manner (Vuylsteker et al., 1998; Reinhardt et al., 2003; Dubrovsky et al., 2008) has prompted the idea that auxin might function as a morphogen in plant development. The concept of a morphogen is derived from a classic idea that multicellular organisms are patterned by gradients of form-providing substances (Boveri 1901; Hörstadius 1935; Spemann and Mangold 1924), later formalized in influential reaction-diffusion models (Turing, 1952; Wolpert, 1969). A morphogen is defined as a substance that forms a concentration gradient and specifies three or more positional values along this gradient in a dose-dependent manner. The Drosophila Bicoid protein, produced from RNA maternally deposited at the apical pole of the egg, is perhaps the best-known example for a morphogen in animal development (Ephrussi et al., 2004) and can specify the formation of distinct segments along the proximal-distal axis. As auxin has been shown to specify three cell fates along the axis of the root in a concentration-dependent manner, it fulfills the criteria of a morphogen (Sabatini et al., 1999; Benková et al., 2003; Blilou et al., 2005). Unlike a morphogen however, auxin gradients are formed primarily by active transport (Friml et al., 2002; Křeček et al., 2009). This difference, and the fact that auxin often induces simple cell fate changes in complex, reiterative patterns in time and space, have led to the proposal that auxin may better be described as a morphogenetic trigger (Benková et al., 2009). Both models seem to provide valid frameworks for interpreting the action of auxin.

In the root meristem, cell fates are regulated by auxin in a concentration-dependent manner through members of the PLETHORA (PLT) family of AP-2 transcription factors (Aida et al., 1999; Galinha et al., 2007). Auxin concentration as well as the accumulation of PLT proteins form similar gradients along the root, with highest values at the tip. PLT transcription in the root is dependent on auxin and requires MP, suggesting that the PLT protein gradient is a

read-out of the auxin gradient (Galhina et al., 2007; Mähönen et al., 2014). High levels of PLT activity promote quiescent center (QC) fate, while lower levels are needed for cell proliferation in the proximal domain, and cell expansion even further proximal (Galinha et al., 2007; Mähönen et al., 2014). Ectopic expression of PLT1 or PLT2 can drive the formation of an ectopic root, indicating that PLT genes are master regulators of root fate (Aida et al., 2004; Galinha et al., 2007; Smith and Long, 2010). PLT proteins also function downstream of MP during root initiation in the embryo, and *plt* multiple mutants have a similar phenotype as *mp* mutants (Aida et al., 2004).

In addition to auxin/MP/PLT function, the QC fates are dependent on the GRAS-family transcription factors, SCARECROW (SCR) and SHORTROOT (SHR). SHR is transcribed in the central provascular cylinder during the early globular stage (Helariutta et al., 2000), while SCR is expressed in the hypophysis and endodermis (Wysocka-Diller et al., 2000). Loss of SCR or SHR does not interfere with QC specification in the embryo, but organizing center fates cannot be maintained. The two genes also regulate an asymmetric division in the meristem that gives rise to the cortex and endodermis, the two layers of the ground tissue in the root, and are required for specifying endodermal fate (Benfey et al., 1993, Scheres et al. 1995). Mutants have only a single layer of ground tissue, which adopts cortex fate in *shr* mutants (Helariutta et al. 2000), and has a hybrid character in scr mutants (Di Laurenzio et al., 1996). The transcription of SCR requires SHR and in the mature root, is initiated by the direct movement of SHR protein from provascular cylinder cells to the surrounding cells of the ground tissue and QC (Helariutta et al. 2000; Nakajima et al. 2001). Overexpressing PLT results in the formation of ectopic QC cells in regions overlapping with SHR and SCR expression (Aida et al. 2004). Similarly, ectopically expressing SCR, or manipulating the basal auxin perception maximum results in an

ectopic or shifted QC which forms where auxin maxima and the expression of these genes overlap (Sabatini et al. 1999). According to current models, the precise position of the stem cell niche is specified by PLT genes, which determine the apical-basal coordinates, and SCR and SHR which determine its position along the radial axis (Aida et al., 2004).

As the expression of SHR and SCR is not influenced by auxin (Aida et al. 2004), it is likely that these two pathways function independently. While the targets of these pathways are unknown, maintenance of the root QC also requires the transcription factor WOX5 (Sarkar et al., 2007). WOX5 is expressed in the organizing cells of the root meristem (Haecker et al., 2004; Sarkar et al., 2007; Stahl et al., 2009; its close homologue, WUSCHEL (WUS) is expressed in the organizing cells of the shoot meristem; see below). Loss of *WOX5* results in premature differentiation of the distal columella stem cells, though root growth and meristem size are not dramatically affected (Sarkar et al., 2007). WOX5 functions in a negative feedback loop with CLAVATA3/ESR-RELATED 40 (CLE40), which counteracts WOX5-mediated stem cell-promoting effects to enable distal stem cell differentiation (Reviewed in Stahl and Simon, 2010; Fig. 1.3). WOX5 is not expressed in *mp* and *scr* mutants (Haecker et al., 2004; Sarkar et al., 2007) and can activate PLT expression in the root meristem (Ding and Friml, 2010). Intriguingly, the expression domain of WOX5 is slightly expanded in *plt* mutants (Sarkar et al., 2007), suggesting that WOX5 is indirectly responsible for restricting its own expression. WOX5 expression in roots is also abolished by treating seedlings with auxin (Ding and Friml, 2010).

Initiation and organization of the shoot apical meristem – phyllotaxis, primordia boundaries, and lateral organ polarity

Shortly following the division of the hypophysis and the initiation of a root at the late globular stage, PIN1 transporters at the flanks of the proembryo become re-localized to the apical side of cells (Benková et al., 2003). This shift funnels auxin toward the incipient cotyledon primordia, and in subsequent development auxin maxima are detected at the tips of developing cotyledons, the central provasculature, and the root tip (Friml et al., 2003; Fig. 1.2). PIN1 localization at the mature shoot meristem is similar to the embryonic apex, with transporters predominantly expressed in the outermost or L1 layer and directing auxin toward specific points in the peripheral zone of the meristem, marking the sites of incipient leaf primordia (Reinhardt et al., 2003; Smith et al. 2006). The regular arrangement of localized auxin sinks is responsible for the arrangement of organ primordia and, eventually, of leaves and flowers around the stem, referred to as phyllotaxis. Loss of PIN1 results in seedlings with a needle-like inflorescence stem, unable to initiate lateral organs (Okada et al., 1991; Gälweiler et al., 1998) and similar to the naked shoots of *mp* or *pid* mutants (Przemeck et al., 1996). Phyllotaxis is more specifically affected in mutants for the auxin importer AUXIN RESISTANT 1 (AUX1), and closely related, LIKE AUX1 (LAX) (Bennett et al. 1996); single mutants produce fertile plants with a disorganized pattern of primordia initiation (Bainbridge et al. 2008). Double and triple *aux1 lax* mutants form primordia in close proximity to one another, and this is accompanied by the loss of defined DR5 signal peaks in the peripheral zone of the shoot apical meristem and the altered spacing of PIN1 proteins (Bainbridge et al., 2008).

Members of the PLT family also play a role in regulating phyllotaxis (Pinon et al., 2013). PLT3, PLT5, and PLT7 are expressed in overlapping domains in the shoot apical meristem and

in organ primordia at different stages of development (Prasad et al., 2011; Nole-Wilson et al., 2005). The concomitant loss of these three genes results in plants unable to establish a normal spiral phyllotaxis (primordia divergence angle of 137.5°), and instead adopt a distichous pattern (primordia divergence angle of 180°) (Kuhlemeier 2007; Prasad et al., 2011; Pinon et al., 2013). By complementing the *plt3 plt5 plt7* phenotype with PLT genes ectopically expressed in either the shoot meristem or organ primordia, Pinon et al. (2013) found that PLT expression in the shoot apical meristem is required for spiral phyllotaxis; but not expression in the organ primordia. PLT activity is required for PIN1 up-regulation at the periphery of incipient primordia (Prasad et al., 2011). Interestingly, *plt3 plt5 plt7* phyllotactic defects are not caused by a reduction in PIN1 transcription, as ectopically expressing PIN1 in the meristem cannot restore spiral phyllotaxis in triple mutants (Pinon et al., 2013).

The separation of cotyledons and lateral organ formation is controlled by antagonistic interactions between transcription factors expressed in developing primordia and the shoot apical meristem. The NAC domain transcription factors CUP-SHAPED COTYLEDON 1 (CUC1) and CUC2 are expressed directly adjacent to the auxin perception maxima at cotyledon primordia in a narrow strip between them (Aida et al., 1999; Takada et al., 2001). These genes activate the expression of the class I KNOX transcription factor SHOOTMERISTEMLESS (STM) in the incipient shoot meristem (Long et al., 1996; Takada et al., 2001), and this coincides with the restriction of the CUC1/2 expression domain to the cotyledon margins (Barton and Poethig, 1993; Aida et al., 1997, 1999). The loss of *cuc1* and *cuc2* results in seedlings with severely fused cotyledons, forming a collar around the shoot apex, and lacking a shoot apical meristem (Aida et al., 1997, 1999). Mutants for *mp*, *pin1*, or *pin1 pid* display similar phenotypes, and misexpress CUC genes; CUC2 expression is lost from cotyledon boundaries, while CUC1 is

ectopically expressed in coyledons (Aida et al., 2002; Furutani et al., 2004). The misexpression of CUC1 is at least partially responsible for the cotyledon defect of *pin1 pid*, as mutating *cuc1* partially restores cotyledon development (Furutani et al., 2004). Thus, auxin-mediated CUC expression establishes bilateral symmetry in Arabidopsis.

Whereas CUC genes are required for separating primordia within the meristem, antagonistic interactions between the KNOTTED (KN) class of homeodomain proteins and the MYB-type transcription factor ASYMMETRIC LEAVES 1 (AS1) regulates stem cell/central vs. differentiation/primordia fate (Byrne et al., 2000). AS1 is one of the earliest molecular markers of lateral organ primordia, and *as1* mutants produce leaves with prominent outgrowths, or lobes from the proximal to distal end (Byrne et al., 2000). AS1 expression in the center of the meristem is restricted by STM; in *stm* mutants AS1 expression expands into the shoot meristem resulting in premature differentiation (Barton and Poethig, 1993; Clark et al., 1996; Byrne et al., 2000). STM expression is unchanged in *as1* mutants, yet another KNOX gene expressed in shoot meristem, KNAT1, becomes ectopically expressed at the base of *as1* leaves, likely causing the local outgrowth observed in mutants (Byrne et al., 2000). While KNAT1 does not affect development of the shoot meristem; mutants display reduced growth in floral internodes, pedicels and the style during reproduction (Douglas et al., 2002; Venglat et al., 2002), KNAT1 can substitute for *STM* in *stm* mutants (Chuck et al., 1996; Byrne et al., 2000, 2002; Scofield et al., 2008). Thus, antagonism between AS1 in organ primordia and KNOX genes expressed in the meristem provides a mechanism for differentiating between stem cells and organ founder cells within the shoot apical meristem.

The size of the stem cell population in the shoot apical meristem is controlled by WUS and CLV genes. WUS encodes a homeodomain transcription factor expressed in the non-

dividing organizing center of the shoot apical meristem (Laux et al., 1996; Mayer et al., 1998; the organizing center of the shoot serves analogous role to the QC of the root). WUS is expressed in the early embryo, prior to the expression of STM, yet *STM* is required to maintain WUS expression (Scofield et al., 2014). The loss of WUS does not affect the initiation of the shoot apical meristem, though the meristem prematurely differentiates after producing just a few organs (Laux et al. 1996). Similar to WOX5 and CLE40 in the root, WUS expression, and thus the size of the meristem, is regulated by negative feedback loop involving a ligand–receptor signaling cascade known as the CLAVATA (CLV) pathway (Reviewed in Stahl and Simon, 2010; Fig. 1.4). WUS protein moves from the organizing center to stem cells at the shoot apex where it directly activates the transcription of the CLV3 signaling peptide (Yadav et al., 2011). CLV3 peptide, secreted in response, can diffuse back to the organizing center, where it triggers the activity of CLV1 surface receptors to represses WUS transcription (Clark et al., 1993, 1995; Kayes et al., 1998; Brand et al., 2000; Schoof et al., 2000; Yadav et al., 2011; Lenhard et al., 2003). Restricting WUS movement results in the premature differentiation of stem cells (Yadav et al., 2011), while the loss of CLV1 or CLV3 causes the shoot apical meristem to expand dramatically (Clark et al., 1995; Fletcher et al., 1999). Interestingly, WUS and WOX5 are functionally interchangeable in shoot and root meristems (Sarkar et al., 2007), and it was found that CLE40 can substitute for CLV3 in the shoot (Sarkar et al., 2007; Hobe et al., 2003). Thus, an evolutionary link exists between shoot and root stem cell regulation.

Antagonistic interactions enforce root and shoot fates

The fundamental differences between shoot and root systems are not only morphologically apparent but also reflected in transcription profiles. Microarray analyses using transcripts

derived from the incipient shoot and root of developing embryos found that these domains become clearly distinct as early as the globular stage (Spencer et al., 2007). The temperature-sensitive, dominant-negative mutant topless-1 suggests that these transcriptional differences are likely enforced by epigenetic mechanisms. At the permissive temperature, *tpl-1* embryos begin to deviate from wild type at the globular stage and over time often form a second root at their apex, resulting in a mirror-image duplication phenotype at the seedling stage (Long et al., 2002). TOPLESS (TPL) encodes a transcriptional co-repressor and functions by recruiting histone deacetylases (Long et al., 2006; Zhu et al., 2010; Krogan et al., 2012) as well as disrupting components of the mediator complex (Malavé and Dent, 2006; Gonzalez et al., 2007). Interestingly, the double root phenotype, requires PLT gene expression. PLT1 and PLT2 are ectopically expressed at the shoot pole of *tpl-1* mutants, and their concomitant loss in a *tpl-1* mutant background suppresses apical root formation (Smith and Long, 2010). Analysis of the *tpl-1* phenotype would suggest that key differences between the shoot and root become fixed at the globular stage, as apical markers like WUS and STM are lost (Long et al., 2002, 2006) and basal markers like PLT and SCR (Smith and Long, 2010) are gained at the apical pole of mutants.

The formation of double roots in *tpl-1* is intriguing because it seems to be occurring at least partially independent of auxin signaling. Formation of the apical root is unaffected by loss of MP, and it remains unclear whether an ectopic auxin maximum initially forms at the apical pole of these mutants (Long et al., 2002). TPL and a small family of TPL-related (TPR) proteins do not bind DNA, and must interact with other transcription factors to recognize their target genes. The repressor of MP, BODENLOS (BDL), for example, interacts with TPL through its EAR motif, providing a molecular mechanism by which the transcription of MP-dependent genes are repressed in the absence of auxin (Szemenyei et al., 2008). While the contribution of

auxin signaling to the double root phenotype of *tpl-1* mutants remains unclear, factors functioning antagonistically to PLT genes in the shoot have recently been identified. Dominant alleles of HD-ZIP III genes, which produce mRNA that is resistant to microRNA165/166-mediated degradation, result in over-accumulation of HD-ZIP III transcripts at the apical pole of developing embryos and can suppress the *tpl-1* double root phenotype (Smith and Long, 2010). Furthermore, forced expression of miRNA-resistant forms of HD-ZIP III genes at the base of the embryo using the PLT promoter transforms the root pole into a second shoot apical meristem, leading to double shoot seedlings (Smith and Long, 2010). These results cast HD-ZIP III and PLT transcription factors as master regulators of shoot and root fate, respectively, and reveal a molecular antagonism between the pathways specifying apical and basal cell fates (Fig. 1.5a).

Embryonic polarity – how are root and shoot fates first established?

PIN localization prior to the establishment of apical-to-basal auxin transport at the 32-cell stage suggests that auxin accumulates in the apical cell and its descendants. Following the division of the zygote, PIN7 becomes localized to the apical membrane of the basal cell and its daughters, and DR5 is weakly expressed in the apical cell lineage (Friml et al., 2003; Fig. 1.2). However, loss of multiple PIN genes or GN, and treatment with auxin transport inhibitors have relatively minor effects prior to the 32-cell stage, such that is currently not clear whether auxin acts as an instructive signal in the apical cell. Embryonic polarity can be traced back to the asymmetric division of the zygote. This division requires a mitogen-activated protein (MAP) kinase signaling pathway, named for its MAPKK kinase gene, YDA (Lukowitz et al., 2004) which is triggered by the receptor-like cytoplasmic kinase SHORT SUSPENSOR (SSP) (Bayer et al., 2009) or signaling peptides expressed in the endosperm (Costa et al., 2014). Suppressing the

pathway by knocking out yda or the functionally equivalent MAP kinases *mpk3* and *mpk6*, block zygote elongation, such that the first division results in an abnormally small basal cell, which typically fails to form a recognizable suspensor (Lukowitz et al., 2004; Wang et al., 2007). It was also found that by constitutively activating YDA, longer suspensors with additional cells form, and in some cases the zygote merely forms a file of cells, suggesting that the apical cell has arrested or adopted a suspensor fate (Lukowitz et al., 2004). Similarly, loss of the RWP-RK factor *grounded* (*grd*), generates embryos with severely reduced suspensors that often fail to express markers for the incipient QC or the basal cell lineage, though like *yda* and *mpk* mutants eventually form seedlings (Lukowitz et al., 2004; Wang et al., 2007; Jeong et al., 2011).

The zinc finger transcription factor WRKY2 is expressed in the zygote, where it facilitates the polar re-distribution of organelles following cell division (Ueda et al., 2011). Loss of *wrky2* causes zygotes to divide equally, rather than asymmetrically, and subsequently often leads to aberrant divisions in the upper-most portion of the suspensor (Ueda et al., 2011). These defects affect patterns of WOX gene expression in the basal proembryo and suspensor. The expression of WOX8, normally present throughout the suspensor and required for normal development, is suppressed in the upper portion in *wrky2* mutants; while WOX9 expression, normally detected in the lower tier of globular stage embryos, expands basally to replace WOX8 in those cells (Ueda et al., 2011). WRKY2 was first identified as a direct activator of WOX8 (Ueda et al., 2011). Members of the MAP kinase signaling pathway also interact with WOX genes. Combining mutations in *yda* or *grd*, with *wox8* and *wox9* causes many embryos to arrest as short zygotes or after the first division, with daughter cells of virtually identical size (Breuninger et al., 2008; Jeong et al., 2011). These observations imply that transcriptional

programs initiated after fertilization are essential for creating embryonic polarity in the first few cell divisions.

Members of the WOX family are differentially expressed in the apical and basal daughter cells of the zygote, and by the 8-cell stage mark four domain along the primary axis (Haecker et al., 2004). WOX2 transcripts are detected in the upper tier, WOX9 in the lower tier, WOX8 and WOX9 in the hypophysis, and WOX8 in the suspensor (Haecker et al., 2004; Fig. 1.1). Mutations in wox8 and wox9 do not affect division of the zygote, but cause aberrant divisions in the suspensor, resulting in enlarged or misshapen cells; and in a small proportion of embryos, the apical cell generates a filamentous stalk of vacuolated cells and eventually arrests (Breuninger et al., 2008). Double mutants fail to establish an auxin gradient, and markers for the proembryo like PIN1 and WOX2 are often undetectable (Breuninger et al., 2008). Even though WRKY2 activates WOX8 expression; the early phenotypes of *wrky2* and *wox8 wox9* are distinct (Ueda et al., 2011). This may indicate that WOX8 expression is at least to some extent independent of WRKY2, a likely scenario given that WOX8 is weakly expressed in *wrky2* zygotes. The promoter of WOX8 contains several cis-regulatory elements required for its expression, only one of which is activated by WRKY2 (Ueda et al., 2011), which implies that other factors may activate WOX8 in the zygote.

The loss of *wox2* causes relatively mild defects in the apical cell lineage. Mutants display aberrant divisions at the apex of globular embryos, and seedlings occasionally form only a single cotyledon (Breuninger et al., 2008). These defects are enhanced by combining *wox2* with mutations in other WOX genes expressed in the proembryo and hypophysis. Combining *wox2* with mutations in *wox1*, which is expressed in the provascular cells of cotyledons; *wox3/pressed flower* (*prs*), which is expressed at the margins of cotyledon primordia; and *wox5*, which is

expressed in the hypophysis, and later the QC, results in rod-like or shootless seedlings (Haecker et al., 2004; Breuninger et al., 2008). As this effect was not observed in *wox1 wox3 wox5* mutants – which form normal seedlings – it is likely that WOX2 is the primary factor controlling shoot fate in the group. Interestingly, combining *wox1 wox2 wox3* mutants with *wox8* markedly increases the number of seedlings with shoot defects (Breuninger et al., 2008). This suggests that WOX genes play overlapping roles in regulating shoot and root development.

In the early embryo, WOX genes appear to be controlling early fate decisions independently of auxin signaling. Auxin has a minor effect on the expression of WOX9; which no longer shifts from the uppermost suspensor cell to the lower tier in *mp* mutants (Haecker at al., 2004), though this was observed at the heart stage, when the primary axis was already established. Similarly, *mp wox2 wox8* mutants display abnormal divisions in the protoderm only slightly more frequently than *wox2 wox8* (Breuninger et al., 2008); and the occasional arrest of *mp wox8* embryos as ball-shaped structures is difficult to interpret, as this was not observed in *mp wox2 wox8* mutants (Breuninger et al., 2008). Intriguing, there is evidence that the establishment of an auxin prepattern – the spatial layout of auxin biosynthesis, transport, and interpretation machinery – requires the action of WOX genes in the early embryo. PIN1 expression is absent in *wox8 wox9* mutants, while the transcription of DR5 expands throughout the proembryo (Breuninger et al., 2008).

The distinct, sometimes overlapping expression domains of WOX genes are reminiscent of the genetic network controlling floral organ identity (Krizek and Fletcher, 2005). Nearly all of these factors are MADS box transcription factors, and their combined function specifies the whorl identity of floral organs in plants (Egea-Cortines et al., 1999; Honma and Goto, 2001). Similar kinds of interactions might occur between WOX genes in the early embryo, and may

provide a bridge between these factors and other families of regional regulators. Such a network might include the AP2-type transcription factors DORNRÖSCHEN (DRN), which is transcribed in the apical cell following the first division, then throughout the embryo, until it becomes restricted to the apical domain during the late globular stage, and later in the tips of cotyledons (Cole et al., 2009). Loss of DRN or its close homolog DRN-LIKE (DRNL) results in fused cotyledons, similar to CUC mutant seedlings (Aida et al., 1997; Kirch et al., 2003; Chandler et al., 2007), while the loss of both genes results in embryos that fail to form cotyledons (Chandler et al., 2007). It is also conceivable that PLT genes might be directly regulated by WOX family members: PLT1 and PLT2 are both expressed in the lower tier of 8-cell stage embryos, then in the lens-shaped cell and provasculature, where they maintain the stem cell niche (Aida et al., 1999, 2004). Intriguingly, the early expression of both DRN and PLT genes occurs independently of MP, though the expression of both requires MP during later stages development (Aida et al., 2004; Chandler et al., 2007; Cole et al., 2009). WOX genes may activate them in the early embryo.

Establishing developmental domains in the pre-globular embryo

The radial organization of tissue types is first apparent following the 8-cell stage, when each cell of the proembryo divides tangentially along the apical-basal axis. These divisions produce the protoderm and ground tissue lineages, and they are controlled at least in part by WOX genes and MP. The loss of WOX2 results in embryos that do not always divide tangentially, and this phenotype is enhanced by the loss of additional WOX genes or MP (Breuninger et al., 2008). As *mp* mutants or embryos expressing an auxin-insensitive version of *bdl* form an epidermis (Berleth and Jürgens, 1993; Hamann et al., 1999), it is unlikely that auxin responses play a

dominant role in specifying this tissue. The separation of inner and outer cell fates coincides with the refinement of gene expression domains. The HD-ZIP IV transcription factors *ARABIDOPSIS THALIANA MERISTEM LAYER 1* (*ATML1*) and closely related *PROTODERMAL FACTOR 1* (*PDF1*) and PDF2, which are initially expressed broadly in the proembryo, become restriced to the protoderm layer (Sessions et al., 1999; Abe et al., 2001, 2003). The loss of *atml1* and *pdf2* causes most embryos to arrest at the globular stage, developing abnormally enlarged surface cells, and occasionally, seedlings without a recognizable epidermis develop as well (Abe et al., 2003; Ogawa et al., 2015). ATML1 expression is sufficient to specify protoderm identity: overexpressing ATML1 results in the formation of epidermis-specific traits, such as guard cells and root hairs in the inner tissues of leaves (Takada et al., 2013). Based on these findings, ATML1 is considered a master regulator for epidermal cell fate.

The GATA-type transcription factor HANABA TARANU (HAN) defines the position of gene expression domains between the suspensor and proembryo. The HAN gene is transcribed more broadly than WOX family members: it is expressed in the zygote, apical cell, and throughout the proembryo, before being restricted to the central cells of the lower tier and then the central provasculature (Zhao et al., 2004; Nawy et al., 2010); yet defects associated with losing HAN function predominantly affect the lower tier and the suspensor. In *han* mutants, many genes normally expressed in the hypophysis and sub-hypophyseal cells expand into the lower tier; conversely, genes normally expressed throughout the proembryo become detectable only in the upper tier (Nawy et al., 2010). These results seem to indicate that the boundary between the embryo and suspensor shifts apically in these mutants, a view supported by the formation of large, vacuolated cells in the lower tier that look similar to the vacuolated cells of

the suspensor (Nawy et al., 2010). Despite these striking abnormalities, *han* mutants often recover and develop into complete seedlings.

The initiation of a root in recovering *han* mutants is unlike root formation in wild-type embryos. Rather than initiated in the formative division of the hypophysis at the base of the proembryo, a root is assembled between the upper and lower tier of *han* embryos (Nawy et al., 2010). As this phenotype is dissimilar to other embryo mutants, it has been difficult to place HAN into any known pathways regulating root formation. One possibility is that the primary function of HAN might be to regulate auxin transport: both PIN1 and PIN7 are expressed in apically shifted domains at the 8-cell stage, PIN7 expanded into the lower tier and PIN1 restricted to the upper tier, and DR5 expression is found throughout the lower tier of *han* embryos instead of the hypophysis. In addition, root formation in *han* does not require MP (Nawy et al., 2010). However, it remains unclear whether HAN functions primarily to configure the auxin transport machinery in the early embryo of whether it has other primary targets.

Transcriptional regulation of polarity

One of the most intriguing results to emerge from investigations into HAN function is the synergistic genetic interaction between *han* and *angustifolia 3* (*an3*) mutations. Combining a weak *han* allele, in which a single leucine residue within a small conserved segment called the HAN domain is substituted with phenylalanine, and a null allele of *an3* results in seedlings that have malformed cotyledons and often an ectopic root at their shoot pole (Kanei et al., 2012). This effect is surprising because both mutations alone cause only rather subtle defects: *an3* is associated with the formation of slightly narrow leaves with fewer cells (Horiguchi et al., 2011; Kanei et al., 2012); *han-30* mutations have no visible defects on the embryo and cause rounder

leaves. Similar double root seedlings are generated by *tpl-1* mutants (Long et al., 2002; Long et al., 2006; Smith and Long, 2010), and HAN was previously shown to bind TPL in yeast (Causier et al., 2012), so it is conceivable that the *tpl-1* phenotype might stem at least in part from the loss of *han* function. The double root phenotype of *tpl-1* mutants is caused by the ectopic expression of PLT genes in the apical embryo (Smith and Long, 2010). PLT expression is similarly misexpressed in *han an3* (Kanei et al., 2012), but is unaffected in *han* single mutant embryos (Nawy et al., 2010); it remains unknown by which mechanism the double mutation affects PLT expression.

The *han* phenotype is distinct from other mutants affecting patterns of gene expression in the early embryo, and for this reason, a more generalized discussion of the establishment of polarity in plants may be instructive. Plant leaves display distinct polarity along their dorsoventral axis, and this is maintained at the laminar margins by the transcription factors, WOX1 and WOX3 (Nakata et al., 2012). These genes are expressed in the middle domain of the leaf, and when they are both mutated transcription factors normally expressed at both the adaxial and abaxial surface, become co-expressed in the laminar margin; resulting in reduced blade outgrowth and the coexistence of adaxial and abaxial-like cell types (Nakata et al., 2012; Fig. 1.5b). This effect mimics what we see in *han* mutants, the loss of HAN expression at the proembryo boundary results in a similar, albeit unidirectional, expansion of hypophysial and subhypophysial fates into the lower tier of the proembryo (Nawy et al., 2010).

The coordinated apical shift of expression domains observed in *han* embryos has been interpreted as a shift in the inductive boundary between suspensor and proembryo, across which the root is initiated, and is reminiscent of the effect seen with mutants of the lateral organ polarity pathway. Simultaneous loss of multiple members of the *KANADI* (*KAN*) gene family,

KANADI 1 (KAN1), KAN2, and KAN3, which are normally expressed on the abaxial side of cotyledons and leaf primordia, produces radially symmetric leaves with adaxial characteristics (Kerstetter et al. 2001; Eshed et al., 2004). This radialized phenotype is caused by the concomitant expansion of HD-ZIP III expression, normally confined to the adaxial leaf surface, into the domain previously occupied by KAN expression (Eshed et al. 2001). The opposite effect is observed with the loss of HD-ZIP III genes, in this case, radially symmetric leaf primordia with abaxial characteristics are formed (Emery et al. 2003; Prigge et al., 2005). Thus, the boundary between abaxial and adaxial domains of the developing leaf is maintained by antagonistic interaction between two families of transcription factors (Fig. 1.5b). Similarly, HAN may be positioning the boundary between suspensor and proembryo by antagonizing genes normally expressed in the suspensor.

Conclusions

Auxin is a pervasive signal for plant growth and development, yet it remains unclear how auxin signaling might regulate the earliest fate decisions during Arabidopsis embryogenesis. In recent years, a number of auxin-independent transcription factors have been reported to play a role in establishing embryonic polarity and the first developmental domains. This would seem to suggest the possibility that early development is driven by a transcriptional network that operates with minimal input from auxin. Important players seem to be the RWP-RK gene GRD and the zinc finger protein WRKY2 gene, which regulate the asymmetric division of the zygote; members of the WOX homeodomain family of proteins, which act in distinct regions along the apical basal axis; the homeodomain-leucine zipper proteins ML1 and PDF, which specify protoderm fate; and the GATA factor HAN, which regulates gene expression at the boundary

between suspensor and proembryo. Identification of these factors provides welcome opportunities for further studies, as our current understanding of the early patterning process is incomplete at best.

References

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

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. The Plant Cell 15, 63-78.

Aida, M., Beis, D., Heidstra, R., Willemsen, V., Blilou, I., Galinha, C., Nussaume, L., Noh, Y., Amasino, R., and Scheres, B. (2004). The *PLETHORA* genes mediate patterning of the *Arabidopsis* root stem cell niche. Cell 119, 109-120.

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M. (1997). Genes involved in the organ separation in *Arabidopsis*: and analysis of the *cup-shaped cotyledon* mutant. The Plant Cell 9, 841-857.

Aida, M., Ishida, T., and Tasaka, M. (1999). Shoot apical meristem and cotyledon formation during *Arabidopsis* embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. Development 126, 1563-1570.

Aida, M., Vernoux, T., Furutani, M., Traas, J., and Tasaka, M. (2002). Roles of *PIN-FORMED1* and *MONOPTEROS* in pattern formation of the apical region of the *Arabidopsis* embryo. Development 129, 3965-3974.

Baima, S., Nobili, F., Sessa, G., Lucchetti, S., Ruberti, I., and Morelli, G. (1995). The expression of *Athb-8* homeobox gene is restricted to provascular cells in *Arabidopsis thaliana*. Development 121, 4171-4182.

Baima, S., Possenti, M., Matteucci, A., Wisman, E., Altamura, M.M., Ruberti, I., and Morelli, G. (2001). The *Arabidopsis* ATHB-8 HD-Zip protein acts as a differentiation-promoting transcription factor of the vascular meristem. Plant Physiology 126, 643-655.

Bainbridge, K., Guyomarc'h, S., Bayer, E., Swarup, R., Bennett, M., Mandel, T., and Kuhlemeier, C. (2008). Auxin influx carriers stabilize phyllotactic patterning. Genes & Development 22, 810-823.

Balzan, S., Johal, G.S., and Carraro, N. (2014). The role of auxin transporters in monocots development. Frontiers in Plant Science 5, 1-12.

Barton, M.K., and Peothig, R.S. (1993). Formation of the shoot apical meristem in *Arabidopsis thaliana*: an analysis of development in the wild type and in the *shoot meristemless* mutant. Development 119, 823-831.

Bayer, M., Nawy, T., Giglione, C., Galli, M., Meinnel, T., and Lukowitz, W. (2009). Paternal control of embryonic patterning in *Arabidopsis thaliana*. Science 323, 1485-1488.

Benfey, P.N., Linstead, P.J., Roberts, K., Schiefelbein, J.W., Hauser, M., and Aeschbacher, R.A. (1993). Root development in *Arabidopsis*: four mutants with dramatically altered root morphogenesis. Development 119, 57-70.

Benjamins, R., Quint, A., Weijers, D., Hooykaas, P., and Offringa, R. (2001). The PINOID protein kinase regulates organ development in *Arabidopsis* by enhancing polar auxin transport. Development 128, 4057-4067.

Benková, E., Ivanchenko, M.G., Friml, J., Shishkova, S., and Dubrovsky, J.G. (2009). A morphogenic trigger: is there an emerging concept in plant developmental biology? Trends in Plant Science 14, 189-193.

Benková, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertová, D., Jürgens, G., and Friml, J. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115, 591-602.

Bennett, S.R.M., Alvarez, J., Bossinger, G., and Smyth, D.R. (1995). Morphogenesis in *pinoid* mutants of *Arabidopsis thaliana*. The Plant Journal 8, 505-520.

Bennett, M.J., Marchant, A., Green, H.G., May, S.T., Ward, S.P., Millner, P.A., Walker, A.R., Schulz, B., and Feldmann, K.A. (1996). Arabidopsis AUX1 gene: a permease-like regulator of root gravitropism. Science 273, 948-950.

Berleth, T., and Jürgens, G. (1993). The role, of the *monopteros* gene in organising the basal region of the *Arabidopsis* embryo. Development 118, 575-587.

Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K., and Scheres, B. (2005). The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots. Nature 433, 39-44.

Boer, D.R., Freire-Rios, A., van den Berg, W.A.M., Saaki, T., Manfield, I.W., Kepinski, S., López-Vidrieo, I., Franco-Zorilla, J.M., de Vries, S.C., Solano, R., Weijers, D., and Coll, M. (2014). Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors. Cell 156, 577-589.

Boot, K.J.M., Libbenga, K.R., Hille, S.C., Offringa, R., and van Duijn, B. (2012). Polar auxin transport: an early invention. Journal of Experimental Botany 63, 4213-4218.

Boveri, T. (1901). Die Polarität von oocyte, ei und Larve des *Strongylocentrotus lividus*. Zoologische Jahrbücher. Abteilung für Anatomie und Ontogenie der Tiere 14, 630-690.

Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M., and Simon, R. (2000). Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by *CLV3* activity. Science 289, 617-619.

Brenuinger, H., Rikirsch, E., Hermann, M., Ueda, M., and Laux, T. (2008). Differential expression of *WOX* genes mediates apical-basal axis formation in the *Arabidopsis* embryo. Developmental Cell 14, 867-876.

Busch, M., Mayer, U., and Jürgens, G. (1996). Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure and intragenic complementation. Molecular and General Genetics 250, 681-691.

Byrne, M.E., Barley, R., Curtis, M., Arroye, J.M., Dunham, M., Hudson, A., and Martienssen, R.A. (2000). *Asymmetric leaves1* mediates leaf patterning and stem cell function in *Arabidopsis*. Nature 408, 967-971.

Byrne, M.E., Simorowski, J., and Martienssen, R.A. (2002). *ASYMMETRIC LEAVES1* reveals *knox* gene redundancy in *Arabidopsis*. Development 129, 1957-1965.

Causier, B., Ashworth, M., Guo, W., and Davies, B. (2012). The TOPLESS interactome: A framework for gene repression in *Arabidopsis*. Plant Physiology 158, 423-438.

Cederholm, H., Iyer-Pascuzzi, A.S., and Benfey, P.N. (2012). Patterning the primary root in Arabidopsis. Wiley Interdisciplinary Reviews: Developmental Biology 1, 675-691.

Chandler, J.W., Cole, M., Flier, A., Grewe, B., and Werr, W. (2007). The AP2 transcription factors DORNRÖSCHEN and DORNRÖSCHEN-LIKE redundantly control *Arabidopsis* embryo patterning via interaction with PHAVOLUTA. Development 134, 1653-1662.

Cheng, Y., Dai, X., and Zhao, Y. (2006). Auxin biosynthesis by the YUCCA Flavin monooxygenases controls the formation of floral organs and vascular tissues in *Arabidopsis*. Genes & Development 20, 1790-1799.

Cheng, Y., Dai, X., and Zhao, Y. (2007). Auxin synthesized by the YUCCA flavin monooxygenases is essential for embryogenesis and leaf formation in *Arabidopsis*. The Plant Cell 19, 2430-2439.

Chuck, G., Lincoln, C., and Hake, S. (1996). *KNAT1* induces lobed leaves with ectopic meristems when overexpressed in *Arabidopsis*. The Plant Cell 8, 1277-1289.

Clark, S.E., Jacobsen, S.E., Levin, J.Z., and Meyerowitz, E.M. (1996). The *CLAVATA* and *SHOOT MERISTEMLESS* loci competitively regulate meristem activity in *Arabidopsis*. Development 122, 1567-1575.

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

Clark, S.E., Running, M.P., and Meyerowitz, E.M. (1995). *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*. Development 121, 2057-2067.

Clowes, F.A.L. (1961). Apical Meristems. Blackwell Scientific, Oxford.

Cole, M., Chandler, J., Weijers, D., Jacobs, B., Comelli, P., and Werr, W. (2009). *DORNRÖSCHEN* is a direct target of the auxin response factor MONOPTEROS in the *Arabidopsis* embryo. Development 136, 1643-1651.

Costa, L.M., Marshall, E., Tesfaye, M., Silverstein, K.A., Mori, M., Umetsu, Y., Otterbach, S.L., Papareddy, R., Dickinson, H.G., Boutiller, K., VandenBosch, K.A., Ohki, S., and Gutierrez-

Marcos, J.F. (2014). Central cell-derived peptides regulate early embryo patterning in flowering plants. Science 344, 168-172.

De Reuille, P.B., Bohn-Caurseau, I., Ljung, K., Morin, H., Carraro, N., Godin, C., and Traas, J. (2006). Computer simulations reveal properties of the cell-cell signaling network at the shoot apex in *Arabidopsis*. Proceedings of the National Academy of Sciences 103, 1627-1632.

De Rybel, B., Möller, B., Yoshida, S., Grabowicz, I., de Reuille, P.B., Boeren, S., Smith, R.S., Borst, J.W., and Weijers, D. (2013). A bHLH complex controls embryonic vascular tissue establishment and indeterminate growth in *Arabidopsis*. Developmental Cell 24, 426-437.

De Smet, I., Lau, S., Voß, U., Vanneste, S., Benjamins, R., Rademacher, E.H., Schlereth, A., De Rybel, B., Vassileva, V., Grunewald, W., Naudts, M., Levesque, M.P., Ehrismann, J.S., Inzé, D., Luschnig, C., Benfey, P.N., Weijers, D., Van Montagu, M.C.E., Bennett, M.J., Jürgens, G., and Beeckman, T. (2010). Bimodular auxin response controls organogenesis in *Arabidopsis*. Proceedings of the National Academy of Sciences 107, 2705-2710.

Dharmasiri, N., Dharmasiri, S., and Estelle, M. (2005). The F-box protein TIR1 is an auxin receptor. Nature 435, 441-445.

Dharmasiri, N., Dharmasiri, S., Weijers, D., Lechner, E., Yamada, M., Hobbie, L., Ehrismann, J.S., Jürgens, G., and Estelle, M. (2005). Plant development is regulated by a family of auxin receptor F box proteins. Developmental Cell 9, 109-119.

Dhonukshe, P., Aniento, F., Hwang, I., Robinson, D.G., Mravec, J., Stierhof, Y., and Friml, J. (2007). Clathrin-mediated constitutive endocytosis of PIN auxin efflux carriers in *Arabidopsis*. Current Biology 17, 520-527.

Di Laurenzio, L., Wysocka-Diller, J., Malamy, J.E., Pysh, L., Helariutta, Y., Freshour, G., Hahn, M.G., Feldmann, K.A., and Benfey, P.N. (1996). The SCARECROW gene regulates an

asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. Cell 86, 423-433.

Ding, Z., and Friml, J. (2010). Auxin regulates distal stem cell differentiation in Arabidopsis roots. Proceedings of the National Academy of Sciences 107, 12046-12051.

Dolan, L., Janmaat, K., Willemsen, V., Linstead, P., Poethig, S., Roberts, K., and Scheres, B. (1993). Cellular organisation of the *Arabidopsis thaliana* root. Development 119, 71-84.

Donner, T.J., Sherr, I., and Scarpella, E. (2009). Regulation of preprocambial cell state acquisition by auxin signaling in *Arabidopsis* leaves. Development 136, 3235-3246.

Douglas, S.J., Chuck, G., Dengler, R.E., Pelecanda, L., and Riggs, C.D. (2002). *KNAT1* and *ERECTA* regulate inflorescence architecture in *Arabidopsis*. The Plant Cell 14, 547-558.

Dubrovsky, J.G., Sauer, M., Napsuclaly-Mendivil, S., Ivanchenko, M.G., Friml, J., Shishkova, S., Celenza, J., and Benková, E. (2008). Auxin acts as a local morphogenic trigger to specify lateral root founder cells. Proceedings of the National Academy of Sciences 105, 8790-8794.

Egea-Cortines, M., Saedler, H., and Sommer, H. (1999). Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. The EMBO Journal 18, 5370-5379.

Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F., an Bowman, J.L. (2003). Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. Current Biology 13, 1768-1774.

Ephrussi, A., and St Johnston, D. (2004). Seeing is believing: The bicoid morphogen gradient matures. Cell 116, 143-152.

Esau, K. (1977). Anatomy of seed plants. John Wiley and Sons, New York.

Eshed, Y., Baum, S.F., Perea, J.V., and Bowman, J.L. (2001). Establishment of polarity in lateral organs of plants. Current Biology 11, 1251-1260.

Eshed, Y., Izhaki, A., Baum, S.F., Floyd, S.F., Floyd, S.K., and Bowman, J.L. (2004). Asymmetric leaf development and blade expansion in *Arabidopsis* are mediated by KANADI and YABBY activities. Development 131, 2997-3006.

Feraru, E., and Friml, J. (2008). PIN polar targeting. Plant Physiology 147, 1553-1559.

Finet, C., Berne-Dedieu, A., Scutt, C.P., and Marlétaz, F. (2012). Evolution of the ARF gene family in land plants: old domains, new tricks. Molecular Biology and Evolution 30, 45-56.

Fletcher, J.C., Brand, U., Running, M.P., Simon, R., and Meyerowitz, E.M. (1999). Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. Science 283, 1911-1914.

Friml, J., Benková, E., Blilou, I., Wisniewska, J., Hamann, T., Ljung, K., Woody, S., Sandberg, G., Scheres, B., Jürgens, G., and Palme, K. (2002). AtPIN4 mediates sink-driven auxin gradients and root patterning in *Arabidopsis*. Cell 108, 661-673.

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwartz, H., Hamann, T., Offringa, R., and Jürgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. Nature 426, 147-153.

Friml, J., Yang, X., Michniewicz, M., Weijers, D., Quint, A., Tietz, O., Benjamins, R., Ouwerkerk, P.B.F., Ljung, K., Sandberg, G., Hooykaas, P.J.J., Palme, K., and Offringa, R. (2004). A PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. Science 306, 862-865.

Fujita, T., and Hasebe, M. (2009). Convergences and divergences in polar auxin transport and shoot development in land plant evolution. Plant Signaling and Behavior 4, 313-315.

Fujita, T., Sakaguchi, H., Hiwatashi, Y., Wagstaff, S.J., Ito, M., Deguchi, H., Sato, T., and Hasebe, M. (2008). Convergent evolution of shoots in land plants: lack of auxin polar transport in moss shoots. Evolution & Development 10, 176-186.

Furutani, M., Vernoux, T., Traas, J., Kato, T., Tasaka, T., and Aida, M. (2004). *PIN-FORMED1* and *PINOID* regulate boundary formation and cotyledon development in *Arabidopsis* embryogenesis. Development 131, 5021-5030.

Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R., and Scheres, B. (2007). PLETHORA proteins as dose-dependent master regulators of *Arabidopsis* root development. Nature 449, 1053-1057.

Galván-Ampudia, C., and Offringa, R. (2007). Plant evolution: AGC kinases tell the auxin tale. Trends in Plant Science 12, 541-547.

Gälweiler, L., Guan, C., Müller, A., Wisman, E., Mendgen, K., Yephremov, A., and Palme, K. (1998). Regulation of polar auxin transport by AtPIN1 in *Arabidopsis* vascular tissue. Science 282, 2226-2230.

Geldner, N., Anders, N., Wolters, H., Keicher, J., Kornberger, W., Muller, P., Delbarre, A., Ueda, T., Nakano, A., and Jürgens, G. (2003). The *Arabidopsis* GNOM ARF-GEF mediates endosomal recycling, auxin transport, and auxin-dependent plant growth. Cell 112, 219-230.

Gonzalez, D., Bowen, A.J., Carroll, T.S., and Conlan, R.S. (2007). The transcription compressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription. Molecular and Cellular Biology 27, 5306-5315.

Gray, W.M., Kepinski, S., Rouse, D., Leyser, O., and Estelle, M. (2001). Auxin regulates SCF^{TIR1}-dependent degradation of AUX/IAA proteins. Nature 414, 271-276.

Grieneisen, V.A., Xu, J., Marée, A.F.M., Hogeweg, P., and Scheres, B. (2007). Auxin transport is sufficient to generate a maximum and gradient guiding root growth. Nature 449, 1008-1013.

Hadfi, K., Speth, V., and Neuhaus, G. (1998). Auxin-induced developmental patterns in *Brassica juncea* embryos. Development 125, 879-887.

Haecker, A., Groß-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrman, M., and Laux, T. (2004). Expression dynamics of *WOX* genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. Development 131, 657-668.

Hamann, T., Mayer, U., and Jürgens, G. (1999). The auxin-insensitive *bodenlos* mutation affects primary root formation and apical-basal patterning in the *Arabidopsis* embryo. Development 126, 1387-1395.

Hardtke, C.S., and Berleth, T. (1998). The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. The EMBO Journal 17, 1405-1411.

Hayward, A., Stirnberg, P., Beveridge, C., and Leyser, O. (2009). Interactions between auxin and strigolactone in shoot branching control. Plant Physiology 151, 400-412

Helariutta, Y., Fukaki, H., Wysocka-Diller, J., Nakajima, K., Jung, J., Sena, G., Hauser, M., and Benfey, P. (2000). The *SHORT-ROOT* gene controls radial patterning of the *Arabidopsis* root through radial signaling. Cell 101, 555-567.

Hobe, M., Müller, R., Grünewald, M., Brand, U., and Simon, R. (2003). Loss of CLE40, a protein functionally equivalent to the stem cell restricting signal CLV3, enhances root waving in Arabidopsis, Development Genes and Evolution 213, 371-381.

Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature 409, 525-529.

Horiguchi, G., Mollá-Morales, A., Pérez-Pérez, J.M., Kojima, K., Robles, P., Ponce, M.R., Micol, J.L., and Tsukaya, H. (2011). Differential contributions of ribosomal genes to *Arabidopsis thaliana* leaf development. The Plant Journal 65, 724-736.

Hörstadius, S. (1935). Über die Determination im Verlaufe der Eiachse beir Seeigeln. Pubblicazioni della Stazione Zoologica di Napoli 14, 251-479.

Jenik, P.D., Gillmor, C.S., and Lukowitz, W. (2007). Embryonic patterning in Arabidopsis thaliana. Annual Review of Cell and Developmental Biology 23, 207-236.

Jeong, S., Bayer, M., and Lukowitz, W. (2011). Taking the very first steps: from polarity to axial domains in the early *Arabidopsis* embryo. Journal of Experimental Botany 62, 1687-1697.

Jones, A.R., Kramer, E.M., Knox, K., Swarup, R., Bennett, M.J., Lazarus, C.M., Leyser, H.M.O., and Grierson, C.S. (2009). Auxin transport through non-hair cells sustains root hair development. Nature Cell Biology 11, 78-84.

Jürgens, G., and Mayer, U. (1994). Arabidopsis. In A Colour Atlas of Developing Embryos (ed. J. Bard). Wolfe, London.

Kanei, M., Horiguchi, G., and Tsukaya, H. (2012). Stable establishment of cotyledon identity during embryogenesis in *Arabidopsis* by *ANGUSTIFOLIA3* and *HANABA TARANU*. Development 139, 2436-2446.

Kayes, J.M., and Clark, S.E. (1998). *CLAVATA2*, a regulator of meristem and organ development in *Arabidopsis*. Development 125, 3843-3851.

Kepinski, S., and Leyser, O. (2005). The *Arabidopsis* F-box protein TIR1 is an auxin receptor. Nature 435, 446-451.

Kerstetter, R.A., Bollman, K., Taylor, R.A., Bomblies, K., and Peothig, R.S. (2001). KANADI regulates organ polarity in *Arabidopsis*. Nature 411, 706-709.

Kidner, C., Sundaresan, V., Roberts, K., and Dolan, L. (2000). Clonal analysis of the *Arabidopsis* root confirms that position, not lineage, determines cell fate. Planta 211, 191-199.

Kirch, T., Sion, R., Grünewald, M., and Werr, W. (2003). The *DORNRÖSCHEN/ENHANCER OF SHOOT REGENERATION1* gene of *Arabidopsis* acts in the control of meristem cell fate and lateral organ development. The Plant Cell 15, 694-705.

Křeček, P., Skůpa, P., Libus, J., Naramoto, S., Tejos, R., Friml, J., and Zažímalová, E. (2009). Genome Biology 10, 249.

Krizek, B.A., and Fletcher, J.C. (2005). Molecular mechanisms of flower development: an armchair guide. Nature Reviews Genetics 6, 688-698.

Krogan, N.T., Hogan, K., and Long, J.A. (2012). APETALA2 negatively regulates multiple floral organ identity genes in *Arabidopsis* by recruiting the co-repressor TOPLESS and the histone deacetylase HDA19. Development 139, 4180-4190.

Kuhlemeier, C. (2007). Phyllotaxis. Trends in Plant Science 12, 143-150.

Lau, S., Shao, N., Bock, R., Jürgens, G., and De Smet, I. (2009). Auxin signaling in algal lineages: fact or myth? Trends in Plant Science 14, 182-188.

Laux, T., Mayer, K.F.X., Berger, J., and Jürgens, G. (1996). The *WUSCHEL* gene is required for shoot and floral meristem integrity in *Arabidopsis*. Development 122, 87-96.

Lenhard, M., and Laux, T. (2003). Stem cell homeostasis in the *Arabidopsis* shoot meristem is regulated by intercellular movement of CLAVATA3 and its sequestration by CLAVATQA1. Development 130, 3163-3173.

Liao, C., Smet, W., Brunoud, G., Yoshida, S., Vernoux, T., and Weijers, D. (2015). Reporters for sensitive and quantitative measurement of auxin response. Nature Methods 12, 207-210.

Liu, C., Xu, Z., and Chua, N. (1993). Auxin polar transport is essential for the establishment of bilateral symmetry during early plant embryogenesis. The Plant Cell 5, 621-630.

Ljung, K., Bhalerao, R.P., and Sandberg, G. (2001). Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth. The Plant Journal 28, 465-474.

Lokerse, A.S., and Weijers, D. (2009). Auxin enters the matrix – assembly of response machineries for specific outputs. Current Opinion in Plant Biology 12, 520-526.

Long, J.A., Moan, E.I., Medford, J.I., and Barton, M.K. (1996). A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. Nature 379 66-69.

Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M. (2006). TOPLESS regulates apical embryonic fate in *Arabidopsis*. Science 312, 1520-1523.

Long, J.A., Woody, S., Poethig, S., Meyerowitz, E.M., and Barton, M.K. (2002). Transformation of shoots into roots in *Arabidopsis* embryos mutant at the TOPLESS locus. Development 129, 2297-2306.

Lukowitz, W., Roeder, A., Parmenter, D., and Somerville, C. (2004). A MAPKK kinase gene regulates extra-embryonic cell fate in *Arabidopsis*. Cell 116, 109-119.

Mähönen, A.P., ten Tusscher, K., Siligato, R., Smetana, O., Díaz-Triviño, S. Salojärvi, J., Wachsman, G., Prasad, K., Heidstra, R., and Scheres, B. (2014). PLETHORA gradient formation mechanism separates auxin responses. Nature 515, 125-129.

Malavé, T.M., and Dent, S.Y.R. (2006). Transcriptional repression by Tup1-Ssn6. Biochemistry and Cell Biology 84, 437-443.

Mano, Y., and Nemoto, K. (2012). The pathway of auxin biosynthesis in plants. Journal of Experimental Botany 63, 2853-2872.

Mansfield, S.G., and Briarty, L.G. (1991). Early embryogenesis in Arabidopsis thaliana: II. The developing embryo. Canadian Journal of Botany 69, 461-476.

Maraschin, F.S., Memelink, J., and Offringa, R. (2009). Auxin induced, SCF^{TIR1}-mediated poly-ubiquination marks AUX/IAA proteins for degradation. The Plant Journal 59, 100-109.

Mattsson, J., Sung, Z.R., and Berleth, T. (1999). Responses of plant vascular systems to auxin transport inhibition. Development 126, 2979-2991.

Mattsson, J., Ckurshumova, W., and Berleth, T. (2003). Auxin signaling in *Arabidopsis* leaf vascular development. Plant Physiology 131, 1327-1339.

Mayer, U., Büttner, G., and Jürgens, G. (1993). Apical-basal pattern formation in the *Arabidopsis* embryo: studies on the role of the *gnom* gene. Development 117, 149-162.

Mayer, K.F.X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. Cell 95, 805-815.

Michniewicz, M., Zago, M.K., Abas, L., Weijers, D., Schweighofer, A., Meskiene, I., Heisler, M.G., Ohno, C., Zhang, J., Huang, F., Schwab, R., Weigel, D., Meyerowitz, E.M., Luschnig, C., and Friml, J. (2007). Antagonistic regulation of PIN phosphorylation by PP2A and PINOID directs auxin flux. Cell 130, 1044-1056.

Möller, B., and Weijers, D. (2009). Auxin control of embryo patterning. Cold Spring Harbor Perspectives in Biology 1, 1-13.

Nakajima, K., Sena, G., Nawy, T., and Benfey, P.N. (2001). Intercellular movement of the putative transcription factor SHR in root patterning. Nature 413, 307-311.

Nakamura, A., Higuchi, K., Goda, H., Fujiwara, M.T., Sawa, S., Koshiba, T., Shimada, Y., and Yoshida, S. (2003). Brassinolide induces *IAA5*, *IAA19*, and DR5, a synthetic auxin response

element in *Arabidopsis*, implying a cross talk point of brassinosteroid and auxin signaling. Plant Physiology 133, 1843-1853.

Nakata, M., Matsumoto, N., Tsugeki, R., Rikirsch, E., Laux, T., and Okada, K. (2012). Roles of the middle domain-specific *WUSCHEL-RELATED HOMEOBOX* genes in early development of leaves in *Arabidopsis*. The Plant Cell 24, 519-535.

Natesh, S., and Rau, M.A. (1984). The embryo. In Embryology of Angiosperms (ed. B.M. Johri), Springer-Verlag.

Nawy, T., Bayer, M., Mravec, J., Friml, J., Bimbaum, K.D., and Lukowitz, W. (2010). The GATA factor HANABA TARANU is required to position the proembryo boundary in the early *Arabidopsis* embryo. Developmental Cell 19, 103-113.

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 57, 613-628.

Ogawa, E., Yamada, Y., Sezaki, N., Kosaka, S., Kondo, H., Kamata, N., Abe, M., Komeda, Y., and Takahashi, T. (2015). *ATML1* and *PDF2* play a redundant and essential role in *Arabidopsis* embryo development. Plant and Cell Physiology 56, 1183-1192.

Okada, K., Ueda, J., Komaki, M.K., Bell, C.J., and Shimura, Y. (1991). Requirement of the auxin polar transport system in early stages of *Arabidopsis* floral bud formation. The Plant Cell 3, 677-684.

Okushima, Y., Overvoorde, P.J., Arima, K., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R., Hughes, B., Lui, A., Nguyen, D., Onodera, C., Quach, H., Smith, A., Yu, G., and Theologis, A. (2005). Functional genomic analysis of the *AUXIN RESPONSE FACTOR* gene family

members in *Arabidopsis thaliana*: unique and overlapping functions of ARF7 and ARF19. The Plant Cell 17, 444-463.

Paciorek, T., Zažímalová, E., Ruthardt, N., Petrášek, J., Stierhof, Y., Kleine-Vehn, J., Morris, D.A., Emans, N., Jürgens, G., Geldner, N., and Friml, J. (2005). Auxin inhibits endocytosis and promotes its own efflux from cells. Nature 435, 1251-1256.

Petrášek, J., Mravec, J., Bouchard, R., Blakeslee, J.J., Abas, M., Seifertová, D., Wiśniewska, J., Tadele, Z., Kubeš, M., Čovanová, M., Dhonukshe, P., Skůpa, P., Benková, E., Perry, L., Křeček, P., Lee, O.R., Fink, G.R., Geisler, M., Murphy, A.S., Luschnig, C., Zažímalová, E., and Friml, J. (2006). PIN proteins perform a rate-limiting function in cellular auxin efflux. Science 312, 914-918.

Pinon, V., Prasad, K., Grigg, S.P., Sanchez-Oerez, G.F., and Scheres, B. (2013). Local auxin biosynthesis regulation by PLETHORA transcription factors controls phyllotaxis in *Arabidopsis*. Proceedings of the National Academy of Sciences 110, 1107-1112.

Prasad, K., Grigg, S.P., Barkoulas, M., Yadav, R.K., Sanchez-Perez, G.F., Pinon, V., Blilou, I., Hofhuis, H., Dhonukshe, P., Galinha, C., Mähönon, A.P., Muller, W.H., Raman, S., Verkleij, A.J., Snel, B., Reddy, G.V., Tsiantis, M., and Scheres, B. (2011). *Arabidopsis* PLETHORA transcription factors control phyllotaxis. Current Biology 21, 1123-1128.

Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N., and Clark, S.E. (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. The Plant Cell 17, 61-76.

Przemeck, G.K.H., Mattsson, J., Hardtke, C.S., Sung, Z.R., and Berleth, T. (1996). Studies on the role of the *Arabidopsis* gene *MONOPTEROS* in vascular development and plant cell axialization. Planta 200, 229-237.

Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Rios, A.F., Borst, J.W., Lukowitz, W., Jürgens, G., and Weijers, D. (2012). Different auxin response machineries control distinct cell fates in the early plant embryo. Developmental Cell 22, 211-222.

Rademacher, E.H., Möller, B., Lokerse, A.S., Llavata-Peris, C.I., van den Berg, W., and Weijers, D. (2011). A cellular expression map of *Arabidopsis AUXIN RESPONSE FACTOR* gene family. The Plant Journal 68, 597-606.

Raven, J.A. (1975). Transport of indoleacetic acid in plant cells in relation to pH and electrical potential gradients and its significance for polar transport. New Phytologist 74, 163-172.

Reinhardt, D., Pesce, E., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J., Friml, J., and Kuhlemeier, C. (2003). Regulation of phyllotaxis by polar auxin transport. Nature 426, 255-260.

Remington, D.L., Vision, T.J., Guilfoyle, T.J., and Reed, J.W. (2004). Contrasting modes of diversification in the *Aux/IAA* and *ARF* gene families. Plant Physiology 135, 1738-1752.

Rubery, P.H., and Sheldrake, A.R. (1974). Carrier-mediated auxin transport. Planta 118, 101-121.

Sabatini, S., Beis, D., Wolkenfelt, H., Murfett, J., Guilfoyle, T., Malamy, J.,Benfey, P., Leyser, O., Bechtold, N., Weisbeek, P., and Scheres, B. (1999). An auxin-dependent distal organizer of pattern and polarity in the *Arabidopsis* root. Cell 99, 463-472.

Sachs, T. (1981). The control of the patterned differentiation of vascular tissues. Advances in Botanical Research 9, 151-262.

Sachs, T. (1989). The development of vascular networks during leaf development. Current Topics in Plant Biochemistry and Physiology 8, 168-183.

Sachs, T. (2003). Collective specification of cellular development. BioEssays 25, 897-903.

Sarkar, A.K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., Scheres, B., Heidstra, R., and Laux, T. (2007). Conserved factors regulate signaling in *Arabidopsis thaliana* shoot and root stem cell organizers. Nature 446, 811-814.

Sato, A., and Yamamoto, K.T. (2008). Overexpression of the non-canonical *Aux/IAA* genes causes auxin-related aberrant phenotypes in *Arabidopsis*. Physiologia Plantarum 133, 397-405.

Saulsberry, A., Martin, P.R., O'Brien, T., Sieburth, L.E., and Pickett, F.B. (2002). The induced sector *Arabidopsis* apical embryonic fate map. Development 129, 3403-3410.

Sauer, M., Balla, J., Luschnig, C., Wiśniewska, J., Reinöhl, V., Friml, J., and Benková, E. (2006). Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. Genes & Development 20, 2902-2911.

Scheres, B., Di Laurenzio, L., Willemsen, V., Hauser, M., Janmaat, K., Weisbeek, P., and Benfey, P.N. (1995). Mutations affecting the radial organization of the *Arabidopsis* root display specific defects throughout the embryonic axis. Development 121, 53-62.

Scheres, B., Wolkenfelt, H., Willemsen, V., Terlouw, M., Lawson, E., Dean, C., and Weisbeek, P. (1994). Embryonic origin of the *Arabidopsis* primary root and root meristem initials. Development 120, 2475-2487.

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

Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F.X., Jürgens, G., and Laux, T. (2000). The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes. Cell 100, 635-644.

Schrader, J., Baba, K., May, S.T., Palme, K., Bennett, M., Bhalerao, R.P., and Sandberg, G. (2003). Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. Proceedings of the National Academy of Sciences 100, 10096-10101.

Scofield, S., Dewitte, W., and Murray, J.A.H. (2008). A model for *Arabidopsis* class-1 *KNOX* gene function. Plant Signaling and Behavior 3, 257-259.

Scofield, S., Dewitte, W., and Murray, J.A.H. (2014). STM sustains stem cell function in the *Arabidopsis* shoot apical meristem and controls *KNOX* gene expression independently of the transcriptional repressor AS1. Plant Signaling and Behavior 9, e28934.

Sessions, A., Weigel, D., and Yanofsky, M.F. (1999). The *Arabidopsis thaliana MERISTEM LAYER 1* promoter specifies epidermal expression in meristems and young primordia. The Plant Journal 20, 259-263.

Shevell, D.E., Leu, W., Gilmor, C.S., Xia, G., Feldmann, K.A., and Chua, N. (1994). *EMB30* is essential for normal cell division, cell expansion, and cell adhesion in *Arabidopsis* and encodes a protein that has similarity to Sec7. Cell 77, 1051-1062.

Smith, R.S, Guyomarc'h, Mandel, T., Reinhardt, D., Kuhlemeier, C., and Prusinkiewicz, P. (2006). A plausible model of phyllotaxis. Proceedings of the National Academy of Sciences 103, 1301-1306.

Smith, Z.R., and Long, J.A. (2010). Control of *Arabidopsis* apical-basal embryo polarity by antagonistic transcription factors. Nature 464, 423-427.

Spemann, H., and Mangold, H. (1924). Über induction von Embryonalanlagen durch Implantation artfremder Organisatoren. Archiv für mikroskopische Anatomie und Entwicklungsmechanik 100, 599-638.

Spencer, M.W.B., Casson, S.A., and Lindsey, K. (2007). Transcriptional profiling of the *Arabidopsis* embryo. Plant Physiology 143, 924-940.

Stahl, Y., and Simon, R. (2010). Plant primary meristems: shared functions and regulatory mechanisms. Current Opinion in Plant Biology 13, 53-58.

Stahl, Y., Wink, R.H., Ingram, G.C., and Simon, R. (2009). A signaling module controlling the stem cell niche in *Arabidopsis* root meristems. Current Biology 19, 909-914.

Steinmann, T., Geldner, N., Grebe, M., Mangold, S., Jackson, C.L., Paris, S., Gälweiler, L., Palme, K., and Jürgens, G. (1999). Coordinated polar localization of auxin efflux carrier PIN1 by GNOM ARF GEF. Science 286, 316-318.

Stepanova, A.N., Robertson-Hoyt, J., Yun, J., Benavente, L.M., Xie, D., Doležal, Schlereth, A., Jürgens, G., and Alonso, J.M. (2008). *TAA1*-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. Cell 133, 177-191.

Szemenyei, H., Hannon, M., and Long, J.A. (2008). TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. Science 319, 1384-1386.

Takada, S., Hibara, K., Ishida, T., and Tasaka, M. (2001). The *CUP-SHAPED COTYLEDON1* gene of *Arabidopsis* regulates shoot apical meristem formation. Development 128, 1127-1135.

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

Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M., and Zheng, N. (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature 446, 640-645.

Tao, Y., Ferrer, J., Ljung, K., Pojer, F., Hong, F., Long, J.A., Li, L., Moreno, J.E., Bowman, M.E., Ivans, L.J., Cheng, Y., Lim, J., Zhao, Y., Ballaré, C.L., Sandberg, G., Noel, J.P., and

Chory, J. (2008). Rapid synthesis of auxin via a new tryptophan-dependent pathway is required for shade avoidance in plants. Cell 133, 164-176.

Thimann, K.V., and Skoog, F. (1933). Studies on the growth hormone of plants: III. The inhibiting action of the growth substance on bud development. Proceedings of the National Academy of Sciences 19, 714-716.

Tiwari, S.B., Hagen, G., and Guilfoyle, T. (2003). The roles of auxin response factor domains in auxin-responsive transcription. The Plant Cell 15, 533-543.

Tiwari, S.B., Hagen, G., and Guilfoyle, T.J. (2004). Aux/IAA proteins contain a potent transcriptional repression domain. The Plant Cell 16, 533-543.

Turing, A.M. (1952). The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society (B) 237, 37-72.

Ueda, M., Zhang, Z., and Laux, T. (2011). Transcriptional activation of *Arabidopsis* axis patterning genes *WOX8/9* links zygote polarity to embryo development. Developmental Cell 20, 264-270.

Ulmasov, T., Hagen, G., and Guilfoyle, T.J. (1999). Activation and repression of transcription by auxin response factors. Proceedings of the National Academy of Sciences 96, 5844-5849.

Ulmasov, T., Murfett, J., Hagen, G., and Guilfoyle, T.J. (1997). Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. The Plant Cell 9, 1963-1971.

Van den Berg, C., Willemsen, V., Hage, W., Weisbeek, P., and Scheres, B. (1995). Cell fate in the *Arabidopsis* root meristem determined by directional signaling. Nature 378, 62-65.

Venglat, S.P., Dumonceaux, T., Rozwadowski, K., Parnell, L., Babic, V., Keller, W., Martienssen, R., Selvaraj, G., and Datla, R. (2002). The homeobox gene

BREVIPEDICELLUS is a key regulator of inflorescence architecture in *Arabidopsis*. Proceedings of the National Academy of Sciences 99, 4730-4735.

Vieten, A., Vanneste, S., Wiśniewska, J., Benková, E., Benjamins, R., Beeckman, T., Luschnig, C., and Friml, J. (2005). Functional redundancy of PIN proteins is accompanied by auxin-dependent cross-regulation of PIN expression. Development 120, 4521-4531.

Vuylsteker, C., Dewaele, E., and Rambour, S. (1998). Auxin induced lateral root formation in chicory. Annals of Botany 81, 449-454.

Wang, H., Ngwenyama, N., Liu, Y., Walker, J.C., and Zhang, S. (2007). Stomatal development and patterning are regulated by environmentally responsive mitogen-activated protein kinases in *Arabidopsis*. The Plant Cell 19, 63-73.

Wiśniewska, J., Xu, J., Seifertová, D., Brewer, P.B., Růžička, K., Blilou, I., Rouquié, D., Benková, E., Scheres, B., and Friml, J. (2006). Polar PIN localization directs auxin flow in plants. Science 312, 883.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. Journal of Theoretical Biology 25, 1-47.

Wysocka-Diller, J.W., Helariutta, Y., Fukaki, H., Malamy, J.E., and Benfey, P.N. (2000). Molecular analysis of SCARECROW function reveals a radial patterning mechanism common to root and shoot. Development 127, 595-603.

Yadav, R.K., Perales, M., Gruel, J., Girke, T., Jönsson, H., and Reddy, G.V. (2011). WUSCHEL protein movement mediates stem cell homeostasis in the *Arabidopsis* shoot apex. Genes & Development 25, 2025-2030.

Žadníková, P., and Simon, R. (2014). How boundaries control plant development. Current Opinion in Plant Biology 17, 116-125.

Zhao, Y., Medrano, L., Ohashi, K., Fletcher, J.C., Yu, H., Sakai, H., and Meyerowitz, E.M. (2004). HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in *Arabidopsis*. The Plant Cell 16, 2586-2600.

Zhu, Z., Xu, F., Zhang, Y., Cheng, Y.T., Wiermer, M., Li, X., and Zhang, Y. (2010). *Arabidopsis* resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. Proceedings of the National Academy of Sciences 107, 13960-13965.

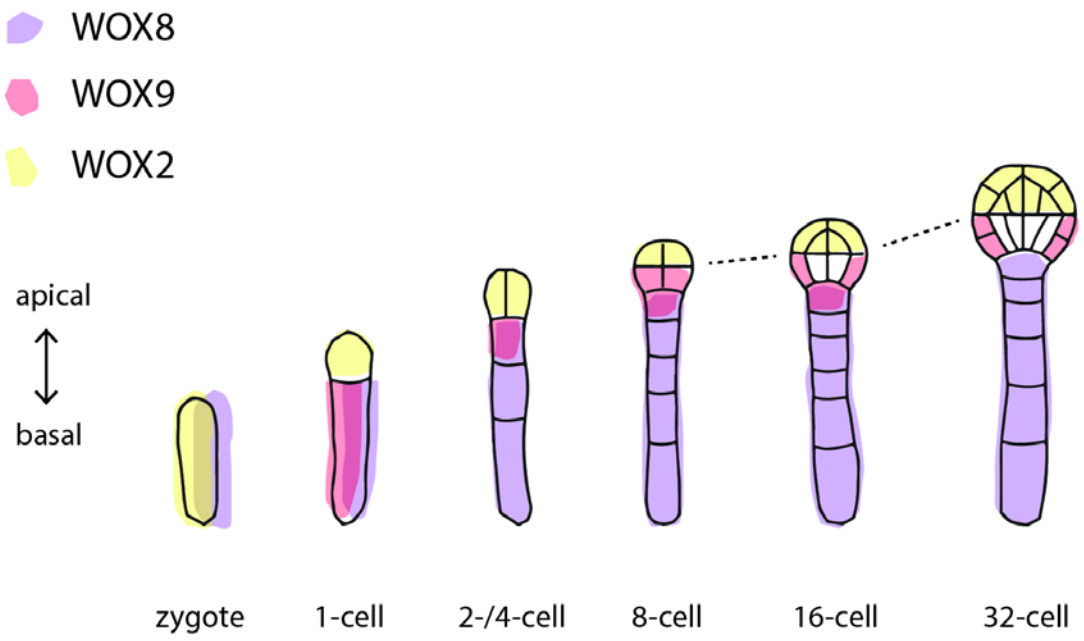

Figure 1.1. Dynamics of WOX gene expression in embryogenesis. Embryo development in *Arabidopsis* proceeds by a regular sequence of cell divisions. Transcriptional regulators of the WOX family define domains along the apical-basal axis from the division of the zygote. Adapted from Jenik et al. (2007).

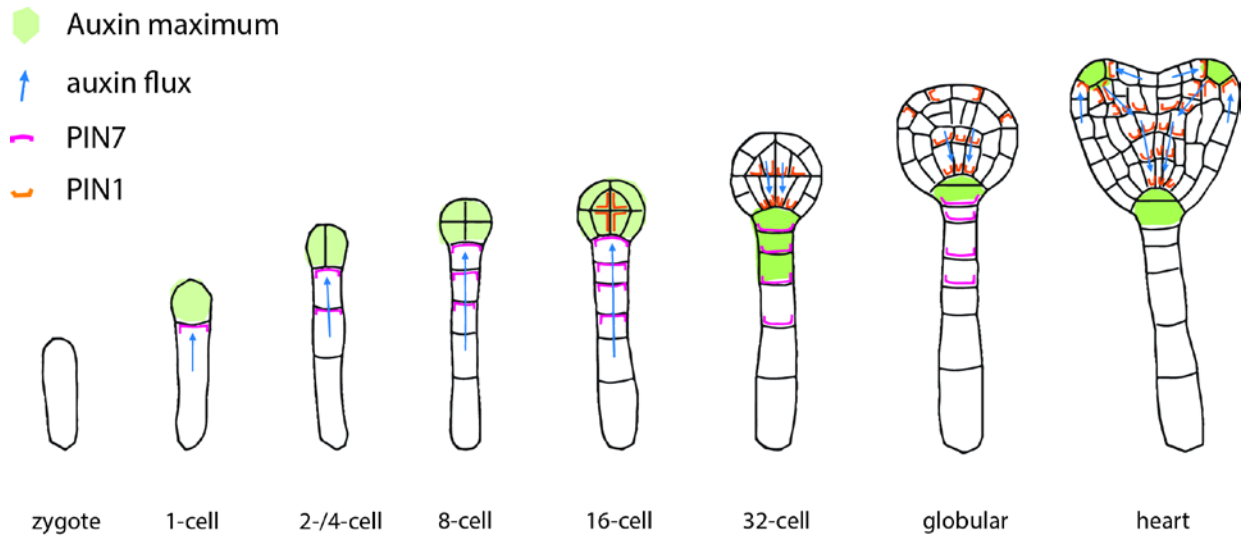

Figure 1.2. Auxin transport and accumulation during embryogenesis. Prior to the 32-cell stage, auxin (blue arrows) is thought to be directed upwards into the proembryo by PIN7, which localizes to the apical membranes of the basal cell and its daughters. At the globular stage, PIN7 localization switches to the basal membranes of the hypophysis and suspensor cells, and PIN1 assembles at the basal membranes of cells in the central domain of the proembryo. This arrangement results in apical-to-basal auxin transport, which continues throughout the life of the plant. Sites of auxin accumulation may be visualized with a synthetic reporter of auxin-dependent transcription, called DR5. Weak expression of DR5 is initially detected in the apical cell and its daughters, until the 32-cell stage when strong expression is found in the hypophysis and upper-most suspensor cells. Beginning at the heart stage, PIN1 reorganized to cells at the flanks of the upper portion of the embryo, where it directs auxin towards the incipient cotyledon primordia, resulting in DR5 expression. Adapted from Jenik et al. (2007).

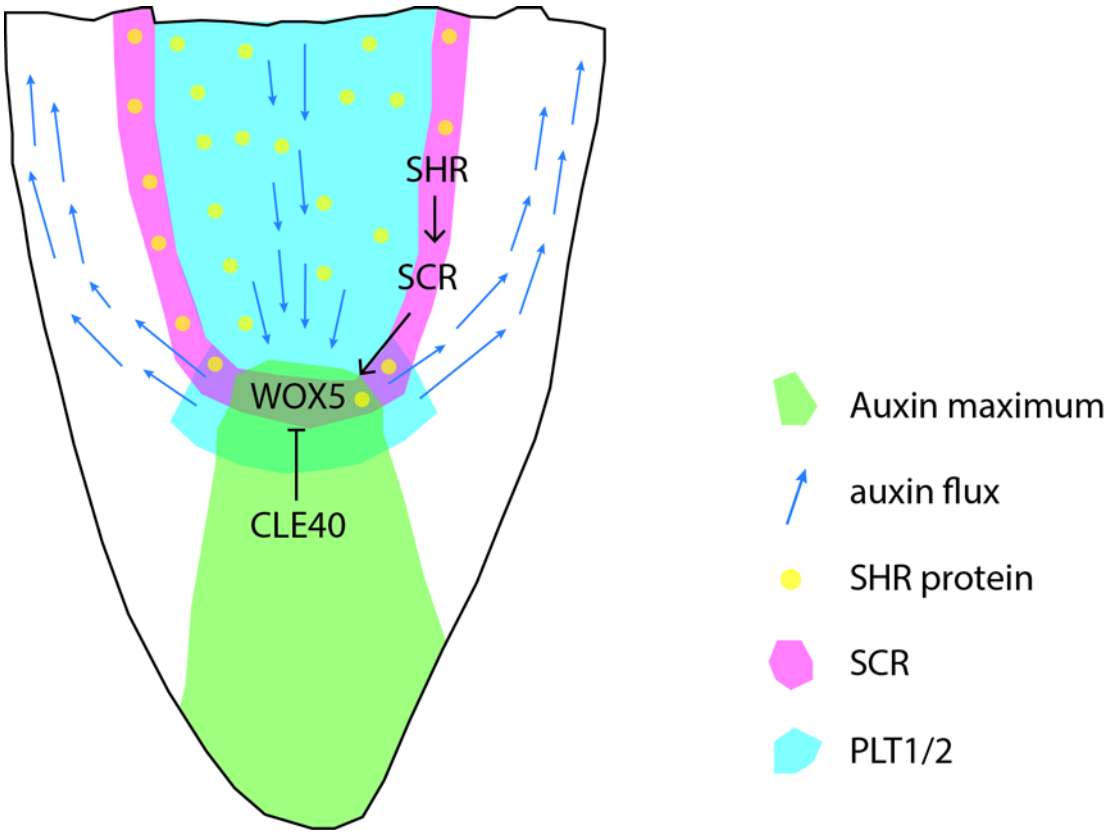

Figure 1.3. Positioning and maintenance of the root stem cell niche. In the root apical meristem, auxin is actively transported through the central provasculature to the tip, resulting in an auxin maximum extending from the quiescent center to the columnella. An auxin gradient with a maximum towards the distal meristem activates the expression of PLT genes, which function to promote quiescent center (QC) fate. In parallel, SHR protein moves from its expression domain in the central provascular cylinder to activate SCR in ground tissue cells and the QC, where they act together to maintain QC fate. Current models suggest that the position of the QC, and thus, the stem cell niche is determined by the overlap between SCR and SHR expression and an auxin maximum or PLT activity. The maintenance of the root QC also requires WOX5, which is expressed specifically in the QC from the late globular stage on. The size of the stem cell pool is restricted by the CLE40 signaling peptide, which represses WOX5 expression in the columnella and in doing so, enables the differentiation of distal stem cells. The illustration is a simplified depiction of interactions between factors promoting root development, not all interactions are direct. Adapted from Cederholm et al. (2012).

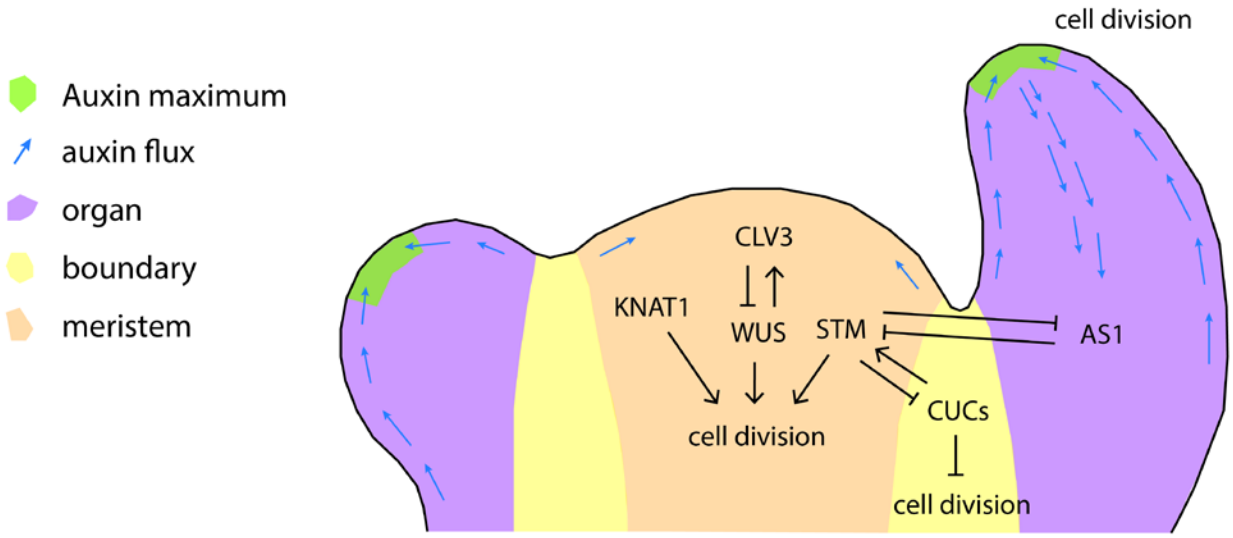

Figure 1.4. Regulation of phyllotaxis and maintenance of the shoot apical meristem. PIN1 transporters in the L1, or epidermis layer of the shoot meristem direct auxin to the periphery, where localized responses promote cell division and the outgrowth of organ primordia. The expression domain of the NAC domain transcription factors CUC1/2 delimits the boundary of organ primordia; their expression is also required to activate the KNOX gene STM in the center of the shoot meristem. STM promotes meristem function, and closely related KNOX genes promote stem cell fate and antagonize AS1, which promotes differentiation in organ primordia. The antagonistic relationship between STM and AS1 balances the recruitment of founder cells for primordia at the periphery with the maintance of stem cells at the tip. The size of the stem cell population of the shoot apical meristem is controlled by a negative feedback mechanism between WUS and the diffusible peptide CLV3. WUS protein moves from its expression domain in the stem cell organizing center to the shoot apex where it directly activates the transcription of CLV3. CLV3 is secreted as a signal for neighboring cells to repress WUS transcription. Adapted from Zadnikova and Simon (2014).

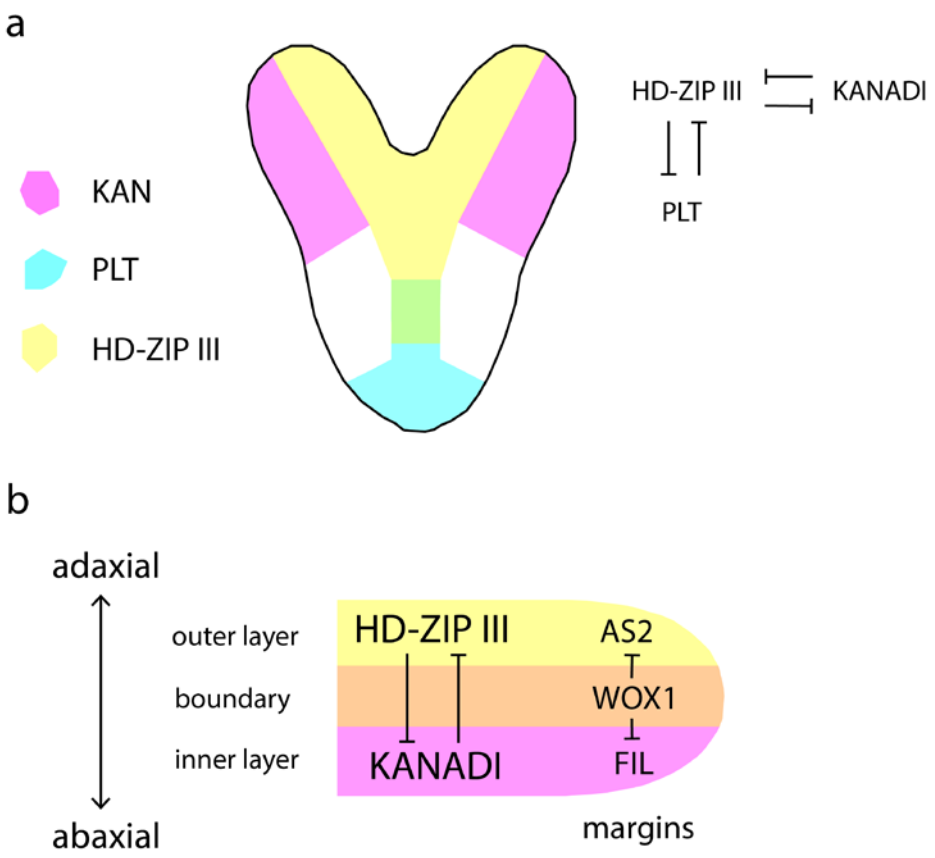

Figure 1.5. Transcriptional control of lateral organ polarity in leaves and embryos. The genes regulating lateral organ polarity in embryos and leaves also provide positional information within the shoot apical meristem and antagonize root development.(a) In the embryo, transcription factors of the HD-ZIP III family are expressed at the adaxial sides of cotyledons and in the central vascular cylinder and repress both PLT genes at the basal pole, and members of the KANADI (KAN) family at the abaxial cotyledon surface. Similarly, KAN and PLT genes antagonize expression of HD-ZIP III genes. As a result, loss of one group of factors causes expanded expression of the other and, in extreme cases, the homeotic transformation of tissues or primordia.(b) Interaction between HD-ZIP III genes and KAN genes persists throughout the development of leaves. As leaf primordia form, antagonism between these factors regulates organ polarity. Importantly, at the margins of developing lateral organs (depicted above), WOX1 becomes expressed at the boundary between the adaxial and abaxial surfaces repressing AS2 and FIL to facilitate laminal expansion.

CHAPTER 2

ANALYSIS OF HAN-FAMILY MUTANTS REVEALS EMBRYOS WITH A PARTIALLY RADIALIZED PRIMARY AXIS[1]

[1] Volny, M.F., and W. Lukowitz. To be submitted to Developmental Cell

Abstract

Arabidopsis embryos display distinct polarity from the earliest stages of development. The asymmetric division of the zygote generates daughter cells that differ in their size and cellular composition, and by the 8-cell stage, homeodomain transcription factors mark four distinct cell identities arranged along the apical-basal axis. In subsequent divisions these domains give rise to basic elements of the plant body plan like the shoot and root meristem and the primary tissues. The molecular network responsible for establishing and refining the coordinates of the primary axis after fertilization remain largely unknown. Previously, we have shown that the GATA transcription factor, HANABA TARANU (HAN), is required for positioning the inductive boundary between the proembryo and suspensor, at which the root initiates. The loss of HAN results in an apical shift in gene expression domains normally demarcating the proembryo boundary. Despite this defect, mutants often recover and organize a root between cells of the upper and lower tier. Here, we show that this recovery is due to the action of two closely related GATA factors, HAN-LIKE 1 (HANL1) and HANL2. Loss of all three HAN family genes causes arrest in late embryogenesis; triple mutants develop into oblong, blimp-shaped structures with abnormally enlarged surface cells and no recognizable apical meristems or provascular cells. Molecular markers of embryonic polarity, such as DRN, WUS, or PLT1, remain unchanged in these embryos, but many regulators of organogenesis or meristem maintenance, such as FIL, TMO5, and CLV3 are never expressed, suggesting a breakdown in the patterning process. Strikingly, several genes normally expressed on the basal side of the proembryo boundary over time become transcribed in domains that surround triple mutant embryos in concentric layers: expression of WOX8 and SUC3, normally restricted to the suspensor and distal root, expands to all surface cells; and WOX5, normally active in the quiescent center, is

transcribed in all sub-surface cells. We interpret this phenotype as a complete loss of the proembryo boundary, resulting in partial radialization of the apical-basal axis.

Introduction

Plants have a flexible body plan generated by branching in response to environmental variables. Elaborate structures such as stems, leaves, and flowers are formed post-embryonically from meristems set at either pole of the plant body in a continuous process throughout the life of the plant. In *Arabidopsis*, the shoot and root meristems can be traced to small groups of founder cells in the early embryo. Following fertilization, the zygote divides into a smaller cytoplasmically dense apical cell, and a larger vacuolated basal cell. These cells differ in their developmental fate: over three rounds of divisions the apical cell generates a spherical 8-cell (octant) proembryo, while the basal cell divides only transversely and gives rise to the filamentous suspensor. At this stage, four different domains can be readily distinguished along the apical-basal axis of the embryo; these include an upper tier, lower tier, the uppermost derivative of the basal cell, called the hypophysis, and the suspensor. These domains will produce the principal organs: the upper tier will form the shoot meristem and most of the embryonic leaves or cotyledons, the lower tier will form the hypocotyl and most of the primary root, while the hypophysis will produce the quiescent center of the root meristem and the distal root cap or columella.

Although many key regulators of the shoot and root meristem have been identified, it remains largely unknown how the apical and basal poles are first specified in the embryo. The earliest markers for precursors of the meristems include members of the WUSCHEL-RELATED HOMEODOMAIN (WOX) family of transcription factors. The proper expression of these genes requires the zinc finger transcription factor WRKY2 (Ueda et al., 2011), and members of the YDA MAP kinase signaling pathways (Lukowitz et al., 2004; Wang et al., 2007; Jeong et al., 2011); loss of these functions affect the asymmetric division of the zygote, resulting in smaller

basal cells and suspensor defects. The WOX genes are differentially expressed in each of the four cell types along the apical-basal axis at the octant stage (Haecker et al., 2004. WOX2 transcripts are detected in the upper tier, WOX9 in the lower tier, WOX8 and WOX9 in the hypophysis, and WOX8 in the suspensor. The specification of these domains is important as mutating WOX genes can have severe consequences. Loss of both WOX8/9 results in zygotes which divide normally, but often display aberrant divisions giving rise to enlarged or misshapen cells; and in a small proportion of embryos the apical cell generates a filamentous stalk of vacuolated suspensor-like cells, and eventually arrests (Breuninger et al., 2008). Similarly, loss of *wox2* in combination with other WOX genes expressed in the proembryo, results in rod-like or shootless seedlings (Breuninger et al., 2008).

The hormone auxin is a pervasive signal for plant growth and development, and is required for many key patterning steps in the embryo (reviewed in Tanaka et al., 2006). Auxin is directionally transported, and strong responses often require localized areas of auxin accumulation, called auxin maxima (Sabatini et al., 1999; Friml et al., 2002; Dubrovsky et al., 2008). These localized sinks of auxin are generated through the activity of the PIN (named for PIN-FORMED 1) family of efflux carriers (Křeček et al., 2009). PINs are transmembrane proteins, the polar subcellular localization of which determines the direction of auxin flux (Friml et al., 2004). Following the division of the zygote, PIN7 localizes to the upper membrane of the basal cell and its daughters, the suspensor cells, likely resulting in auxin accumulation in the apical cell and its daughters; consistent with this view, a synthetic auxin response gene, termed DR5, is weakly expressed in the apical cell and subsequently the proembryo (Ulmasov et al., 1997; Friml et al., 2003). By the early globular stage, PIN7 localization switches to the basal side of suspensor cells and PIN1 assembles at the basal membranes of cells in the lower half of

the proembryo, the precursors of the vascular cylinder. This re-localization establishes apical-to-basal auxin transport as it persists throughout the life cycle and drives robust DR5 expression in the hypophysis and sub-hypophyseal cell.

Auxin accumulation is interpreted in plant cells by a short signaling pathway that directly influences transcription. The binding of auxin to auxin signaling F-box components of E3 ubiquitin ligase complexes enables their interaction with members of the Aux/IAA family of corepressors and targets them for degradation (Dharmasiri et al., 2005; Kepinski and Leyser, 2005; Tan et al., 2007; Maraschin et al., 2009). Aux/IAA proteins bind a large and diverse family of transcriptional regulators called AUXIN RESPONSE FACTORS (ARFs); which when freed by the auxin-dependent degradation of Aux/IAAs induce transcriptional changes (Ulmasov et al. 1999; Tiwari et al. 2003). Embryos that are unable to synthesize, transport, or respond to auxin fail to establish a main body axis, root meristem, vascular cylinder or cotyledons (reviewed in Möller and Weijers, 2009). For example, mutants lacking multiple PIN transporters display abnormal divisions in the basal embryo; they are often rootless, and occasionally form ball-shaped seedlings without any indication of polarity (Friml et al., 2003). Similar phenotypes are observed for embryos treated with chemical inhibitors of polar auxin transport (Liu et al., 1993, Hadfi et al., 1998), or in transgenic lines where auxin flux is redirected by reversing the subcellular localization of PIN1 from the basal to the apical side of proembryo cells (Friml et al., 2004). Inhibiting auxin responses, by mutating the ARF MONOPTEROS (MP), or expressing auxin-resistant variants of its repressor BODENLOS (BDL) cause vascular defects, and a failure to develop a hypocotyl and a root (Berleth and Jürgens, 1993; Hardtke and Berleth, 1998; Hamann et al., 1999). One of the earliest abnormalities of these mutants is that the uppermost cell of the suspensor, or hypophysis, fails to produce a small, lens-shaped apical daughter at the

late globular stage; since the lens-shaped cell is the precursor of the quiescent center, around which the meristem is organized, this asymmetric division marks the initiation of the embryonic root.

While the importance of polar auxin transport for axis and root formation is well established, its function in the initial establishment of embryonic polarity is less clear. Before the 32-cell stage, expression of the DR5 reporter is low, and auxin mutants have either weak or no visible phenotypes (Berleth and Jürgens, 1993; Cheng et al., 2007). For example, *mp* and *gn* mutants only show clear abnormalities from the mid globular stage; *pin* multiple mutants occasionally display dramatic early defects, including early arrest phenotypes, but they have low penetrance, while most *pin* multiple mutants resemble *gn* embryos (Friml et al., 2003). Thus, it is difficult to assess the role of auxin in the earliest stages of embryogenesis. The establishment of the auxin biosynthesis, transport, and interpretation machinery with appropriate spatial configuration in the early embryo requires the action of WOX genes, which, in turn, are regulated by the WRKY2 transcription factor, and a MAP kinase signaling pathway including the MAPKK kinase YDA and the RWP-RK transcription factor GRD. For example, PIN1 expression in the proembryo is absent in *wox8 wox9* mutants, while the transcription of DR5 expands throughout the embryo (Breuninger et al., 2008). WOX8 is a direct target of WRKY, and in *wrky2* mutants WOX8 expression is lost from the upper-most suspensor cells; WOX9 expression, normally detected in the lower tier of globular stage embryos, expands basally in these mutants to largely replace WOX8 function in those cells (Ueda et al., 2011). Similarly, *yda* or *grd* mutants often fail to express WOX genes in the suspensor and incipient QC (Lukowitz et al., 2004; Jeong et al., 2011). The loss of single *wox* genes has either a minimal or no effect on early development, while multiple mutants tend to completely disrupt the process, causing early

arrest phenotypes that are inherently difficult to interpret; this effect has hampered pinpointing their exact contribution to the process.

Loss of the GATA-TYPE transcription factor HANABA TARANU (HAN) dramatically alters the arrangement of gene expression domains in the early embryo (Nawy et al., 2010). HAN is expressed more broadly than early WOX genes: in the zygote, apical cell, and throughout the proembryo. In *han* mutants, transcripts normally found in the suspensor expand into the lower tier, while transcripts normally present in the entire proembryo become confined to the upper tier (Nawy et al., 2010). Like WOX genes, HAN influences the spatial expression of PIN transporters; both PIN1 and PIN7 are expressed in domains that are more apical along the primary axis at the 8-cell stage, PIN7 expands into the lower tier and PIN1 becomes restricted to the upper tier, and this coincides with DR5 expression detected throughout the lower tier of *han* embryos instead of in the hypophysis. These observations have been interpreted as a coordinated apical shift in the boundary between the proembryo and suspensor, and in support of this view the base of *han* embryos form swollen, suspensor-like cells which divide in aberrant planes. Despite lacking the anatomical hallmarks of a normal root primordium, *han* mutants often recover to form complete seedlings. These plants develop normally until they produce flowers, which are comprised of fewer, sometimes fused floral organs, and they are almost completely sterile (Zhao et al., 2004). Combining a weak allele of *han* with *angustifolia 3* (*an3*), a gene expressed in leaf primordia and required for regulating organ size (Horiguchi et al., 2005; Kawade et al., 2013), results in seedlings with malformed cotyledons that frequently form an ectopic root at the shoot pole (Kanei et al., 2012). Similar double root seedlings are produced by plants harboring a conditional, dominant-negative allele of the TOPLESS (TPL) gene (Long et

al., 2002; Long et al., 2006; Smith and Long, 2010). The synergistic interaction of *han* and *an3* is striking but its molecular basis is not well understood.

The GATA gene family in *Arabidopsis* comprises 30 members, which have been subdivided based on their exon/intron structure and their zinc finger into four classes named A–D (Reyes et al., 2004); HAN is in the B-class, and two closely related HAN-like genes, HANL1 and HANL2 have been identified in Arabidopsis; the defining feature of the HAN genes is a conserved 14 amino acid sequence towards the N-terminus, termed a HAN domain (Reyes et al., 2004; Zhao et al., 2004; Whipple et al., 2010). Our previous work on *han* mutants left open a key question: how does a proembryo boundary (and a root meristem) form between the upper and lower tier, why is the boundary not lost entirely? To address this question, we have investigated the potential contribution of HANL1 and HANL2 to this process. We found that the HAN-like genes are biochemically equivalent to HAN and required for the recovery of *han* single mutants. Loss of all three *han* family genes results in oblong, blimp-shaped embryos, which form abnormally enlarged surface cells, and lack the anatomical hallmarks of a vascular cylinder. Early molecular markers of embryonic polarity are generally unchanged in these embryos, while several regulators of organogenesis or meristem maintenance are never expressed, suggesting a breakdown in the patterning process. Strikingly, several genes normally restricted to the basal side of the proembryo boundary over time surround triple mutant embryos in concentric layers. These include WOX8 and IAA10, the transcripts of which are normally found in the suspensor and distal root, and expand to all surface cells; and WOX5, which is normally active in the quiescent center, which expands to all sub-surface cells.

Results

Previous studies have noted that the *Arabidopsis* genome encodes two close homologues to HAN, named HANL1 and HANL2 (Zhao et al., 2004). HAN family proteins are defined by a near identical Zinc-finger DNA binding domain and a short stretch of conserved amino acids at their N-terminal end, called a HAN domain, but share little sequence similarity otherwise (Zhao et al., 2004; Whipple et al., 2010; Figure 1a,b). HAN, HANL1, and HANL2 reside in large syntenic blocks that can be traced back to genome duplication and triplication events in the evolutionary history of Arabidopsis (Reyes et al., 2004; Vanneste et al., 2015; summarized in Fig. 2.1a): HANL1 and HANL2 were separated in the α whole genome duplication (WGD), approximately 45 mya ago, while HAN dates back to the γ triplication at the base of the eudicots, approximately 125 mya ago (the β WGD, approximately 60 mya ago, did not increase HAN family genes in Arabidopsis). Similarly, the origin of HAN-family genes in other angiosperms can be dated to WGDs in their respective lineages (Fig. 2.1a; see legend for details). The sequenced genome of *Amborella*, the apparent sister of all other extant angiosperms, contains a single HAN gene as well as an entry that either represents a truncated gene or an incomplete record (not included in Fig. 2.1a).

Arabidopsis HAN-family proteins can provide equivalent function

Loss-of-function *han* mutants typically recover from their early defects and develop into complete seedlings. To determine whether this effect can be explained by functional redundancies within the HAN family, we examined embryo development in *han* mutants transformed with the HAN-like genes. We constructed a T-DNA construct containing ~6.8 kbp segment upstream of the HAN translational start site (Nawy et al., 2010; previously shown to be

sufficient for *han* rescue), followed by the HANL1 or HANL2 coding sequences. All T-DNAs were transformed into *han/+ hanl1 hanl2* plants, and the embryos of at least 10 independent primary transgenic lines were examined by DIC microscopy. In both cases, the expression of a single copy of either HAN-like gene was sufficient for rescue, and facilitated the recovery of fully fertile *han-16* homozygous plants. These results indicate that the HAN-like genes are biochemically equivalent to HAN in embryo and flower development.

Expression of HAN-LIKE genes is below detection before the globular stage and similar to HAN expression thereafter

We next assessed whether the expression domains of HAN and HAN-like genes overlap in embryos using fluorescent reporter genes and *in situ* hybridization. HAN transcripts have been detected in the zygote, the apical cell, and all cells of the early proembryo, before they become gradually restricted to provascular cells at the globular stage (Nawy et al., 2010). Fluorescent reporters of HAN, HANL1, and HANL2 were constructed by fusing their prospective promoters (~6.8 kbp, ~3.7 kbp, and ~2.9 kbp 5'UTR fragments, respectively, extending from the start codon to the next transcription unit) to the coding sequence of a triple Venus fluorescent protein with a nuclear localization signal (see methods for details). The HAN reporter faithfully recapitulated HAN expression as previously described on the basis of *in situ* hybridization (Fig. 2.1a,f). HANL1 and HANL2 expression was below detection prior to the globular stage but then closely mimicked the pattern of HAN expression (Fig. 2.1c-e). All three genes are broadly expressed in the early globular embryo, but are strongest in central cells directly adjacent to the proembryo boundary. Following the division of the hypophysis, HANL1 and HANL2 become restricted to the central cells of the lower tier until the heart stage, when expression is detected

throughout the provasculature of both the hypocotyl and cotyledon primordia. Subsequent expression remains strong in provascular cells and eventually also covers the shoot meristem and the proximal initials of the root meristem. In situ hybridization with antisense RNA probes directed to non-conserved regions of each HAN-family gene confirmed these observations (Fig. 2.1f,g,h; Appendix, Fig. S2.1). Our findings imply that HAN-like genes are unlikely to contribute to early stages of embryonic development but may provide redundant function to HAN after the globular stage.

The recovery of han is due to the activity of HAN-like genes after the globular stage

Loss of *han* results in severe anatomical defects in the basal proembryo and suspensor – but many embryos recover to form essentially normal seedlings. To determine whether this effect is due to the activity of HAN-like genes after the early globular stage, we first examined whether HANL1 and HANL2 expression was dependent on HAN activity. *In situ* hybridization revealed that both HANL1 and HANL2 remain strongly expressed in the central upper tier of *han* embryos, suggesting that their expression domains are shifted apically in a similar manner to other genes (Fig. 2.1i,j). We next investigated the consequence associated with loss of HAN-like gene function in the embryo. Point mutations in HANL1 and HANL2 were identified by the Arabidopsis TILLING facility (Till et al., 2003), and predicted loss-of-function alleles were recovered for both genes (Fig. 2.1b): the *hanl1-56* allele introduces an amino acid substitution within the zinc finger domain (T^{102}>I); the same substitution at the homologous position of the HAN zinc finger is found in the *han-18* allele (T^{162}>I) which closely mimics the effects of the *han-16* truncation allele (Nawy et al., 2010). The *hanl2-79* allele introduces a stop codon upstream of the zinc finger domain (W^{55}>Z).

Loss of *hanl1* and *hanl2* separately, or in combination does not cause an appreciable effect on plant growth; however, there is a subtle but significant effect on suspensor cell number (Fig. 2.2t): at the 32-cell stage, wild type embryos have 7.4 ± 0.8 suspensor cells (n = 118) versus 6.5 ± 0.9 cells in *hanl1 hanl2* (n = 166; $p \leq 0.05$, student-t test). Loss of *han* in combination with either *hanl1* or *hanl2* results in embryos that by anatomy resemble *han-16*; again, these embryos display fewer suspensor cells at the 32-cell stage: we counted 4.7 ± 0.7 cells for *han-16* (n = 109) compared to 4.1 ± 0.9 cells for *han hanl1* (n = 118), and 3.9 ± 0.8 cells in *han hanl2* (n = 123; $p \leq 0.05$ for both comparisons to *han-16*, student-t test). These subtle effects imply that HANL1 and HANL2 make noticeable contributions to embryonic development prior to the 32-cell stage, before robust expression first becomes detectable. However, more severe, novel developmental defects arise only in *han hanl1 hanl2* mutants.

Triple mutant embryos are initially similar to *han-16* (Fig. 2.2g-i,n-p), distinguishable only by their reduced number of suspensor cells (2.9 ± 0.8 cells; n = 112; $p \leq 0.05$, student-t test). Unlike *han*, triple mutants deteriorate steadily during later stages of embryogenesis. Perhaps the most obvious defect is seen in the surface cells: division rates are low compared to wild type (Fig. 2.2u; by the bent cotyledon stage, when they typically arrest, we counted fewer surface cells across the median section of triple mutants than in wild type embryos at the transition stage), such that most *han hanl1 hanl2* embryos become entirely surrounded by abnormally large surface cells. These mutants display no anatomical hallmarks of cotyledon primordia, apical meristems, or a vascular cylinder and arrest as slightly oblong, blimp-shaped structures (Fig. 2.2s), with an overall size that is significantly smaller than wild type (Fig. 2.2v; approximately the size of wild type embryos at the heart stage)

To ensure that the observed effects were indeed due to the loss of HAN-family genes, we transformed *han-16/+ hanl1-56 hanl2-79* plants with T-DNA constructs containing the entire HANL1 or HANL2 locus (6.2 kbp and 5.5 kpb, respectively, extending to the neighboring transcription units on both sides) and assayed transgenic plants for molecular complementation (see methods for details). A single copy of either HAN-like gene was sufficient for converting the phenotype of triple mutant embryos to the phenotype of *han-16* embryos (Appendix A, Table S2.1). Taken together, our results establish that the recovery of *han* embryos is due to the activity of HANL1 and HANL2. Furthermore, loss of all HAN function results in a striking, novel arrest phenotype, suggesting that HAN-family genes may have a more fundamental role in embryonic patterning than previously appreciated.

Triple mutant embryos retain early markers of embryonic polarity

As triple mutant embryos lack the anatomical hallmarks of normal development, we probed the patterning process with a panel of fluorescent reporter genes. Early events we aimed to visualize were the establishment of embryonic polarity; specification of the epidermis primordium, or protoderm; auxin transport and the formation of an auxin response maximum. Later events included the specification of provascular tissue, initiation of the shoot meristem stem cell niche, cotyledon formation and the establishment of lateral organ polarity, and the onset of the maturation program.

Our survey reveals that developmental regulators acting prior to the globular stage are expressed fairly normally in *han-16 han1-56 hanl2-79* triple mutants. WOX2, in conjunction with other WOX genes, is required for normal development of the shoot structures and transcribed specifically in the apical daughter cell of the zygote, throughout the upper tier at the

8-cell stage, then gradually restricted to surface cells at the adaxial surface of developing cotyledons and the shoot meristem (Haecker et al., 2004; Breuninger et al., 2008); as their wild type counterparts, triple mutant embryos expressed a WOX2 reporter in surface cells covering their apical pole (Appendix A, Fig. S2.2a,b). MERISTEM LAYER1 (ML1) encodes a transcription factor specifying epidermal fate (Abe et al., 2003), and its expression becomes confined to the protoderm layer of wild type embryos when it is formed at the 16-cell stage (Lu et al., 1996); similarly, an ML1 reporter marked the surface cells of triple mutant embryos (Fig. 2.3a,b). WUSCHEL (WUS), required for maintaining the stem cell population of the shoot apical meristem, is transcribed in the inner cells of the upper tier at the 16-cell stage and later in a small cluster of cells underneath the embryonic apex (Mayer et al., 1998); triple mutant embryos showed a somewhat larger but otherwise analogous expression domain (Fig. 2.3c,d). DORNRÖSCHEN (DRN) regulates cotyledon formation, and transcripts accumulate in the apical domain of the proembryo from the 4-cell stage on, as well as at the tips of the cotyledon primordia in wild type (Chandler et al., 2007); similarly, a DRN reporter is expressed in the surface and sub-surface cells at the apical pole of triple mutants (Appendix A, Fig. S2.2c,d). PLETHORA1 (PLT1) encodes a master regulator of cell fates in the root meristem (Galinha et al., 2007); PLT expression marks the root pole, becoming first detectable in the hypophysis of 8-cell stage embryos, and is essentially unchanged in triple mutants (Fig. 2.3e,f). These results indicate that triple mutants retain molecular markers of embryonic polarity and protoderm specification.

Basic elements of the body plan are incompletely specified in triple mutant embryos

Other regulators of the patterning process were not detectable in *han-16 han1-56 hanl2-79* embryos; these included: CLV3, which controls the size of the stem cell population and is first observed in the incipient shoot meristem of heart stage embryos (Fletcher, 2002; Fig. 2.3i); FILAMENTOUS FLOWER/YABBY 1 (FIL/YAB1), which regulates lateral organ polarity and is initially expressed at the abaxial side of cotyledon primordia (Siegfried et at., 1999; Chen et al., 1999; Bonaccorso et al., 2012; Fig. 2.3n); and TARGET OF MONOPTEROS 5 (TMO5), which is required for the periclinal division generating provascular cells at the globular stage as well as vascular development in general, and is first detectable in the central cells of the proembryo (Schlereth et al., 2010; De Rybel et al., 2013; Fig. 2.3l). In contrast, all of these genes are expressed in han single mutants in domains that reflect their expression in wild-type (Fig. 2.3h,k). The absence of these regulators is in concordance with the apparent absence of apical meristems, cotyledon primordia, and a vascular cylinder in the triple mutant embryos.

However, not all genes contributing to the organization of tissues and organ primordia are similarly affected. For example, PHABULOSA (PHB) and HOMEOBOX GENE 8 (HB8), both members of the HD-ZIP III family, and shown to antagonize PLT genes in apical/basal patterning and FIL/YAB genes in the lateral organ polarity pathway, are transcribed predominantly in provascular cells; in addition, PHB transcripts are found in parts of the shoot apical meristem and the adaxial surface cells of cotyledon and leaf primordia (McConnell et al., 2001; Prigge et al., 2005; Smith and Long, 2010; Bowman et al., 2002). Reporters for both genes were robustly expressed in the central domain of *han-16* and *han-16 han1-56 hanl2-79* embryos – although expression of the PHB reporter was not observed in surface cells at the apex of triple mutants, and ATHB8 was occasionally absent from the central domain (Fig. 2.4b,c,e,f).

SHORTROOT (SHR), one of the earliest known markers of the provasculature and required for tissue specification in the vascular cylinder and endodermis (Helariutta et al., 2000), as well as PINHEAD/ZWILLE (PNH/ZLL), involved in the maintenance of the shoot stem cell population and transcribed in the center of developing embryos, are also expressed fairly normally in triple mutants (Fig. 2.4h,j).

Although they show little appreciable growth by the bent cotyledon stage, triple mutant embryos do initiate the transcriptional program associated with seed maturation. Reporters of OLEOSIN 2 (OLEO2), SEED STORAGE ALBUMIN 2 (SESA2), and CRUCIFERIN 3 (CRU3), which become expressed broadly in wild type at this stage (unpublished data), also show strong expression in *han hanl1 hanl2* mutants (Appendix A, Fig. S2.2e-f).

Triple mutants maintain the ability to transport and perceive auxin

Apical-to-basal auxin flux establishes the primary axis and is required for the formation of the vascular cylinder and root. To determine whether losing HAN-family function affects auxin responses, we assessed the expression of DR5, which in wild type marks the hypophysis at the 32-cell stage, and later the tips of coyledons (Friml et al., 2003); strong DR5 expression is associated with root and cotyledon initiation. As in *han* mutants, DR5 is expressed robustly in the lower tier of triple mutant embryos, but they never show expression at the embryo apex (Fig. 2.5a,b,c). Auxin responses in the proembryo are dependent on MP, which is broadly expressed at the globular stage and then becomes confined to provascular cells (Hardtke and Berleth, 1998). Similarly, expression of a MP reporter is generally limited to the central cells of triple mutants, although a broader expression domain has occasionally been observed (Appendix A, Fig. S2.3a,b). IAA10, which marks suspensor cells, and at the heart stage, the tips of cotyledon

primordia, is thought to mediate auxin responses in the suspensor (Rademacher et al., 2012). Interestingly, IAA10 appears more strikingly affected by the loss of HAN function than DR5 and MP and becomes expressed in surface cells along the periphery and the apex of triple mutants (Fig. 2.5d,e).

We next surveyed genes required for the directional transport of auxin. Several PIN genes are known to be active in the embryo; such as PIN1, which is broadly expressed prior to the globular stage, then confined to provascular cells and cotyledon tips (Steinmann et al., 1999); PIN7, which is initiated in the suspensor, then found in provascular cells (Friml et al., 2003); and PIN4, which is first detected in the hypophysis at the early globular stage, then in the QC and provascular cells (Friml et al., 2002). All were expressed in roughly equivalent domains in the triple mutants (Fig. 2.5f-I; Appendix A, Fig. S2.3c,d). We also looked at the expression of two PIN genes with poorly described roles in embryogenesis: PIN2, which in our hands marks the embryonic root tip, QC, and the adaxial surface of cotyledons; and PIN3, which is typically confined to the QC from the early heart stage (our results; Friml et al., 2003). Reporters for both genes displayed unusual expression patterns in triple mutants. PIN2 expression was found in the enlarged surface cells, surrounding the triple mutants, while PIN3 expression expanded into the central domain and, in some cases, appeared to be stronger in the sub-surface cells (Appendix A, Fig. S2.3e-i). These results suggest that triple mutants maintain the capacity for auxin transport and maintain an auxin maximum at the basal pole or lower tier.

Markers for the QC and distal root cap form concentric rings towards the periphery of triple mutant embryos

Triple mutant embryos present no visible indicators of a main axis, meristems or lateral organs, yet patterns of gene activity indicate that they are polar structures with at least rudimentary spatial organization. The expression domain of many genes are relatively normal. The most striking abnormalities we uncovered in our survey were with genes that are normally expressed below the proembryo boundary. Many of these genes showed a much more pronounced expansion of their spatial domains than in *han* single mutants. Specifically, expression of WOX8, SUCROSE TRANSPORTER 3 (SUC3) and IAA10 (mentioned above), which is normally confined to the suspensor and, after division of the hypophysis, the incipient columella (Breuninger et al., 2008; Meyer et al., 2004; Rademacher et al., 2012), expands into the lower tier in *han* single mutants but comes to cover the entire surface of triple mutants (Fig. 2.5j,k; Appendix A, Fig. S2.3j,k); WOX8 reaches the cells at the very apex only rarely, perhaps because it is antagonized by WOX2 (compare Appendix A, Fig. S2.3k with S2.2b) Similarly, WOX5, which in wild type is expressed in the root quiescent center (Sarkar et al., 2007), shows a broad expression domain at the base of *han* single mutant embryos at the globular stage before becoming confined as a root meristem is re-generated; in contrast, WOX5 expression expands into the layer directly beneath the surface cells of triple mutant embryos (Fig. 2.5l,m). In wild type, SCARECROW (SCR) is transcribed in a cup-shaped domain marking the QC and endodermis of the hypocotyl; in triple mutants, a SCR reporter is also expressed in the sub-surface cells at the apical pole, resulting in a closed, spherical domain (Appendix A, Fig. S2.3l,m). Thus, genes normally expressed in basal domains at the root pole (the quiescent center and the

columella) become expressed in concentric layers surrounding triple mutant embryos. We interpret this effect as a partial radialization of the apical-basal pattern.

Discussion

Our investigation of two closely related HAN paralogs, HANL1 and HANL2, found that they can provide biochemically equivalent function to HAN in embryo development. The expression of HAN-like genes is below detection before the globular stage but substantially overlaps with HAN expression thereafter and is not dependent on the activity of HAN. Consistent with this finding, the transcription of HANL2 was found to be elevated in *han* mutants (Zhang et al., 2013). Loss of all HAN family members results in embryos that are similar to *han* mutants prior to the globular stage; yet they deteriorate as many *han* mutants recover to produce complete seedlings. Triple mutants eventually arrest as blimp-shaped structures with fewer, abnormally enlarged surface cells; they lack cotyledons, apical meristems, and the anatomical hallmarks of a vascular cylinder. These results show that the recovery of han embryos is dependent on the activity of HANL1 and HANL2.

A revised model for the role of HAN in embryo patterning

A phenotypic analysis of han single mutants found a coordinated apical shift in many gene expression domains normally delimited by the boundary between the suspensor and proembryo (Nawy et al., 2010). For example, PIN7 expression expands from the suspensor into the lower tier, while PIN1 expression, normally throughout the proembryo, becomes confined to the upper tier. These changes in the spatial layout of PIN transporters are correlated with an apical shift of the auxin perception maximum from the hypophysis to the cells of the lower tier. At the globular

stage, *han* embryos show none of the anatomical features associated with axis and root initiation, and many arrest, however, others gradually recover, forming a root meristem in the center of the proembryo. On the basis of these findings it was proposed that HAN regulates positioning of the proembryo boundary, across which the embryonic root is initiated, and that in *han* single mutants an inductive boundary eventually forms between the upper and lower tier.

How does the triple mutant phenotype relate to this model? The loss of all three members of the HAN family results in blimp-shaped embryos, which form abnormal enlarged surface cells, and lack a discernible vascular cylinder. Many molecular markers of embryonic polarity are unchanged in these embryos, yet several genes normally restricted to the basal side of the proembryo boundary – marking the root quiescent center and the distal root tip – over time become expressed in concentric layers in surface and subsurface cells. We interpreted this phenotype as a partial radialization of the apical-basal axis. In the absence of all HAN function, a proemrbyo boundary cannot be maintained and basal functions expand throughout the apical-basal axis. To the best of our knowledge, a similar effect has not been reported in embryos; but parallels exist to mutants affecting lateral organ polarity.

Plant lateral organs, such as leaves, display distinct polarity along their abaxial/adaxial axis; this polarity is established in the early embryo by auxin and antagonistic interactions between FIL, KANADI *(*KAN*)*, and HD-ZIP III genes (reviewed in Kidner and Timmermans, 2010). Loss of *kan1*, *kan2*, and *kan3*, which are normally expressed at the abaxial surface of cotyledons and leaf primordia, produces radially symmetric leaves with adaxial characteristics (Kerstetter et al. 2001; Eshed et al., 2004). This radialized phenotype is caused by the expansion of HD-ZIP III expression, which is normally restricted to the adaxial leaf surface, into the domain previously occupied by KAN expression (Eshed et al. 2001). Concordantly, loss of the

HD-ZIP III genes *phb*, *phavoluta* (*phv*), and *revoluta* (*rev*) produces radially symmetric leaf primordia with abaxial characteristics (Emery et al. 2003; Prigge et al., 2005). Similarly, the loss of HAN-family genes from the proembryo boundary, results in a unidirectional expansion of QC and root columella fates into surface and subsurface cell layers; the domains maintain their tissue specificity, suggesting that positional cues of the radial axis are recognized while apical-basal cues are lost (Fig. 2.5n,o). According to this interpretation, triple mutants no longer have a proembryo boundary.

Which factors may be responsible for excluding QC and columnella transcripts from the central cylinder? Recently, a PHD-containing protein called REPRESSOR OF WUSCHEL 1 (ROW1) was found to directly repress WOX5 transcription in the initials of the root (Zhang et al., 2015). ROW1 is expressed in cells directly adjacent to the root QC, and in row1 mutants WOX5 transcription expands broadly into the periphery of the meristem (Zhang et al., 2015). This is in contrast to *han hanl1 hanl2* embryos, where WOX5 expression expands apically, but remains confined to a single cell layer below the protoderm. ROW1 activity in the vascular cylinder of the proembryo might explain this effect. Thus, the role of HAN genes in setting an apical limit for the specification of QC and distal root identities at the flanks, might be supported by factors like ROW1 acting to suppress their expansion in the central cylinder.

The triple mutant phenotype is likely not caused by defects in auxin signaling

Loss of *han* causes a re-configuration of the auxin transport machinery, resulting in an apically shifted auxin expression maximum. Indeed, it was proposed that these effects on auxin flux and perception may explain the aberrant development of *han* mutants (Nawy et al. 2010), implying that PIN auxin transporters may be the primary targets of HAN. Our results disagree with this

model. Prior to the globular stage, DR5 is expressed throughout the lower tier in both *han* and triple mutants. The expression of DR5 does not change however, even as transcripts normally confined to the root columella and QC become expressed in the enlarged surface cells and subsurface cells, respectively, of triple mutants. Thus, it seems difficult to reconcile the triple mutant phenotype with the observed defects in auxin transport and perception. PIN genes and MP are present and not drastically changed compared to *han* embryos. An exception however, was presented in the lack of TMO5 expression. TMO5 is essential for the production of vascular tissue, and its absence supports our observation that triple mutants fail to form a recognizable vascular cylinder. It is also a direct target of MP in the early embryo (Schlereth et al., 2010); the absence of TMO5 despite the presence of MP and an auxin maximum might indicate that TMO5 expression requires MP-independent regulatory inputs, and this idea is supported by weak TMO5 expression in *mp* mutants (Schlereth et al., 2010).

Similarities between radialization and mirror-image duplication phenotypes

Combining *an3*, a putative transcriptional co-activator expressed in leaf primordia (Horiguchi et al., 2005; Kawade et al., 2013), and the weak *han-30* allele frequently results in the formation of malformed leaves and an ectopic root at the shoot pole of seedlings (Kanei et al., 2012). Similarly, the *tpl-1* allele causes seedlings to form a root in place of the shoot (Long et al., 2002). The phenotype of *han hanl1 hanl2* embryos is different, but its radialization effect shares some obvious similarity to the mirror-image duplication phenotype associated with *tpl-1* or *han-30 an3* mutations: in both cases, genes that would normally mark the QC and root columnella identities are expressed in the apical as well as basal regions of the embryo. Perhaps the most apparent parallel is in the expression of SCR, which displays a symmetrical pattern throughout the ground

tissue of *tpl-1* as well as *han hanl1 hanl2* triple mutants (Long et al., 2006; this study). Markers for apical cell fate such as SHOOTMERISTEMLESS are missing in *tpl-1* (Long et al., 2002), similar to the loss of CLV3 and FIL at the apex of triple mutants.

Unlike *han an3* and *tpl-1* mutants however, *han hanl1 hanl2* mutants do not express PLT, which is a master regulator of root fate, at both the apical and basal pole. PLT expression is critical for the duplication effect, as double root seedlings fail to form in *tpl-1 plt1 plt2* triple mutants (Smith and Long, 2010). The expression domains of other genes are also different in *han* triple mutants. WOX5 is ectopically expressed in *han an3* embryos, yet its expression pattern is limited to the apical and basal pole (Kanei et al., 2012), not in a single sub-epidermal cell layer, as it is in oblong triple mutants.

However, the most prominent distinction between double root mutants and *han* triple embryos is that the latter arrest and never develop into viable seedlings. Our analysis indicates that the growth of triple mutants is sluggish even at early stages of embryogenesis and they produce markedly fewer surface cells than either wild type or *han* single mutants, suggesting a strong effect on cell division rates. This is possibly a consequence of their patterning defects; but it should be noted that HAN family genes are expressed in the initials of the shoot and root meristem, and the loss of *han* is correlated with a reduction in the size of the shoot apical meristem and more diffuse expression of WUS (Zhao et al., 2004). Thus, there may be a more direct connection between HAN function and the cell division machinery.

Conclusions

HAN family transcription factors cannot be placed into any of the known developmental pathways controlling patterning in the embryo. These genes appear to set an apical limit to gene

expression domains at the boundary between suspensor and proembryo, positioning QC and distal root cap identities. HAN-family members are strongly expressed throughout the lower tier of embryos, suggesting they may prevent the expansion of basal genes into this domain.

Placing HAN family members into our expanding framework for embryonic patterning will necessitate identifying factors that regulate HAN or HAN target genes. Ironically, the best approach to identifying such targets, might be by searching in the shoot or root apical meristem. HAN family members are expressed in the initials of the root meristems, and at the boundaries that separate the central zone of the shoot meristem from the incipient primordia (Zhao et al., 2004), and HAN is required for normal function of the inflorescence and flower meristems (Zhao et al., 2004). Additionally, induced over-expression of HAN in seedlings results in swollen roots, with a diminished root cap; eventually resulting in seedling death (Zhao et al., 2004). As these results suggest that HAN functions in both the shoot and root apical meristem, these tissues would likely provide better source material for studying protein-DNA binding or transcriptional profiling. Targeting the entire family by inducible RNAi or inducible CRISPR/Cas (Wielopolska et al., 2005; Jiang et al., 2013) would provide a complimentary approach for overcoming embryonic lethality and assessing the role of HAN and HAN-like genes in either of these tissue types.

Materials and Methods

Arabidopsis Lines and Growth Conditions

Plants were grown under constant illumination (~100 μmol/m^2/s) at 23°–25°C and 40%–50% relative humidity using commercial potting mix (RediEarth, Sun Gro Horticulture) containing

systemic insecticide (Marathon 1% G, Olympic Horticultural Products) with slow-release fertilizer (19/12/6 Osmocote, Miracle-Gro).

The *han-16* and *han-18* alleles as well as a co-dominant marker for *han-16* were described previously (Nawy et al., 2010). The *hanl1-56* and *hanl2-79* mutations were identified in the *Arabidopsis* TILLING facility following EMS-mutagenesis of Col *er-105* (Till et al., 2003). Co-dominant PCR-based markers were used to follow the mutations independently of visible phenotypes: for *hanl1-56,* a ~280 bp fragment was amplified with the primer pair aatgacgaccaaaaccaccac / gtccttttggtccgtttcttc and digested with TasI, yielding fragments of ~150+90+40 bp with wild type and ~120+90+40+30 with mutant template; for *hanl2-79,* a ~250 bp fragment was amplified with the primer pair ggtttctcaatgttcttctccc / agaagtggtgtcgcagttgg and digested with PflMI, yielding fragments of ~160+90 bp with wild type and 250 bp with mutant template.

Fluorescent reporter lines introduced into the triple mutant background by crossing are described in: *DR5rev-GFP*, Col accession (Friml et al., 2003); *ML1:H2B-YFP*, Col accession (Roeder et al., 2010); *pPLT1:CFP*, L*er* accession (Galinha et al., 2007); *pWUS*::*mGFP5-ER*, Ler accession (Yadav et al., 2009); *pSCR:GFP*, L*er* accession (Wysocka-Diller et al., 2000); *pSHR:GFP*, L*er* accession (Helariutta et al., 2000); *pWOX8::nYFP*, *pWOX2::nYFP* & *ZLL-YFP*, Col accession (Breuninger et al., 2008); *pWOX5:GFP*, Col accession (Blilou et al., 2005); *SUC3:GFP*, C24 accession (Meyer et al., 2004); pCLV3**::**mGFP5-ER &, Ler accession (Reddy & Meyerowitz, 2005); *pIAA10::3xGFP-NLS*, Col accession (Rademacher et al., 2012); pMP::MP-GFP, Col accession (Krogan et al., 2012); *pDRNL::erGFP*, Col accession (Chandler et al., 2011). A *PHB-2xYFP* reporter created by recombineering and introduced into the L*er* accession was kindly made available by J. Long (University of California, Los Angeles).

Reporter constructs directly transformed into triple mutant lines are described below except for: *pFIL::dsRed-N7* (Heisler & al., 2005); *pTMO5::3XGFP* (De Rybel et al., 2014).

T-DNA construction

The HANL1 rescue construct contains a 6.2kbp FspI-EheI genomic DNA fragment that was obtained from the BAC clone T30D6 and covers the entire locus, extending ~3.7kbp upstream of the start codon. The HANL2 rescue construct contains a 5.5kbp HindII-EcoRI fragment obtained from the BAC clone C7A10 and extends ~2.9kbp upstream of the start codon. Both fragments were inserted into a modified pCambia3300 T-DNA vector conferring resistance to glufosinate (Basta). The HAN rescue construct consists of a 9.1 kbp EcoRI-SalI fragment extending ~6.8 kbp upstream of the start codon and was described previously (Nawy et al. 2010).

For the purpose of gene swapping, the HAN rescue construct was mutagenized to introduce NcoI and HindIII restriction sites at the HAN start and stop codons, respectively (cc**ATG**G---**TAA**gctt); the second codon was changed from atg to gtg in the process, resulting in an M2>V substitution. As control, a PCR-generated HAN cDNA fragment containing corresponding restriction sites was inserted into the modified T-DNA; the resulting construct fully complemented the phenotype of *han-16* embryos, indicating that neither the M2>V mutation nor the absence of intron sequences significantly affected its function. The coding sequences of HANL1 and HANL2 were amplified from genomic DNA with PCR primers introducing NcoI and HindIII sites at their start and stop codons, respectively, and inserted into the modified T-DNA to generate the chimeric genes (see Appendix A, Table S2.2 for details).

Transcriptional reporters of HAN, HANL1, and HANL2 were created by introducing restriction sites at the start codons of the sub-cloned genomic fragments (BglII in the case of

HAN: **ATG**CCGAGATCT; NcoI in the case of HANL1 and HANL2: cc**ATG**G), and then combining the 5' upstream fragments of the genes with the coding sequence of a triple Venus FP with an N-terminal SV40 nuclear localization signal (kindly provided by T. Vernaux, INRA, Versailles). Reporters of HB8 and PIN genes were constructed in a similar manner from PCR-generated fragments containing the 5' upstream sequences of these genes including the start codon. All T-DNAs contain a modified pCambia3300 vector, in which the 35S promoter of the Basta resistance gene was replaced by a Nos promoter to avoid potential interference. Primers used for the purpose are listed in Appendix A, Table S2.2; annotated listings of the predicted plasmid sequences are available upon request.

Microscopy & Molecular Complementation

For light microscopy, whole-mount immature seeds were dissected from siliques, embedded in modified Hoyer's solution (70% chloral hydrate, 4% glycerol, 5% gum arabic; diluted 2:1 with 20% gum arabic), and cleared for 2 hours at room temperature (pre-globular stages) or overnight at 4^{o}C (older stages). The embryos contained in the cleared seeds were then imaged by differential interference contrast (DIC) microscopy using a Leica DMR microscscope equipped with a Qimaging micropublisher 5.0 digital camera.

Fluorescent reporter genes were examined in a *han-16*/+ *hanl1-56 hanl2-79* background, as *han* mutants are sterile. Seeds containing embryos of the desired stage were dissected from siliques, placed in 15% glycerol containing propidium iodide as a counter stain (10 mg/l). Slides were imaged immediately upon release of the embryos with a Zeiss LSM 710 confocal microscope using a 20× lens and standard filter sets. GFP, YFP, and PI were excited with the 488 nm line of an argon laser and CFP with the 457 nm line. Triple mutant embryos were often

slow in taking up the counter stain, such that the two fluorescent channels could not be imaged at the same time and had to be overlaid manually.

The rescue activity of HAN::HANL1 and HAN::HANL2 constructs was tested in the T1 generation of transgenic plants with the genotype *han-16/+ han1-56 hanl2-79* by counting the number of blimp-shaped arrest mutants. As the *han-16* allele is recessive, plants segregate approximately 25% mutant embryos; complementation resulted in a reduction in the number of triple mutant phenotypes, typically to ~6%, indicating that one copy is sufficient for HAN function. At a minimum 10 independent lines per construct, and at least 100 embryos of each line were examined. Complementation in the T2 generation was confirmed for select lines.

In Situ Hybridization

Siliques of various developmental stages were fixed overnight at 4°C (4% formaldehyde in phosphate-buffered saline) and prepared for sectioning and hybridization as described (Long et al., 1996). Digoxigenin-labeled RNA probes were generated by in vitro transcription from PCR fragments containing a T7 promoter sequence at their ends (see Appendix A, Table S2.2 for details; the HAN probe is described in Zhao et al., 2004). Reagents for labeling and detection were obtained from Roche and used according to the included instructions. The hybridization was performed at 60-65°C overnight. The HAN probes produced robust signals after incubation over night. In contrast, HANL1 and HANL2 probes generated weaker signals; color reactions were allowed to continue for up to five days, with daily change of the substrate solution.

References

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. The Plant Cell 15, 63-78.

Berleth, T., and Jürgens, G. (1993). The role, of the *monopteros* gene in organising the basal region of the *Arabidopsis* embryo. Development 118, 575-587.

Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K., and Scheres, B. (2005). The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots.Nature 433, 39-44.

Bonaccorso, O., Lee, J.E., Puah, L., Scutt, C.P., and Golz, J.F. (2012). FILAMENTOUS FLOWER controls lateral organ development by acting as both an activator and a repressor. BMC Plant Biology 12, 176.

Bowman, J.L., Eshed, Y., and Baum, S.F. (2002). Establishment of polarity in angiosperm lateral organs. Trends in Genetics 18, 134-141.

Breuninger, H., Rikirsch, E., Hermann, M., Ueda, M., and Laux, T. (2008). Differential expression of *WOX* genes mediates apical-basal axis formation in the *Arabidopsis* embryo. Developmental Cell 14, 867-876.

Chandler, J.W., Cole, M., Flier, A., Grewe, B., and Werr, W. (2007). The AP2 transcription factors DORNRÖSCHEN and DORNRÖSCHEN-LIKE redundantly control *Arabidopsis* embryo patterning via interaction with PHAVOLUTA. Development 134, 1653-1662.

Chandler, J.W., Cole, M., Jacobs, B., Comelli, P., and Werr, W. (2011). Genetic integration of *DORNRÖSCHEN and DORNRÖSCHEN-LIKE* reveals hierarchical interactions in auxin

signaling and patterning of the *Arabidopsis* apical embryo. Plant Molecular Biology 75, 223-236.

Chen, Q., Atkinson, A., Otsuga, D., Christensen, T, Reynolds, L., and Drews, G.N. (1999). The *Arabidopsis FILAMENTOUS FLOWER* gene is required for flower formation. Development 126, 2715-2726.

Cheng, Y., Dai, X., and Zhao, Y. (2007). Auxin synthesized by the YUCCA flavin monooxygenases is essential for embryogenesis and leaf formation in *Arabidopsis*. The Plant Cell 19, 2430-2439.

De Rybel, B., Möller, B., Yoshida, S., Grabowicz, I., de Reuille, P.B., Boeren, S., Smith, R.S., Borst, J.W., and Weijers, D. (2013). A bHLH complex controls embryonic vascular tissue establishment and indeterminate growth in *Arabidopsis*. Developmental Cell 24, 426-437.

Dharmasiri, N., Dharmasiri, S., and Estelle, M. (2005). The F-box protein TIR1 is an auxin receptor. Nature 435, 441-445.

Dubrovsky, J.G., Sauer, M., Napsuclaly-Mendivil, S., Ivanchenko, M.G., Friml, J., Shishkova, S., Celenza, J., and Benková, E. (2008). Auxin acts as a local morphogenic trigger to specify lateral root founder cells. Proceedings of the National Academy of Sciences 105, 8790-8794.

Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F., an Bowman, J.L. (2003). Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. Current Biology 13, 1768-1774.

Eshed, Y., Baum, S.F., Perea, J.V., and Bowman, J.L. (2001). Establishment of polarity in lateral organs of plants. Current Biology 11, 1251-1260.

Eshed, Y., Izhaki, A., Baum, S.F, Floyd, S.K., and Bowman, J.L. (2004). Asymmetric leaf development and blade expansion in Arabidopsis are mediated by KANADI and YABBY activities. Development 131, 2997-3006.

Fletcher, J.C. (2002). Shoot and floral meristem maintenance in *Arabidopsis*. Annual Review of Plant Biology 53, 45-66.

Friml, J., Benková, E., Blilou, I., Wisniewska, J., Hamann, T., Ljung, K., Woody, S., Sandberg, G., Scheres, B., Jürgens, G., and Palme, K. (2002). AtPIN4 mediates sink-driven auxin gradients and root patterning in *Arabidopsis*. Cell 108, 661-673.

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwartz, H., Hamann, T., Offringa, R., and Jürgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. Nature 426, 147-153.

Friml, J., Yang, X., Michniewicz, M., Weijers, D., Quint, A., Tietz, O., Benjamins, R., Ouwerkerk, P.B.F., Ljung, K., Sandberg, G., Hooykaas, P.J.J., Palme, K., and Offringa, R. (2004). A PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. Science 306, 862-865.

Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R., and Scheres, B. (2007). PLETHORA proteins as dose-dependent master regulators of *Arabidopsis* root development. Nature 449, 1053-1057.

Hadfi, K., Speth, V., and Neuhaus, G. (1998). Auxin-induced developmental patterns in *Brassica juncea* embryos. Development 125, 879-887.

Haecker, A., Groß-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrman, M., and Laux, T. (2004). Expression dynamics of *WOX* genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. Development 131, 657-668.

Hamann, T., Mayer, U., and Jürgens, G. (1999). The auxin-insensitive *bodenlos* mutation affects primary root formation and apical-basal patterning in the *Arabidopsis* embryo. Development 126, 1387-1395.

Hardtke, C.S., and Berleth, T. (1998). The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. The EMBO Journal 17, 1405-1411.

Heisler, M.G., Ohno, C., Das, P., Sieber, P., Reddy, G.V., Long, J.A., and Meyerowitz, E.M. (2005). Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the *Arabidopsis* inflorescence meristem. Current Biology 15, 1899-1911.

Helariutta, Y., Fukaki, H., Wysocka-Diller, J., Nakajima, K., Jung, J., Sena, G., Hauser, M., and Benfey, P. (2000). The *SHORT-ROOT* gene controls radial patterning of the *Arabidopsis* root through radial signaling. Cell 101, 555-567.

Horiguchi, G., Kim, G., and Tsukaya, H. (2005). The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. The Plant Journal 43, 68-78.

Jeong, S., Bayer, M., and Lukowitz, W. (2011). Taking the very first steps: from polarity to axial domains in the early *Arabidopsis* embryo. Journal of Experimental Botany 62, 1687-1697.

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nature Biotechnology 31, 233-239.

Jiao, Y., Leebens-Mack, J., Ayyampalayam, S., Bowers, J.E., McKain, M.R., McNeal, J., Rolf, M., Ruzicka, D.R., Wafula, E., Wickett, N.J., Wu, X., Zhang, Y., Wang, J., Zhang, Y., Carpenter, E.J., Deyholos, M.K., Kutchan, T.M., Chanderbali, A.S., Soltis, P.S., Stevenson,

D.W., McCombie, R., Pires, J.C., Wong, G.K., Soltis, D.E., and dePamphilis, C.W. (2012). A genome triplication associated with early diversification of the core eudicots. Genome Biology 13, R3.

Kanei, M., Horiguchi, G., and Tsukaya, H. (2012). Stable establishment of cotyledon identity during embryogenesis in *Arabidopsis* by *ANGUSTIFOLIA3* and *HANABA TARANU*. Development 139, 2436-2446.

Kawade, K., Horiguchi, G., Usami, T., Hirai, M.Y., and Tsukaya, H. (2013). ANGUSTIFOLIA3 signaling coordinates proliferation between clonally distinct cells in leaves. Current Biology 23, 788-792.

Kepinski, S., and Leyser, O. (2005). The *Arabidopsis* F-box protein TIR1 is an auxin receptor. Nature 435, 446-451.

Kerstetter, R.A., Bollman, K., Taylor, R.A., Bomblies, K., and Peothig, R.S. (2001). KANADI regulates organ polarity in *Arabidopsis*. Nature 411, 706-709.

Kidner, C.A., and Timmermans, M.C.P. (2010). Chapter five – signaling sides: adaxial-abaxial patterning in leaves. Current Topics in Developmental Biology 91, 141-168.

Křeček, P., Skůpa, P., Libus, J., Naramoto, S., Tejos, R., Friml, J., and Zažímalová, E. (2009). Genome Biology 10, 249

Krogan, N.T., Hogan, K., and Long, J.A. (2012). APETALA2 negatively regulates multiple floral organ identity genes in *Arabidopsis* by recruiting the co-repressor TOPLESS and the histone deacetylase HDA19. Development 139, 4180-4190.

Lie, C., Kelsom, C., and Wu, X. (2012). *WOX2* and *STIMPY-LIKE/WOX8* promote cotyledon boundary formation in Arabidopsis. The Plant Journal 72, 674-682.

Liu, C., Xu, Z., and Chua, N. (1993). Auxin polar transport is essential for the establishment of bilateral symmetry during early plant embryogenesis. The Plant Cell 5, 621-630.

Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M. (2006). TOPLESS regulates apical embryonic fate in *Arabidopsis*. Science 312, 1520-1523.

Long, J.A., Woody, S., Poethig, S., Meyerowitz, E.M., and Barton, M.K. (2002). Transformation of shoots into roots in *Arabidopsis* embryos mutant at the TOPLESS locus. Development 129, 2297-2306.

Lu, P., Porat, R., Nadeau, J.A., and O'Neill, S.D. (1996). Identification of a meristem L1 layer-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox genes. The Plant Cell 8, 2155-2168.

Lukowitz, W., Roeder, A., Parmenter, D., and Somerville, C. (2004). A MAPKK kinase gene regulates extra-embryonic cell fate in *Arabidopsis*. Cell 116, 109-119.

McConnell, J.R., Emery, J., Eshed, Y., Bao, N., Bowman, J., and Barton, M.K. (2001). Role of *PHABULOSA* and *PHAVOLUTA* in determining radial patterning in shoots. Nature 411, 709-713.

Maraschin, F.S., Memelink, J., and Offringa, R. (2009). Auxin induced, SCF^{TIR1}-mediated poly-ubiquination marks AUX/IAA proteins for degradation. The Plant Journal 59, 100-109.

Mayer, K.F.X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. Cell 95, 805-815.

Meyer, S., Lauterbach, C., Niedermeier, M., Barth, I., Sjolund, R.D., and Sauer, N. (2004). Wounding enhances expression of AtSUC3, a sucrose transporter from *Arabidopsis* sieve elements and sink tissues. Plant Physiology 134, 684-693.

Möller, B., and Weijers, D. (2009). Auxin control of embryo patterning. Cold Spring Harbor Perspectives in Biology 1, 1-13.

Nawy, T., Bayer, M., Mravec, J., Friml, J., Bimbaum, K.D., and Lukowitz, W. (2010). The GATA factor HANABA TARANU is required to position the proembryo boundary in the early *Arabidopsis* embryo. Developmental Cell 19, 103-113.

Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N., and Clark, S.E. (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. The Plant Cell 17, 61-76.

Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Rios, A.F., Borst, J.W., Lukowitz, W., Jürgens, G., and Weijers, D. (2012). Different auxin response machineries control distinct cell fates in the early plant embryo. Developmental Cell 22, 211-222.

Reddy, G.V., and Meyerowitz, E.M. (2005). Stem-cell homeostasis and growth dynamics can be uncoupled in the *Arabidopsis* shoot apex. Science 310, 663-667.

Reyes, J.C., Muro-Pastor, M.I., and Florencio, F.J. (2004). The GATA family of transcription factors in *Arabidopsis* and rice. Plant Physiology 134, 1718-1732.

Roeder, A.H.K., Chickarmane, V., Cunha, A., Obara, B., Manjunath, B.S., and Meyerowitz, E.M. (2010). Variability in the control of cell division underlies sepal epidermal patterning in Arabidopsis thaliana. PLoS Biol 8(5): e1000367. Doi:10.1371/journal.pbio.1000367.

Sabatini, S., Beis, D., Wolkenfelt, H., Murfett, J., Guilfoyle, T., Malamy, J.,Benfey, P., Leyser, O., Bechtold, N., Weisbeek, P., and Scheres, B. (1999). An auxin-dependent distal organizer of pattern and polarity in the *Arabidopsis* root. Cell 99, 463-472.

Sarkar, A.K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., Scheres, B., Heidstra, R., and Laux, T. (2007). Conserved factors regulate signaling in *Arabidopsis thaliana* shoot and root stem cell organizers. Nature 446, 811-814.

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

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

Smith, Z.R., and Long, J.A. (2010). Control of *Arabidopsis* apical-basal embryo polarity by antagonistic transcription factors. Nature 464, 423-427.

Steinmann, T., Geldner, N., Grebe, M., Mangold, S., Jackson, C.L., Paris, S., Gälweiler, L., Palme, K., and Jürgens, G. (1999). Coordinated polar localization of auxin efflux carrier PIN1 by GNOM ARF GEF. Science 286, 316-318.

Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M., and Zheng, N. (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature 446, 640-645.

Tanaka, H., Dhonukshe, P., Brewer, P.B., and Friml, J. (2006). Spatiotemporal asymmetric auxin distribution: A means to coordinate plant development. Cellular and Molecular Life Sciences 63, 2738-2754.

Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K., Taylor, N.E., Henikoff, J.G., Comai, L., and Henikoff, S. (2003).

Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Research 13, 524-530.

Tiwari, S.B., Hagen, G., and Guilfoyle, T. (2003). The roles of auxin response factor domains in auxin-responsive transcription. The Plant Cell 15, 533-543.

Ueda, M., Zhang, Z., and Laux, T. (2011). Transcriptional activation of *Arabidopsis* axis patterning genes *WOX8/9* links zygote polarity to embryo development. Developmental Cell 20, 264-270.

Ulmasov, T., Hagen, G., and Guilfoyle, T.J. (1999). Activation and repression of transcription by auxin response factors. Proceedings of the National Academy of Sciences 96, 5844-5849.

Ulmasov, T., Murfett, J., Hagen, G., and Guilfoyle, T.J. (1997). Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. The Plant Cell 9, 1963-1971.

Vanneste, K., Sterck, L., Myburg, A.A., Van de Peer, Y., and Mizrachi, E. (2015). Horsetails are ancient polyploids: evidence from *Equisetum giganteum*. The Plant Cell, tpc-15.

Wang, H., Ngwenyama, N., Liu, Y., Walker, J.C., and Zhang, S. (2007). Stomatal development and patterning are regulated by environmentally responsive mitogen-activated protein kinases in *Arabidopsis*. The Plant Cell 19, 63-73.

Whipple, C.J., Hall, D.H., DeBlasio, S., Taguchi-Shiobara, F., Schmidt, R.J., and Jackson, D.P. (2010). A conserved mechanism of bract suppression in the grass family. The Plant Cell 22, 565-578.

Wielopolska, A., Townley, H., Moore, I., Waterhouse, P., and Helliwell, C. (2005). A high-throughput inducible RNAi vector for plants. Plant Biotechnology Journal 3, 583-590.

Wysocka-Diller, J.W., Helariutta, Y., Fukaki, H., Malamy, J.E., and Benfey, P.N. (2000). Molecular analysis of SCARECROW function reveals a radial patterning mechanism common to root and shoot. Development 127, 595-603.

Yadav, R.K., Girke, T., Pasala, S., Xie, M., Reddy, G.V. (2009). Gene expression map of *Arabidopsis* shoot apical meristem stem cell niche. Proceedings of the National Academy of Sciences 106, 4941-4946.

Zhang, Y., Jiao, Y., Liu, Z., and Zhu, Y. (2015). ROW1 maintains quiescent centre identity by confining WOX5 expression to specific cells. Nature Communications, doi: 10.1038/ncomms7003.

Zhang, X., Zhou, Y., Ding, L., Wu, Z., Liu, R., and Meyerowitz, E.M. (2013). Transcription repressor HANABA TARANU controls flower development by integrating the actions of multiple hormones, floral organ specification genes, and GATA3 family genes in *Arabidopsis*. The Plant Cell 25, 83-101.

Zhao, Y., Medrano, L., Ohashi, K., Fletcher, J.C., Yu, H., Sakai, H., and Meyerowitz, E.M. (2004). HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in *Arabidopsis*. The Plant Cell 16, 2586-2600.

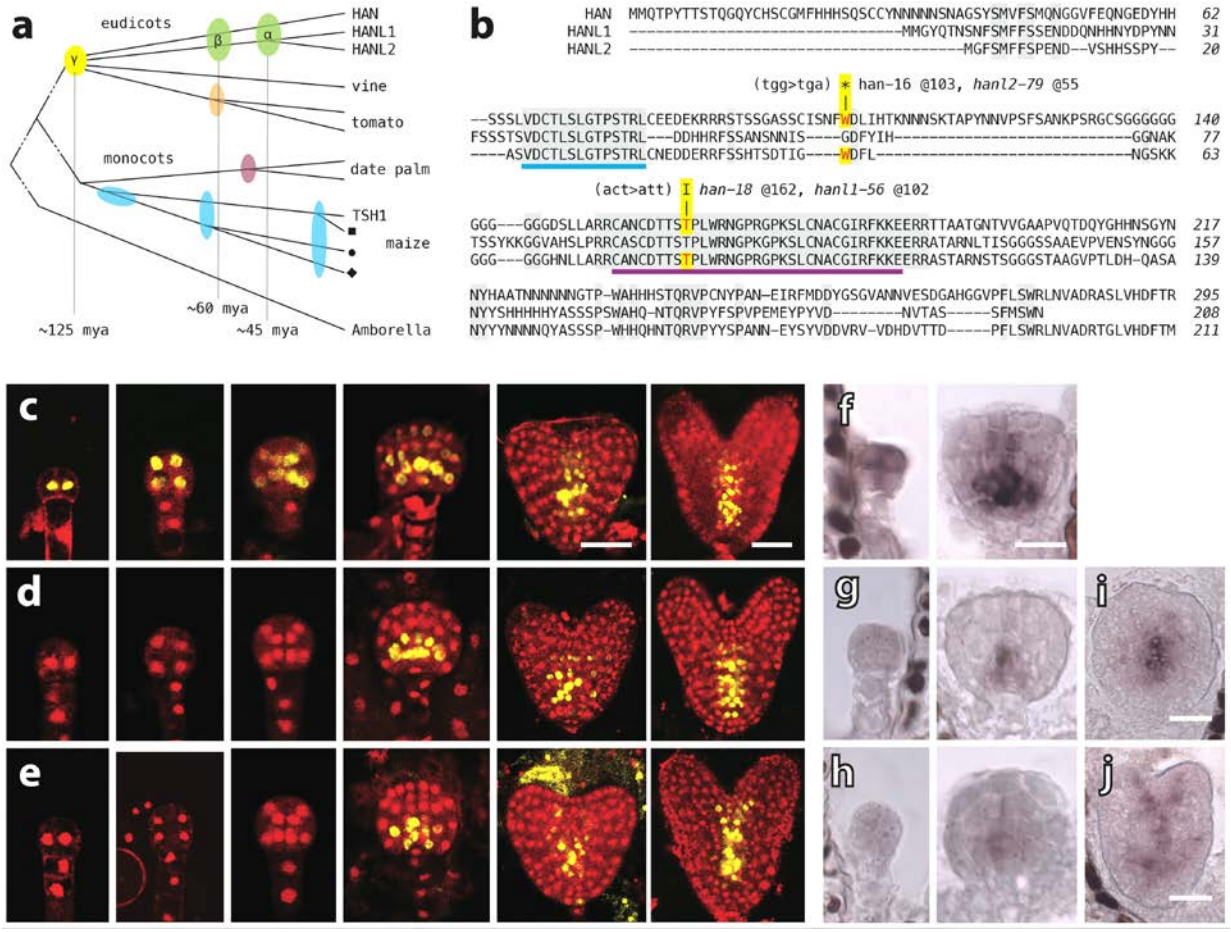

Figure 2.1. Divergence and expression of HAN family genes in *Arabidopsis*

(a) Evolutionary history of the HAN family in representative angiosperms. Arabidopsis HAN genes reside in syntenic chromosome segments that originated in the α WGD (~45 mya ago; split of HANL1 and HANL2) and the γ triplication event at the base of the eudicots (~125 mya ago; Reyes et al., 2004; Jiao et al., 2012). Due to rapid radiation after this event, the phylogenetic relationships between HAN genes of eudicot clades remain poorly resolved. Only a single gene was retained in the vine and tomato lineages; a later WGD (~60 mya) gave rise to the two genes of tomato. Maize contains four HAN genes likely dating to WGD events at the base of the grasses (Paterson et al., 2004; timing of this event is still debated), ~65 mya ago, and 15 mya ago. The two genes of date palms were separated in a WGD ~55 mya ago. Mapping and timing of WGD events follows Vanneste et al. (2014), basal branches are not to scale; the Phytozome and NCBI portals were used for sequence searches; gb identifiers: CBI17512 (vine);

XP_004233721 & XP_004233970 (tomato); XP_008800719, XP_008805393 (date palm); XP_008650055, XP_008656064, XP_008675252, & XP_008657512 (maize *TSH1*, ■, •□◆)
(b) Alignment of *Arabidopsis* HAN proteins. Identical positions shaded in gray; little sequence conservation is apparent outside the HAN motif (Whipple et al., 2010; underlined in cyan) and the GATA Zn-finger domain (CDD record cd00202; underlined in purple). Substitutions found in mutant alleles are highlighted in yellow and detailed above the listing.
(c-e) Expression of HAN family genes during embryogenesis. HAN expression (c, top row) is detected in the apical cell and its daughters until the globular stage, when HAN accumulates in cells adjacent to the proembryo boundary and subsequently, the central provasculature. Alternatively, HANL1 (d, middle row) and HANL2 expression (e, bottom row) is undetectable prior to the globular stage, however from the globular stage their expression closely mimics HAN. Gene expression domains were validated by in situ hybridizations (f,g,h; Appendix A, Fig. S2.1a), which were also used to show the expression of HANL1/2 in *han-16* at the bent cotyledon stage (i,j).

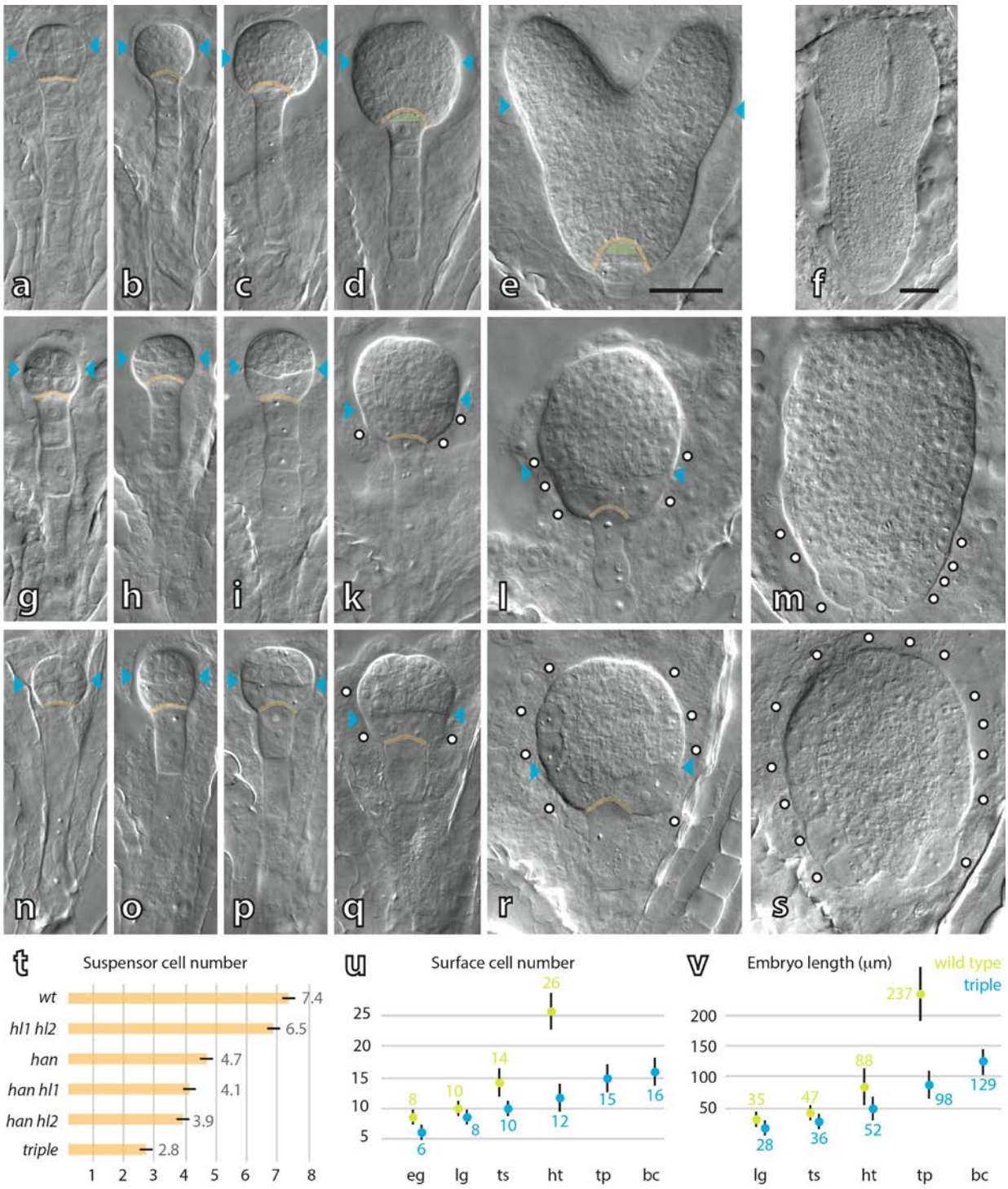
a
b
c
d
e
f
g
h
i
k
l
m
n
o
p
q
r
s
t
Suspensor cell number
wt
hl1 hl2
han
han hl1
han hl2
triple
7.4
6.5
4.7
4.1
3.9
2.8
1 2 3 4 5 6 7 8
u
Surface cell number
25
20
15
10
5
26
14
10
8
8
6
12
15
16
eg lg ts ht tp bc
v
Embryo length (µm)
wild type
triple
200
150
100
50
237
88
35
47
98
129
28
36
52
lg ts ht tp bc

Figure 2.2. *han hanl1 hanl2* mutants arrest as blimp-like structures with abnormally enlarged surface cells

(a–s): Nomarski images of cleared whole-mount seed containing wild type (top row), *han* (middle row) and triple mutant embryos (bottom row) at the 8-cell (a,g,n), 16-cell (b,h,o), early globular (c,i,p), late globular (d,k,q), heart (e,l,r) and torpedo stage (f,m,s); the boundary between suspensor and proembryo is highlighted in orange, and the boundary between upper and lower tier by blue arrowheads; the lens-shaped daughter produced by the hypophysis cell at the late globular stage and the root quiescent center are highlighted in green (d,e); enlarged surface cells at the base of *han* embryos and surrounding triple mutant embryos are marked by white dots; all images to scale with exception of (f); scale bars in (e, f), 30 μm.

(t–v): Effect of HAN genes on cell division and overall growth; loss of HAN genes results in a reduced number of suspensor cells at the 32-cell stage (t); triple mutant embryos contain fewer surface cells along the median section, arresting with a similar number as wild type at the transition stage (u); their length is also diminished, reaching about half the length of wild type torpedo embryos (v); imaging as above was used as a basis for the analysis, and mutants were staged according to their wild type siblings from the same silique; averages of >100 (t) or >50 (u,v) measurements are shown, black bars represent double standard error; abbreviations: *hl1*, hanl1; *hl2*, *hanl2*; eg, early globular; lg, late globular; ts, transition; ht, heart; tp, torpedo; bc, bent cotyledon.

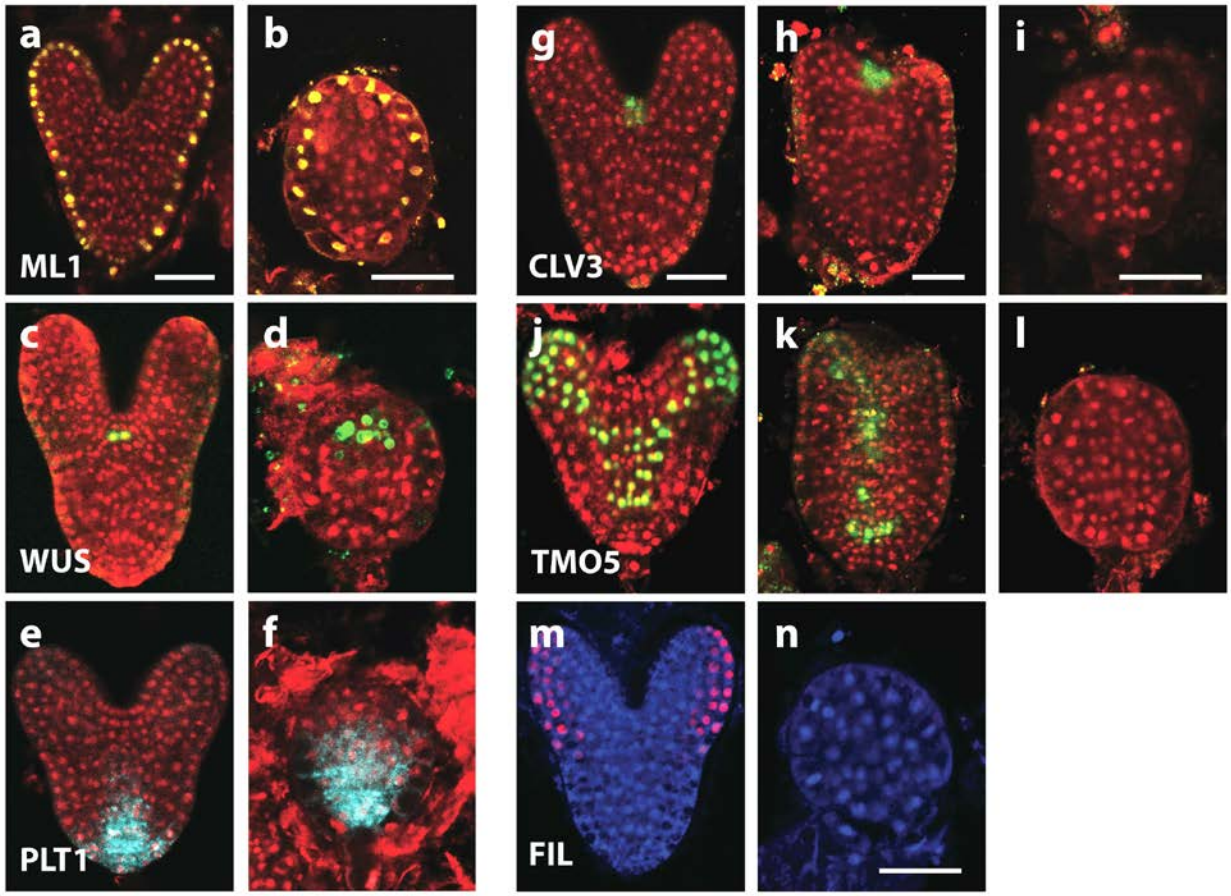

Figure 2.3. Expression of genes regulating embryo polarity, tissue formation, and organogenesis in triple mutant embryos. Regulators of epidermal development, *ML1:H2B-YFP* (Roeder et al., 2010), and regulators of the shoot and root apical meristems with an early onset of expression, *pWUS*::*mGFP5-ER* (Yadav et al., 2009) and *pPLT1:CFP* (Galinha et al., 2007), are expressed in similar overall patterns in wild type (a,c,e) and triple mutants (b,d,f) at the torpedo stage. In contrast, factors contributing to shoot meristem maintenance with a late onset of expression, pCLV3**::**mGFP5-ER (Reddy & Meyerowitz, 2005), a factor promoting provascular fate, *pTMO5::3XGFP* (De Rybel et al., 2014), and a component of the lateral organ polarity pathway, *pFIL::dsRed-N7* (Heisler et al., 2005), are absent in mutants at all stages of development (i,l,n). All scale bars are 30 μm.

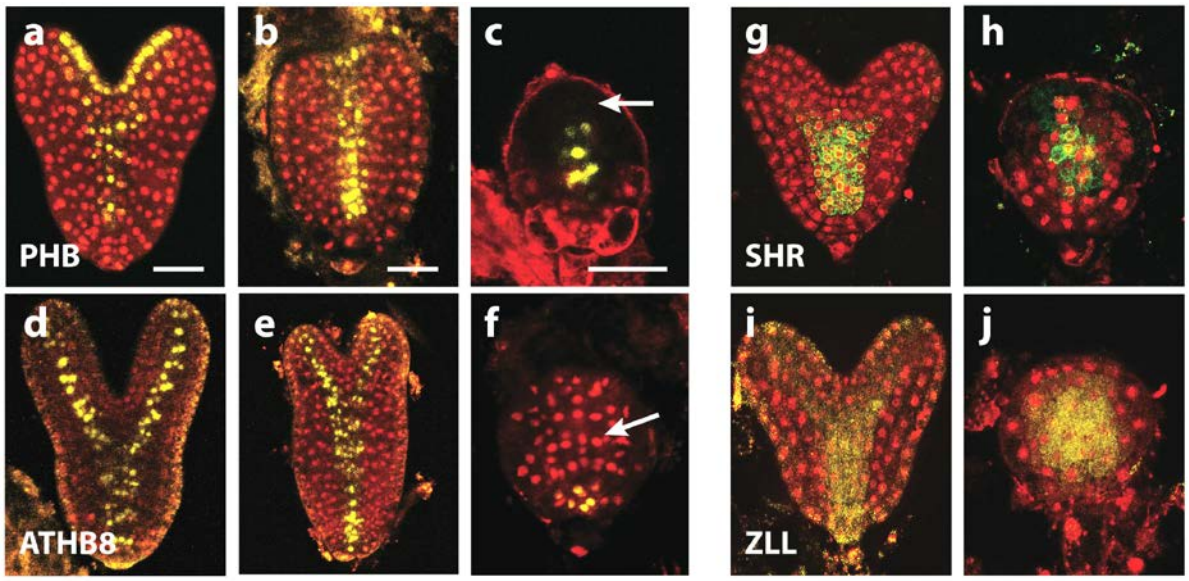

Figure 2.4. Genes expressed in the provascular domain. The expression of PHB-2xYFP and ATHB8::YFP (Yadav et al., 2014) in the central domain of torpedo embryos is similar in wild type (a,d), *han-16* (b,e), and triple (c,f) mutant embryos. However, whereas PHB is also expressed in surface cells between cotyledon primordia of wild type, expression at the embryonic apex is missing in triple mutants (c, see arrow). In addition, triple mutants occasionally express ATHB8 only at their very base (f, see arrow). The expression of *pSHR:GFP* (Helariutta et al., 2000) and *ZLL-YFP* (Breuninger et al., 2008) is also similar in wild type (g,i) and triple mutants (h,j) at the torpedo stage. Scale bars are 30 μm.

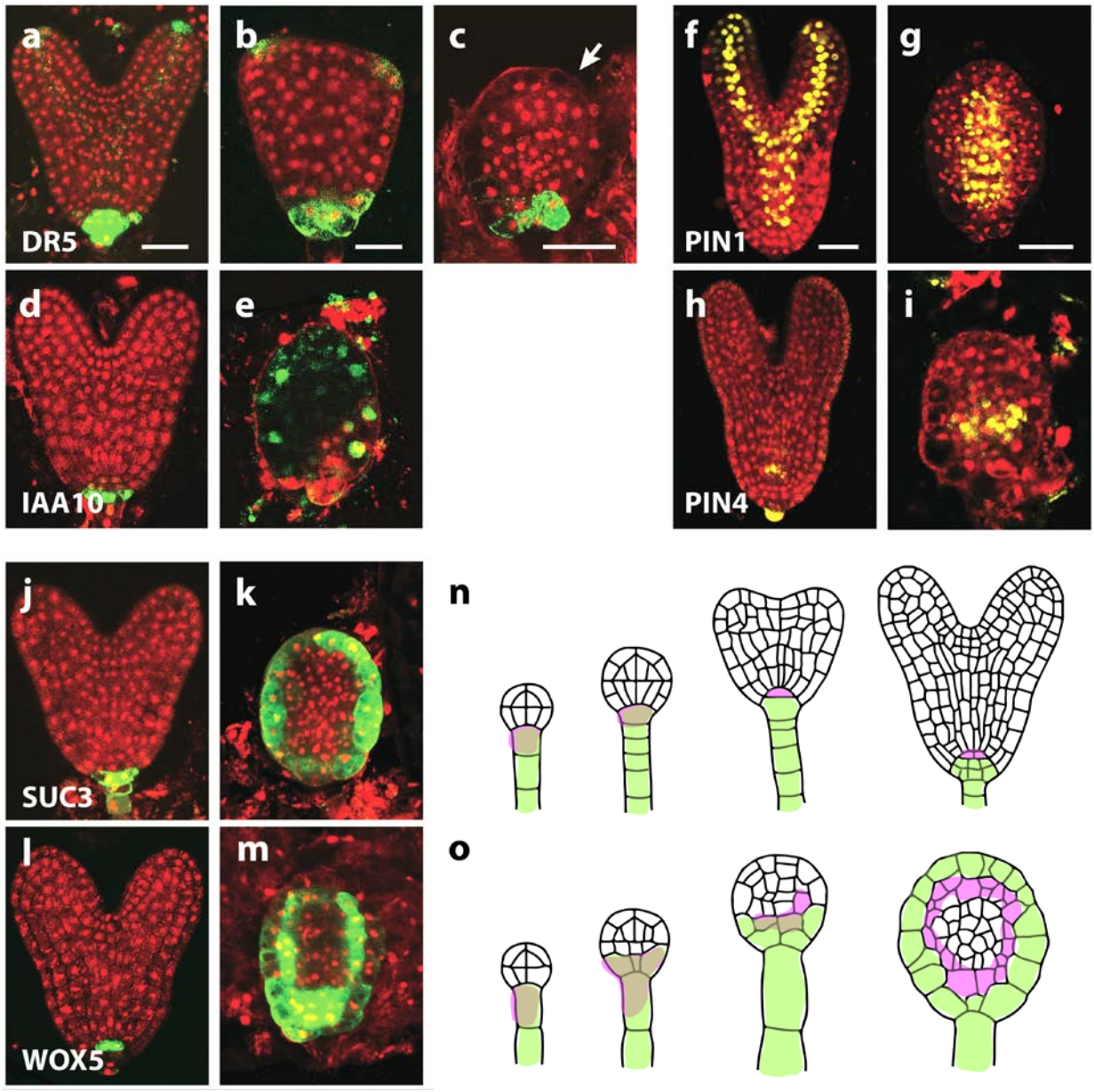

Figure 2.5. Auxin responses and radialization of the apical/basal axis.

The presence of an auxin response maximum, as visualized by *DR5rev-GFP* (Friml et al., 2003), is detected at the base of embryos in wild type (a), *han-16* (b), and triple mutants (c) at the torpedo stage; however, triple mutant embryos never express DR5 at the embryo apex, as occurs during the initiation of primordia (c, see arrow). The expression domain of *pIAA10::3xGFP-NLS* (Rademacher et al., 2012) is expanded in triple mutants, including surface cells up to the apical pole (d,e). The auxin transporter PIN::PIN1-Venus-NLS is expressed in the central domain of triple mutants (g), and in the provascular of wild type (f). PIN4::PIN4-NLS, confined

to the QC of wild type embryos (h), expands to a broader domain between the upper and lower tier in triple mutants (i, similar to *han* mutants). Like IAA10, the expression domain of *SUC3::GFP* (Meyer et al., 2004) and *pWOX5::GFP* (Blilou et al., 2005) is dramatically expanded in triple mutants. SUC3 expression is detected in enlarged surface cells from the base of embryos to the apex (j,k), while WOX5 expression is strongly expressed in sub-surface cells in a concentric ring (l,m). Schematic representation showing the effect of *han hanl1 hanl2* on gene expression; domains confined to the apical and basal daughter of the hypophysis, and thus delimited by the proembryo boundary in wild-type (n), expand into all surface and subsurface cells in triple mutants (o). The apical daughter of the hypophysis normally gives rise to the QC (pink) whereas the basal daughter is the precursor of the root columnella (green). Scale bars are set to 30 μm.

CHAPTER 3

STRUCTURAL BASIS OF HAN FUNCTION IN *ARABIDOPSIS THALIANA* EMBRYOGENESIS[1]

[1] Volny, M.F., Gendrau, J., and Lukowitz. To be submitted to Plant Physiology

Abstract

Transcription factors of the GATA family facilitate plant responses to their environment and regulate developmental processes. The defining feature of this family is a single highly conserved zinc finger, the sequence of which can be used to identify ancient sub-clades. Some family members also share other structural features, yet it remains unclear what biochemical role they might serve. Additionally, we still lack a comprehensive understanding of functional overlap between closely and less closely related GATA factors, a problem that has hampered mutant analyses. We have previously characterized the contributions of three *Arabidopsis* GATA factors containing a B-class zinc finger as well as a short, conserved feature named the HAN domain to embryogenesis. These factors, HANABA TARANU (HAN), HAN-LIKE1 (HANL1), and HANL2, can provide equivalent activity in positioning the inductive boundary between the proembryo and suspensor, at which the root initiates. To establish the structural basis of HAN function and elucidate phylogenetic relationships within the GATA family, we have tested all Arabidopsis B-class GATA factors, as well as chimeric and mutant variants for their ability to rescue the mutant phenotype of embryos lacking HAN function. Expression of GATA29 fully complemented mutant embryos, while GATA21/GNC was partially effective. GATA29 is not closely related to HAN genes, and its activity is dependent on the presence of an N-terminal segment conforming to the consensus of an EAR motif. Clear orthologs of GATA29 can only be found within the rosids suggesting relatively recent neo-functionalization. Our results further reveal that loss or mutation of the HAN domain abolishes HAN activity; that the HAN domain can be replaced with the EAR motif of an unrelated protein, IAA17; and that adding a HAN domain to the N-terminus of GATA16 creates chimeras with equivalent function to HAN. EAR motifs encode short amphiphilic segments found in a large number of plant

transcription factors and function to bind co-repressors of the TOPLESS (TPL) family. Similarly, the HAN domain mediates interaction with TPL proteins in a yeast two-hybrid assay. Taken together, our results suggest that HAN function requires the presence of a B-type zinc finger and a repressor domain, implying that HAN mediates down-regulation of its target genes.

Introduction

GATA factors are an evolutionarily conserved family of transcriptional regulators that recognize the consensus sequence WGATAR (W=T, A; R=G, A) (Ko and Engel, 1993; Lowry and Atchley, 2000; Reyes et al., 2004). Plants have substantially more GATA factors than other eukaryotic lineages, with 30 in Arabidopsis and 28 in rice, suggesting a high level of diversification (Lowry and Atchley, 2000; Reyes et al., 2004). Members of the GATA family in *Arabidopsis* have a single type IV zinc finger with the consensus sequence C-X_2-C-X_{17-20}-C-X_2-C (C=cysteine; X=any residue), often followed by a basic domain (Reyes et al., 2004). In contrast, the GATA factors of fungi and animals typically have two Zn-fingers (Lowry and Atchley, 2000; Scazzocchio, 2000). The binding of target DNA is facilitated by hydrophobic interactions between the zinc finger and the major groove, while the basic stretch contacts the negatively charged phosphate backbone (Omichinski et al. 1993; Starich et al., 1998). GATA factors in flowering plants have been subdivided based on their zinc finger domains into four classes, class A through class D; it seems likely that these clades are of more ancient origin within the land plants (Reyes et al., 2004), but a more comprehensive phylogeny of the family is still lacking. Comparisons of GATA genes from plants and other eukaryotic lineages, such as fungi and animals, indicate that the plant genes are monophyletic (Reyes et al., 2004).

The role of plant GATA factors in environmental responses is well established. The promoters of light and circadian clock controlled genes are enriched for GATA motifs (Giuliano et al., 1988; Lam and Chua, 1989; Buzby et al., 1990; Donald and Cashmore, 1990; Kehoe et al., 1994), and the CCT domain (named for CONSTANS, CONSTANS-LIKE, and TOC1) of C-class GATA factors is found in light-response regulators (Robson et al., 2001; Griffiths et al., 2003; Reyes et al., 2004; Gendron et al., 2012). The transcription of *Arabidopsis* GATA factors

responds to the duration and quality of light, nutrients, the circadian cycle, and phytohormones (Manfield et al., 2007; Luo et al., 2010; Richter et al., 2010; Naito et al., 2007). Mutational analysis suggests that many A-class GATA factors are involved in adjusting growth in response to environmental cues; in particular: GATA2 and GATA4, which promote de-etiolation (Luo et al., 2010); BLUE MICROPYLAR END 3 (*BME3)/*GATA8, which reinforces the breaking of dormancy (Liu et al., 2005); and the B-class genes, GATA, NITRATE-INDUCIBLE, CARBON-METABOLISM INVOLVED (GNC)/GATA21 and GNC-LIKE (GNL)/GATA22, which facilitate chlorophyll biosynthesis, carbon/nitrogen metabolism, and regulate flowering time (Bi et al., 2005; Mara and Irish, 2008; Chiang et al., 2012; Richter et al., 2010, 2013).

Some B-class GATA factors have also been implicated in developmental patterning events. GATA23, specifies the founder cells of lateral root primordia and is required for root branching (De Rybel et al., 2010). HANABA TARANU (HAN)/GATA18 and two closely related paralogs, HAN-LIKE 1 (HANL1)/GATA20 and HAN-LIKE 2 (HANL2)/GATA19, regulate root initiation in the embryo as well as development of the inflorescence and flower meristems (Zhao et al., 2004; Nawy et al., 2010). Loss of *han* function results in reduced shoot meristems, fused stamens and carpels, and fewer floral organs (Zhao et al., 2004), while the concomitant loss of all three genes results in partially radialized embryos that arrest without forming a shoot, root, or provasculature (chapter 2 of this dissertation). Three HAN genes of grasses, maize TASSEL SHEATH 1 (TSH1), rice NECK LEAF 1 (NL1), and barley THIRD OUTER GLUME (TRD), have been implicated in the suppression of bract outgrowth (Wang et al., 2009; Whipple et al., 2010). Several other HAN paralogs are present in these grasses but their function has not been reported. The common feature of HAN genes is the presence of a conserved, 14 amino acid stretch N-terminal of the zinc finger called the HAN domain (Whipple

et al., 2010; Behringer et al., 2014). Although the HAN domain is virtually identical in the HAN genes of angiosperms and has been recognized as important for function (Whipple et a., 2010; Wang et al., 2009; Kanei et al., 2012), a molecular activity has not been assigned to it.

Animal GATA factors often act combinatorially, in complexes with other GATA factors or interacting proteins, to control cell fates and differentiation pathways (Patient and McGhee, 2002; Chlon and Crispino, 2012). Plants however, lack gene homologues for many of the GATA-interacting factors of animal regulatory networks (Tsang et al., 1997; Fox et al., 1999), which make it difficult to extrapolate combinatorial models to plants. Relatively few GATA factors of plants are associated with a visible mutant phenotype, and attempts to investigate their genetic interactions have provided results that are difficult to interpret. For example, neither loss of GNC, GNL or both affects embryonic development, but many *han gnc* and *han gnl* embryos arrest as disorganized clusters of cells (Zhang et al., 2013) – is this an additive effect, or do these genes co-regulate the same developmental pathway? Interestingly, the expression of GNC, GNL, and HANL2 is responsive to the loss of *han* and HAN was found to occupy the promoter of HANL2 (Zhang et al., 2013). Additionally, HAN interacts with GNC, GNL, and HANL2 in yeast two hybrid assays, and in bimolecular fluorescence complementation experiments in tobacco leaves (Zhang et al., 2013), which suggests the possibility of cross-regulatory interactions.

We have previously shown that the HAN-family of B-class GATA factors are biochemically equivalent in positioning the inductive boundary at which the root forms in the embryo. Our aim here is to define the structural basis of HAN function and delineate it from the function of other factors in the B-class. To assess the equivalency of these factors, we tested each B-type protein, in addition to mutant and chimeric variants, for their ability to complement

the *han hanl1 hanl2* mutant phenotype. We found that GATA29 expression fully complemented mutant embryos, while GATA21/GNC was incompletely effective. GATA29 is structurally distinct, not closely related to members of the HAN family, and likely the result of an evolutionarily recent neo-functionalization event. The activity of GATA29 requires an N-terminal EAR motif, as deleting or mutating it makes the protein inactive. We found that HAN is inactivated by mutating or deleting the HAN domain, and that this domain can be replaced by the EAR motif from the Aux/IAA protein, IAA17. Addition of a HAN domain is also sufficient to confer HAN function to GATA16, which contains a B-type zinc finger but no other recognizable domains. Finally, we show that the HAN domain facilitates interactions between HAN and TPL family proteins in yeast two hybrid assays. These results indicate that HAN activity requires a B-class zinc finger and a repressor domain and suggest that HAN functions as a transcriptional repressor.

Results

We previously generated a T-DNA construct for expressing any coding sequence of interest under the control of HAN regulatory sequences; this construct contains the entire HAN locus, extending from ~6.8 kbp upstream of the translational start site to ~1 kbp downstream of the translational stop site, with restriction sites introduced at both ends of the coding sequence (Fig. 3.1a; NcoI at the start codon: ccATGg; BglII and HindIII at the stop codon: agatcTAAgctt; chapter 2 of this dissertation). Expression of a nuclear fluorescent protein faithfully reflects HAN transcription (chapter 2 of this thesis), and the expression of a HAN cDNA fully complemented the phenotype of han mutant embryos, indicating that neither the absence of the single intron of HAN nor alterations due to the introduction of cloning sites had a negative

impact. Essentially the same T-DNA construct was used here to assess the function of B-type GATA factors as well as mutant and chimeric variants of HAN, GATA29 and GATA16. All T-DNAs were transformed into han/+ hanl1 hanl2 plants, and the embryos of at least 10 independent primary transgenic lines were examined by DIC microscopy. Triple mutant embryos are normally produced at a rate of 25% upon self-fertilization and show a striking arrest phenotype that can be reliably scored (compare Fig. 3.1b to c). Molecular complementation, i.e. rescue of this phenotype in the transgenic lines, provided our assay for protein function (see methods for details). HANL1 and HANL2 fully complement the phenotype of triple mutants when expressed from this construct (Appendix B, Table S3.1); this was as expected, since we had previously shown that similar constructs complement han single mutants (chapter 2 of this thesis). Substitutions in the zinc finger domain of Arabidopsis HAN (*han-2*, *-18*, and *-21* alleles; Zhao et al., 2004; Nawy et al., 2010) are associated with strong mutant phenotypes, underscoring the functional importance of this domain. The only other conserved feature of HAN proteins is the HAN domain; a substitution in this domain is found in the *Arabidopsis han-30* allele (Kanei et al., 2012), which is comparatively weak and has no noticeable effect on embryogenesis.

GATA29 can substitute for HAN function in embryogenesis

To assess whether any of the other 11 Arabidopsis GATA factors can provide equivalent function to HAN in the embryo, we amplified their coding sequences from genomic DNA and expressed them under the control of HAN regulatory sequences. None of them complemented the phenotype of triple mutant embryos, with the exception of GATA21/GNC and GATA29 (Fig. 3.1a; Appendix B, Table S3.1). Plants expressing the GATA21/GNC T-DNA produced a large number of embryos which appeared neither similar to wild-type nor triple mutants. Instead,

the embryos showed variable intermediate phenotypes that were somewhat reminiscent of *han* single mutants: some remained small, but showed a flat apex or a small indentation marking the appearance of cotyledons (Fig. 1d), while others developed an elongated hypocotyl and larger cotyledon primordia (Fig. 2e). In contrast, the rescue activity of GATA29 was indistinguishable to what we observed for HAN and HAN-like genes. We conclude that GATA21/GNC has related but not equivalent function to HAN proteins, and GATA29 provides equivalent function.

Loss of the HAN domain abolishes HAN function

GATA factors with a B-class zinc finger and a HAN domain can be found in the genomes of all land plants for which sequence information is available (see Fig. 3.2a for representative examples and consensus). Prominent characteristics of the domain are an invariant cysteine and a series of regularly spaced hydrophobic residues (VxCxLxLxT; the first leucine residue is mutated in Arabidopsis *han-30*, and the second in maize *tsh-1* and rice *nl1*; Wang et al., 2009; Whipple et al., 2010; Kanei et al., 2012). However, GATA29 does not contain a HAN domain, calling into questions its relevance for HAN function in the embryo. To experimentally assess its importance, we generated HAN variants in which the cysteine and two of the hydrophobic residues of the HAN domain were substituted or the entire domain deleted (Fig. 3.2e; C69>A, L71>A, and L73>A). These variants were inactive, confirming that loss of the HAN domain abolishes HAN function (Appendix B, Table S3.1).

We next investigated the phylogenetic origin of GATA29 and its relationship to HAN genes with the hope of uncovering potential cues for understanding the basis of their similar activity. The GATA29 zinc finger domain is unique in that the first pair of cysteine residues is separated by four instead of two positions, indicative of an insertion event (Reyes et al., 2004).

Reciprocal BLAST searches identified GATA29 orthologs harboring the same insertion in cacao and poplar but no such proteins were found outside of the malvids or eurosids II (Fig. 3.2c). The grape vine genome contains an apparent ortholog lacking the insertion, suggesting relatively recent diversification. GATA29 and HAN proteins show limited similarity outside of the zinc finger, and the GATA29 clade stands isolated from HAN genes or other B-class GATA factors in a phylogenetic analysis, such that its origin remains unresolved (Fig. 3.2d; see Appendix B, Fig. S3.1, S3.2, S3.3 for more detailed representation). However, we found that a short N-terminal segment of GATA29 conserved in Arabidopsis, cacao and poplar but absent from the vine ortholog was often aligned to the HAN domain in the course of this analysis (Fig. 3.2a), perhaps more likely due to incidental similarity than a shared origin. This segment conforms to the consensus of an ethylene response factor-associated amphiphilic (EAR) motif, minimally defined by the sequence LxLxL (Ohta et al., 2001; Fig. 3.2b).

The HAN domain is equivalent to an EAR motif

EAR motifs are found in many plant proteins, including members of the Aux/IAA family of transcriptional repressors (Kagale et al., 2010; Ohta et al., 2001; Hiratsu et al., 2003, 2004; Hill et al., 2008; Causier et al., 2012), and function by binding co-repressors of the TOPLESS (TPL) or SAP18 family (Song and Galbraith, 2006; Hill et al., 2008; Szemenyei et al., 2008; Causier et al., 2012). Interestingly, HAN has been identified as a potential binding partner of Arabidopsis TPL in a large-scale interaction study (Causier et al., 2012). A shared ability to recruit co-repressors may explain the similar properties of GATA29 and HAN proteins in our assay. To test this idea, we generated HAN variants containing the well-characterized EAR motif of IAA17 in place of the HAN domain; as a control, an inactive EAR motif, in which the conserved leucine

residues were substituted with alanine residues (Fig. 3.2b; Tiwari et al., 2001, 2004), was also introduced. The chimeric protein with a wild type EAR motif fully complemented the phenotype of triple mutant embryos, while no rescue was observed with the mutant EAR motif (Fig. 3.2e; Appendix B, Table S3.1). We also introduced a HAN domain to the N-terminal end of GATA16, which contains a B-class zinc finger but no other recognizable structural feature and had previously failed to complement triple mutants; and replaced the segment of GATA29 containing an EAR motif with a HAN domain. In both cases, the chimeric proteins fully rescued the phenotype of triple mutant embryos (Fig. 3.2e; Appendix B, Table S3.1). These results demonstrate that the HAN domain is functionally equivalent to an EAR motif and further suggests that any GATA factor with a B-class zinc finger and a repressor domain will have similar properties to HAN.

The HAN domain mediates interactions with TPL proteins in yeast

Arabidopsis TPL and TOPLESS-RELATED (TPR) proteins promote transcriptionally inactive states of chromatin by interacting with histone deacetylases (Long et al., 2006; Zhu et al., 2010; Krogan et al., 2012) as well as disrupting components of the mediator complex (Malave´ and Dent, 2006; Gonzalez et al., 2007). TPL proteins are recruited to their target genes by their affinity for the EAR motifs of DNA-binding transcription factors (Szemenyei et al., 2008; Causier et al., 2012). To determine whether the HAN domain serves a similar function, we investigated protein-protein interactions between HAN variants and members of the TPL family in a yeast two-hybrid assay. Arabidopsis TPL and TPR proteins were fused to the strong transcriptional activation domain of Herpes simplex viral protein 16 and combined with a series of potential interactors fused to the GAL4 DNA-binding domain: the N-terminal end of HAN

(from methionine to the beginning of the zinc finger); the mutated form of the HAN N-terminus (same substitutions as described above); the wild type HAN domain (S^{64}-D^{84} of the amino acid sequence); and the mutated form of the HAN domain. As a control, both the wild type and mutated forms of the EAR motif of IAA17 (M^{2}-G^{29} of the amino acid sequence; same substitutions as described above) provide a point of comparison. We then assessed the strength of individual interactions as growth of the respective yeast strain on selective media (see methods for details).

Proteins containing the HAN domain showed robust interaction with all members of the TPL family except for TPR3 (Fig. 3.3). The HAN domain by itself was sufficient, with interactions slightly weaker in strength than those involving the longer N-terminal HAN segment. Control experiments with an unrelated human protein indicated that the longer N-terminal HAN segment caused some auto-activation; the addition of 50 mM 3-amino-1,2,4 triazole (3-AT), a competitive inhibitor of the HIS3 selective marker, was required to suppress this activity. In contrast, proteins containing the HAN motif alone or the IAA17 EAR motif generated little or no autoactivation, even at the lowest concentrations of 3-AT. Interestingly, interactions with the HAN domain appeared substantially stronger than interactions between the EAR motif of IAA17 and TPL family proteins. Mutating the HAN domain or the EAR motif disrupted all interactions between HAN and TPL family members (Fig. 2). We conclude that the HAN motif is sufficient for mediating interactions with TPL proteins in yeast, and that the hydrophobic amino acids at its core are required for this interaction.

Discussion

Transcription factors of the GATA family play diverse roles in mediating environmental responses and regulating plant development, yet little is known about functional relationships between them, or the biochemical activity of their structural features outside of the zinc finger. To address these questions, we conducted a comprehensive survey assessing the functional equivalency of B-class GATA proteins and the structural requirements for HAN function in embryonic patterning. We found that GATA29 expression fully complements *han hanl1 hanl2* mutants, and GATA21/GNC provides partial rescue. The GATA29 zinc finger is unusual in that the first two cysteine residues are separated by 4 instead of 2 positions; this configuration does not appear to affect gene function. A similar case is presented by the *Caenorhabditis elegans* GATA factor END-1, which has an equivalent insertion of the zinc finger, yet retains its ability to recognize DNA motives (Shoichet et al., 2000).

The B-class clade forms a deep branch of the family, predating the evolution of flowering plants. However, few domains are recognizable outside of the zinc finger of these genes and sequence similarity is limited. A phylogeny of Arabidopsis and rice GATA factors suggests that relationships within the B-class genes are difficult to resolve (Reyes et al., 2004). Similarly, only three groups were consistently supported in our own phylogeny of a broader sample of species: the HAN clade, the GNC/GNL clade, and a clade consisting of GATA29 genes from Arabidopsis, cacao, poplar, and vine; the relationships between these clades is open, and many other GATA factors showed no clear alliance (Fig. 3.2d; Appendix B, Fig. S3.1, S3.2, S3.3). In addition, gene loss and gene duplication events appeared to be frequent, suggesting rapid and ongoing diversification. The GATA29 clade is a good example: no members were identified in most of the species included in our analysis, pointing to a strong possibility of rapid evolution

and/or gene loss; in particular, two prominent features, an insertion in the zinc finger and a consensus EAR motif, were only found in the genes of *Arabidopsis*, cacao, and poplar but absent from the vine gene, suggesting their recent origin.

The partial activity of GNC in our molecular complementation experiments was characterized by the occurrence of variable, intermediate phenotypes, reminiscent of *han* single mutant embryos. GNC lacks a canonical LxLxL EAR motif or the less common DLNxxP variant; however, GNC has been shown to interact with the co-repressor protein SIN3-LIKE 1 (SNL1) in yeast (Bowen et al., 2010). SNL1 is expressed in the developing seed and like TPL family members, recruits HISTONE DEACETYLASE 19 to convert chromatin from an active to an inactive state (Krogan et al., 2012; Long et al., 2006; Zhu et al., 2010; Wang et al., 2013). Thus, GNC appears to have similar biochemical properties as conferred by an EAR motif. Given its similar zinc finger, the interaction of GNC with SNL1 opens the possibility of functional overlap with HAN genes in the embryo and may explain why the loss of both HAN and GNC results in stronger embryo defects than the loss of *han* alone (Zhang et al., 2013). However, stronger mutant phenotypes have also been observed by loss of *han* in combination with *gnl*, which does not have similar activity to HAN in our assay and has not been reported to interact with SNL1. GNC and GNL have overlapping functions in regulating chlorophyll biosynthesis and are predominately expressed in leaves; with no embryonic phenotypes reported for single or double mutants (Bi et al., 2005; Chiang et al., 2012). Thus, the genetic interactions between GNC, GNL, and HAN and the importance of GNC for development of the embryo remain difficult to interpret.

Our results show that the HAN domain can be replaced with an EAR motif from IAA17, shown to mediate interactions with TPL family co-repressors in yeast (Tiwari et al., 2004);

strongly suggesting that HAN functions as a transcriptional repressor. This view is consistent with microarray studies of gene transcription in response to induced over-expression of HAN (Zhang et al., 2013). It is also compatible with our interpretation of the developmental defects observed in *han* single and *han hanl1 hanl2* multiple mutant embryos: loss of *han* results in a coordinated apical shift of gene expression patterns at the boundary between suspensor and proembryo, with genes normally transcribed in the suspensor expanding into the lower tier of the proembryo and genes normally transcribed throughout the proembryo becoming confined to the upper tier (Nawy et al., 2010); loss of all three members of the *han* family results in blimp-shaped embryos, in which several genes normally restricted to the basal side of the proembryo boundary – marking either the distal root tip or the root quiescent center – over time become expressed in concentric layers of either surface or subsurface cells. On the basis of these findings, we have argued that HAN is required to position the inductive boundary between suspensor and proembryo, across which the root is initiated, and that the radialization of gene expression patterns observed in triple mutants is caused by a loss of this boundary (chapter 2 of this dissertation). According to this view, the role of HAN in the embryo is somewhat analogous to the role of lateral organ polarity genes in leaf development. Leaf polarity is maintained by antagonistic interactions of KANADI (KAN) and HD-ZIP III genes. For example, loss of KAN genes results in radially symmetric leaves with adaxial features covering the entire surface. This effect is due to the expansion of HD-ZIP II function, normally confined to the adaxial surface of developing leaves, to the abaxial domain (Eshed et al., 2001, 2004). Similarly, loss of HD-ZIP III function results in radially symmetric leaf primordia with abaxial characteristics (Emery et al. 2003; Prigge et al., 2005), presumably for the opposite reason. By analogy, HAN functions to repress the expression of genes normally active in the suspensor or the basal side of the

proembryo boundary in its own expression domain, the apical side of the proembryo boundary. It is unclear whether HAN antagonizes a positive regulator of basal fates in this context, as abaxial and adaxial factors antagonize each other in the leaf, or by directly repressing a number of relevant targets.

The radialized phenotype of *han hanl1 hanl2* mutants also share similarities to mirror-image duplications of the root caused by the temperature sensitive dominant-negative *tpl-1* allele (Long et al., 2002; Long et al., 2006; Smith and Long, 2010). Combining a weak *han* allele in which a single leucine residue within the HAN domain is mutated, with an *angustifolia 3* (*an3*) allele results in seedings with malformed cotyledons and occasionally, an ectopic root at the position where the shoot should form (Kanei et al., 2012). As *an3* mutants typically have rather subtle defects, such as narrow leaves comprised of fewer cells (Horiguchi et al., 2011; Kanei et al., 2012), and the *han-30* allele causes no visible defects in embryos, their combined effect reflects a rather striking synergism. Intriguingly, our characterization of the HAN domain implies that the *han-30* protein has reduced affinity to TPL family co-suppressors, suggesting a mechanistic link between the effects of *tpl-1* and *han-30 an3*. Indeed, ectopic expression of PLT genes in the apical domain of the embryo is associated with the double root phenotype in both cases (Smith and Long, 2010; Kanei et al., 2012). Our studies of *han* and *han hanl1 hanl2* mutants showed that PLT1 is not regulated by HAN, such that is unclear why PLT expands to the apex of *han-30 an3* embryos. However, similarities between the radialization caused by *han hanl1 hanl2* and the mirror-image duplication caused by *tpl-1* do exist. For example, markers for apical fate are missing in both backgrounds (SHOOTMERISTEMLESS in *tpl-1*; Long et al., 2002; CLV3 and FIL in *han hanl1 hanl2*; chapter 2 of this thesis). Additionally, *tpl-1* as well as *han hanl1 hanl2* mutant embryos express SCR throughout the ground tissue, in a continuous

spherical domain (Long et al., 2006; chapter 2 of this thesis). Considering these observations, the genetic interaction of *han-30* with *tpl-1* may hold additional cues for better understanding the basis of the double root phenotype.

Perspective and open questions

Our finding that GATA29 can provide equivalent function to HAN in embryos, and GNC/GATA21 what appears to be partial function, opens the door to investigations into the roles of these factors in embryo development. The most obvious approach would be to construct fluorescent reporters and perform *in situ* hybridizations to determine whether these genes are transcribed in the embryo. The analysis of GATA29 mutants would be the first necessary step, which could be followed by a more comprehensive work-up of *han hanl1 hanl2 gata29* multiple mutants. A similar course of action could be taken for GNC; analyzing expression in the embryo followed by a comprehensive work-up of multiple mutants. One open question regarding triple mutants expressing GNC is whether the large number of embryos with variable intermediate phenotypes ever recover to form seedlings; this has been established for triple mutants expressing HAN-like genes and GATA29 but is an open question for triple mutants expressing GNC. Finding such plants would strongly confirm the rescue of triple mutants by GNC, and enhance our understanding of phylogenetic relationships between the HAN- and GNC-clade of B-class GATA factors.

Materials and Methods

Plant Strains and Culture

Plants were grown under constant illumination (~100 μmol/m^2/s) at 23°–25°C and 40%–50% relative humidity using commercial potting mix (RediEarth, Sun Gro Horticulture) containing systemic insecticide (Marathon 1% G, Olympic Horticultural Products) with slow-release fertilizer (19/12/6 Osmocote, Miracle-Gro). The *han-16/+;han1-56;hanl2-79* line was described previously (chapter two of this dissertation).

T-DNA construction

A *HAN* rescue construct that extends ~6.8 kbp upstream of the start codon with and NcoI restriction site and BglII/HindIII restriction sites introduced at the start and stop codons, respectively (ccATGg---TAAcgtt), was described previously (chapter 2 of this thesis). For ease of molecular cloning, the fragment was moved into a modified pCambia3300 T-DNA vector conferring resistance to glufosinate (Basta) and supporting higher plasmid yields from bacterial cultures. A HAN cDNA inserted between the NcoI and HindIII sites of this plasmid was sufficient to fully complement the phenotype of *han-16 hanl1-56 hanl2-79* embryos (see results). The coding sequences of the remaining 10 B-class GATA factors was then amplified from Arabidopsis genomic DNA with primers introducing restriction sites at the start and stop codon, inserted into a small cloning plasmid for Sanger sequencing, and then moved into the HAN T-DNA to generate chimeric genes (see Appendix B, Table S3.2 for details; for GATA17-like and GATA22/GNC, a BoXI site was introduced at the stop codon and blunt-end cloned into the HindIII site of the T-DNA, restoring the HindIII site).

Four ~300 bp synthetic NcoI-PflMI fragments purchased from Integrated DNA Technologies Inc. were used to further manipulate the HAN domain of HAN rescue constructs: codons for conserved residues were mutated, C^{69}>A (tgc>gct), L^{71}>R (ctc>aga), and L^{73}>L (ctt>gct), and a diagnostic BglI site added; the entire domain was deleted, V^{67}-L^{80}>EF (gaa ttc), introducing a diagnostic EcoRI site; the entire domain was substituted with the EAR motif of IAA17 (Tiwari et al., 2004), codons 3-29 of IAA17 replacing codons 2-85 of HAN, with a diagnostic SmaI site added in the process; as a control a mutant EAR motif was introduced as well, with the codons for leucine residues at position 14, 16 and 18 of the IAA17 sequence mutated to alanine codons and an NsiI site added for diagnostic purposes.

Four ~220 bp synthetic NcoI-MluI fragments were used to add both wild type and mutant forms of the HAN domain as well as wild the EAR motif of IAA17 to the N-terminus of GATA29. The codons 5-14 of GATA16 were replaced with the codons for residues 64-85 of HAN or a homologous fragment containing the mutations described above; diagnostic SalI and Spe I sites were introduced with the wild type and mutant form as well as an additional XbaI site located after the HAN motif. Similarly, codons 5-14 were exchanged with the same EAR motif-encoding sequences described above; diagnostic SmaI and AgeI sites were introduced with the wild type and mutant fragments, respectively. All segments derived from synthetic fragments were Sanger sequenced to ensure they did not harbor artifacts.

Three ~200 bp synthetic NcoI-BglII fragments were used to mutate the consensus EAR motif at the N-terminus of GATA29 as well as replace it with wild type and mutant forms of a HAN domain: the codons for leucine 7, 9 and 11 were mutated to alanine codons, introducing a SacI site in the process; in addition, codons 5-19 of GATA29 were replaced with codons for a wild type and mutant HAN domain as described above, adding diagnostic SalI and SpeI sites,

respectively. A complete listing of the HAN cDNA rescue construct (pCN-HAN) as well as all variants outlined here (segments between the NcoI and HindIII cloning sites only) can be found in Appendix B, Table S3.3.

Yeast two hybrid assays

Protein interactions were studied using pDEST32 BD and pDEST22 AD vectors from the ProQuest Two Hybrid System (Invitrogen Life Technologies). The N-terminal segment of HAN, encoding amino acids 1-148, were amplified from plasmids harboring a wild type copy of HAN and a variant with a mutant HAN domain by PCR; the primers (5'-GAGAGAGAGAAGTcgacAAGAGCC**ATG**-3' / 5'-GCACAGCGTCTtctaga**cta**GGAGTCAC-3') introduced a SalI site (underlined) upstream of the ATG start codon (bold) as well as a TAG stop codon after codon 148 followed by an XbaI site (underlined; bases listed in small letters are mismatched with the plasmid template). The fragments were inserted into a derivative of pENTR using the two cloning sites, and the resulting plasmids then recombined into the shuttle vector pDEST32, containing a GAL4 DNA-binding domain; in the two bait constructs, a spacer of 11 amino acids separates the attB recombination site from the start of the HAN segment. The remaining bait constructs were generated by conventional cloning of synthetic DNA fragments; first, the SalI-XbaI fragment of the HAN bait clone described above was replaced with a small linker containing an in-frame NcoI site (deleting the SalI site); NcoI and XbaI were then used to introduce the following segments: a ~80 bp fragment encoding MVDH followed by a wild type HAN domain (residues 64-84 of the HAN protein), introducing a diagnostic SalI site; a similar fragment with HAN domain as harboring the same three substitutions described above, introducing a diagnostic SpeI site; a ~ 90 bp fragment encoding the EAR motif of IAA17 (Tiwari

et al., 2004; residues 3-29 of IAA17), introducing a diagnostic SmaI site; a similar fragment with substitutions in the core leucine residues of the EAR motif as described above, introducing a diagnostic AgeI site.

Prey plasmids consisting of the entire TPL/TPR1/2/3/4 coding sequence in pDEST22 were provided as a gift from Dave Nelson (University of Georgia, Athens). For a strong positive control, we used pDEST32/Krev1 with pDEST22/RALGDS (Herrmann et al., 1996; Serebriiskii et al., 1999), and for negative controls we used pDEST32/HAN/IAA17 with pDEST22/RALGDS. The bait and prey constructs were transformed into the yeast strains MaV203 and selected on SD-Leu-Trp agar medium. Three Individual transformants were grown for one day in liquid culture, then dropped in several dilutions on SD-Leu-Trp-His supplemented with either 25, 50, or 100 mM 3-amino-1,2,4 triazole for selection and to test for autoactivation. Plates were grown at 30°C for one week, with growth checked at two day intervals.

Light Microscopy & Molecular Complementation

Embryos were examined by dissecting immature seeds from siliques, transferring them to a modified Hoyer's solution (70% chloral hydrate, 4% glycerol, 5% gum arabic) diluted 2:1 with 20% gum arabic, and cleared for 2 hours at room temperature (pre-globular stages) or overnight at 4^{o}C (globular stages and older). Imaging was performed on whole mount embryos with a DMR microscope (Leica) using differential interference contrast (DIC). Staging of mutant embryos was accomplished by comparing them to their wild-type siblings from the same silique.

The rescue activity of constructs was assayed in the progeny of primary transgenic plants (T1 generation) with the genotype *han-16/+ han1-56 hanl2-79* by counting the number of oblong-shaped arrest mutants. The *han-16* mutation is recessive, meaning that these plants

segregate approximately 25% mutant embryos; molecular complementation typically resulted in a reduction of triple mutant phenotypes to ~6%, a value indicating that once copy of the transgene is sufficient for providing HAN function. At least 10 independent lines per T-DNA construct and at least 100 embryos per line were examined. Rescue activity was confirmed in the T2 generation of select lines.

Phylogenetic analysis

The protein and DNA sequences of B-class GATA factors were retrieved from GenBank by BLAST searches. The following species and entries were included in the analysis:

Theobroma cacao (7 entries)

TCM_000376/XP_007046928, TCM_004753/XP_007051045, TCM_021303/XP_007035717, TCM_026732/XP_007031095, TCM_037285/XP_007012281, TCM_037666/XP_007012845, TCM_040387/XP_007014829;

Citrus sinensis (7 entries)

LOC102613347/XP_006472168, LOC102623069/XP_006480034, LOC102626300/XP_006475925, LOC102626792/XP_006475929, LOC102626980/XP_006475428, LOC102628775/XP_006466800, LOC102629993/XP_006488880;

Populus trichocarpa (10 entries)

POPTR_0002s20080g/XP_002302743, POPTR_0003s21770g/XP_002304024, POPTR_0005s02040g/XP_006382425, POPTR_0005s12440g/XP_006383192, POPTR_0006s24560g/XP_002308561, POPTR_0007s12930g/XP_002310237, POPTR_0008s22950g/XP_006380209, POPTR_0010s00360g/XP_006378028,

POPTR_0014s11940g/XP_002320320, POPTR_0018s05160g/XP_006371885;

Medicago truncatula (8 entries)

MTR_1g016170/XP_003588994, MTR_2g038720/XP_003595135,

MTR_5g020230/XP_003612004, MTR_7g112330/XP_003626174,

MTR_8g077510/XP_003629440, MTR_4g121110/KEH32333, MTR_4g126350/ KEH32529,

AFK42954;

Vitus vitis (8 entries)

LOC100247474/XM_002279247, LOC100247862/XM_010650837,

LOC100253375/XM_002275462, LOC100255968/XM_010662532,

LOC100256170/XM_002278333, LOC100257206/XM_010653922,

LOC100261004/XM_010657895, LOC104879590/XM_010653013;

Solanum lycopersicum (9 entries)

LOC101248643/XP_004233970, LOC101258225/XP_004233721,

LOC101261657/XP_004247383, LOC101262132/XP_004243958,

LOC101265163/XP_010318767, LOC101265416/XP_004252918,

LOC101266030/XP_010321383, LOC101266764/XP_004230282,

LOC101268354/XP_004251667.

Alignments were generated using the clustalo option provided by SeaView (Gouy et al., 2010) and phylogenetic trees with the BIO-NJ option using default parameters. Very similar topologies and confidence values were obtained for trees based on the full length amino acid sequence, the amino acid sequence of the zinc finger domain only, or the DNA sequence of the zinc finger domain only. Adding the B-class GATA factors of species outside of the eudicots (rice, date palm, Amborella, Selaginella, and Physcomitrella) to the analysis had little or no

effect on the GATA29 group and did not resolve its relationship to other GATA factors; the HAN clade remained strongly supported (all these species contain one or more HAN genes).

References

Behringer, C., Bastakis, E., Ranftl, Q.L., Mayer, K.F.X., and Schwechheimer, C. (2014). Functional diversification within the family of B-GATA transcription factors through the leucine-leucine-methionine domain. Plant Physiology 166, 293-305.

Behringer, C., and Schwechheimer, C. (2015). B-GATA transcription factors – insights into their structure, regulation, and role in plant development. Frontiers in Plant Science 6, 1-12.

Bi, Y., Zhang, Y., Signorelli, T., Zhao, R., Zhu, T., and Rothstein, S. (2005). Genetic analysis of *Arabidopsis* GATA transcription factor gene family reveals a nitrate-inducible member important for chlorophyll synthesis and glucose sensitivity. The Plant Journal 44, 680-692.

Bowen, A.J., Gonzalez, D., Mullins, J.G.L., Bhatt, A.M., Martinez, A., and Conlan, R.S. (2010). PAH-domain-specific interactions of the *Arabidopsis* transcription coregulatory SIN3-LIKE1 (SNL1) with telomere-binding protein 1 and ALWAYS EARLY2 Myb-DNA binding factors. Journal of Molecular Biology 395, 937-949.

Buzby, J.S., Yamada, T., and Tobin, E.M. (1990). A light-regulated DNA-binding activity interacts with a conserved region of a *Lemna gibba rbcS* promoter. The Plant Cell 2, 805-814.

Causier, B., Ashworth, M., Guo, W., and Davies, B. (2012). The TOPLESS interactome: A framework for gene repression in *Arabidopsis*. Plant Physiology 158, 423-438.

Chiang, Y., Zubo, Y.O., Tapken, W., Kim, H.J., Lavanway, A.M., Howard, L., Pilon, M., Kieber, J.J., and Schaller, G.E. (2012). Functional characterization of the GATA transcription factors GNC and CGA1 reveals their key role in chloroplast development, growth, and division in *Arabidopsis*. Plant Physiology 160, 332-348.

Chlon, T.M., and Crispino, J.D. (2012). Combinatorial regulation of tissue specification by GATA and FOG factors. Development 139, 3905-3916.

De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W., Audenaert, D., Van Campenhout, J., Overoorde, P., Jansen, L., Vanneste, St., Möller, B., Wilson, M., Holman, T., Van Isterdael, G., Brunoud, G., Vuylsteke, M., Vernoux, T., De Veylder, L., Inzé, D., Weijers, D., Bennett, M.J., and Beeckman, T. (2010). A novel Aux/IAA28 signaling cascade activates GATA23-dependent specification of lateral root founder cell identity. Current Biology 20, 1697-1706.

Donald, R.G.K., and Cashmore, A.R. (1990). Mutation of either G box or I box sequences profoundly affects expression from *Arabidopsis rbcS-1A* promoter. The EMBO Journal 9, 1717-1726.

Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F., an Bowman, J.L. (2003). Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. Current Biology 13, 1768-1774.

Eshed, Y., Baum, S.F., Perea, J.V., and Bowman, J.L. (2001). Establishment of polarity in lateral organs of plants. Current Biology 11, 1251-1260.

Eshed, Y., Izhaki, A., Baum, S.F., Floyd, S.F., Floyd, S.K., and Bowman, J.L. (2004). Asymmetric leaf development and blade expansion in *Arabidopsis* are mediated by KANADI and YABBY activities. Development 131, 2997-3006.

Fox, A.H., Liew, C., Holmes, M., Kowalski, K., Mackay, J., and Crossley, M. (1999). Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. The EMBO Journal 18, 2812-2822.

Gendron, J.M., Pruneda-Paz, J.L., Doherty, C.J., Gross, A.M., Knag, S.E., and Kay, S.A. (2012). *Arabidopsis* circadian clock protein, TOC1, is a DNA-binding transcription factor. Proceedings of the National Academy of Sciences 109, 3167-3172.

Giuliano, G., Pichersky, E., Malik, V.S., Timko, M.P., Scolnik, P.A., and Cahsmore, A.R. (1988). An evolutionary conserved protein binding sequence upstream of a plant light-regulated gene. Proceedings of the National Academy of Sciences 85, 7089-7093.

Gonzalez, D., Bowen, A.J., Carroll, T.S., and Conlan, R.S. (2007). The transcription compressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription. Molecular and Cellular Biology 27, 5306-5315.

Griffiths, S., Dunford, R.P., Coupland, G., and Laurie, D.A. (2003). The evolution of CONSTANS-like gene families in barley, rice, and *Arabidopsis*. Plant Physiology 131, 1855-1867.

Hill, K., Wang, H., and Perry, S.E. (2008). A transcriptional repression motif in the MADS factor AGL15 is involved in recruitment of histone deacetylase complex components. The Plant Journal 53, 172-185.

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

Hiratsu, K., Mitsuda, N., Matsui, K., and Ohme-Takagi, M. (2004). Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in *Arabidopsis*. Biochemical and Biophysical Research Communications 321, 172-178.

Horiguchi, G., Mollá-Morales, A., Pérez-Pérez, J.M., Kojima, K., Robles, P., Ponce, M.R., Micol, J.L., and Tsukaya, H. (2011). Differential contributions of ribosomal genes to *Arabidopsis thaliana* leaf development. The Plant Journal 65, 724-736.

Kagale, S., Links, M.G., and Rozwadowski, K. (2010). Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in *Arabidopsis*. Plant Physiology 152, 1109-1134.

Kanei, M., Horiguchi, G., and Tsukaya, H. (2012). Stable establishment of cotyledon identity during embryogenesis in *Arabidopsis* by *ANGUSTIFOLIA3* and *HANABA TARANU*. Development 139, 2436-2446.

Kehoe, D.M., Degenhardt, J., Winicov, I., and Tobin, E.M. (1994). Two 10-bp regions are critical for phytochrome regulation of a *Lemna gibba Lhcb* gene promoter. The Plant Cell 6, 1123-1134.

Kirik, V., Schnittger, A., Radchuk, V., Adler, K., Hülskamp, M., and Bäumlein, H. (2001). Ectopic expression of the *Arabidopsis AtMYB23* gene induces differentiation of trichome cells. Developmental Biology 235, 366-377.

Ko, L.J., and Engel, J.D. (1993). DNA-binding specificities of the GATA transcription factor family. Molecular and Cellular Biology 13, 4011-4022.

Krogan, N.T., Hogan, K., and Long, J.A. (2012). APETALA2 negatively regulates multiple floral organ identity genes in *Arabidopsis* by recruiting the co-repressor TOPLESS and the histone deacetylase HDA19. Development 139, 4180-4190.

Lam, E., and Chua, N. (1989). ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in *Cab* promoters. The Plant Cell 1, 1147-1156.

Lee, J.E., and Golz, J.F. (2012). Diverse roles of Groucho/Tup1-like co-repressors in plant growth and development. Plant Signaling & Behavior 7, 1-7.

Liu, P., Koizuka, N., Martin, R.C., and Nonogaki, H. (2005). The *BME3 (Blue Micropylar End 3)* GATA zinc finger transcription factor is a positive regulator of *Arabidopsis* seed germination. The Plant Journal 44, 960-971.

Liu, Z., and Karmarkar, V. (2008). Groucho/Tup1 family co-repressors in plant development. Trends in Plant Science 13, 137-144.

Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M. (2006). TOPLESS regulates apical embryonic fate in *Arabidopsis*. Science 312, 1520-1523.

Long, J.A., Woody, S., Poethig, S., Meyerowitz, E.M., and Barton, M.K. (2002). Transformation of shoots into roots in *Arabidopsis* embryos mutant at the TOPLESS locus. Development 129, 2297-2306.

Lowry, J.A., and Atchley, W.R. (2000). Molecular evolution of the GATA family of transcription factors: Conservation within the DNA-binding domain. Journal of Molecular Evolution 50, 103-115.

Luo, X., Lin, W. Zhu, S., Zhu, J., Sun, Y., Fan, X., Chang, M., Hao, Y., Oh, E., Tian, M., Liu, L., Zhang, M., Xie, Q., Chong, K., and Wang, Z. (2010). Integration of light- and brassinosteroid-signaling pathways by a GATA transcription factor in *Arabidopsis*. Developmental Cell 19, 872-883.

Malavé, T.M., and Dent, S.Y.R. (2006). Transcriptional repression by Tup1-Ssn6. Biochemistry and Cell Biology 84, 437-443.

Manfield, I. W., Devlin, P.F., Jen, C., Westhead, D.R., and Gilmartin, P.M. (2007). Conservation, convergence, and divergence of light-responsive, circadian-regulated, and tissue-specific expression patterns of the *Arabidopsis* GATA gene family. Plant Physiology 143, 941-958.

Mara, C.D., and Irish, V.F. (2008). Two GATA transcription factors are downstream effectors of floral homeotic gene action in *Arabidopsis*. Plant Physiology 147, 707-718.

Naito, T., Kiba, T., Koizumi, N., Yamashino, T., and Mizuno, T. (2007). Characterization of a unique GATA family gene that responds to both light and cytokinin in *Arabidopsis thaliana*. Bioscience, Biotechnology, and Biochemistry 71, 1557-1560.

Nawy, T., Bayer, M., Mravec, J., Friml, J., Bimbaum, K.D., and Lukowitz, W. (2010). The GATA factor HANABA TARANU is required to position the proembryo boundary in the early *Arabidopsis* embryo. Developmental Cell 19, 103-113.

Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H., and Ohme-Takagi, M. (2001). Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. The Plant Cell 13, 1959-1968.

Omichinski, J.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S.J., and Gronenborn, A.M. (1993). NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. Science 261, 438-446.

Oppenheimer, D.G., Herman, P.L., Sivakumaran, S., Esch, J., and Marks, M.D. (1991). A *myb* gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. Cell 67, 483-493.

Patient, R.K., and McGhee, J.D. (2002). The GATA family (vertebrates and invertebrates). Current Opinion in Genetics & Development 12, 416-422.

Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N., and Clark, S.E. (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. The Plant Cell 17, 61-76.

Reyes, J.C., Muro-Pastor, M.I., and Florencio, F.J. (2004). The GATA family of transcription factors in *Arabidopsis* and rice. Plant Physiology 134, 1718-1732.

Richter, R., Behringer, C., Müller, I.K., and Schwechheimer, C. (2010). The GATA-type transcription factors GNC and GNL/CGA1 repress gibberellin signaling downstream from DELLA proteins and PHYTOCHROME-INTERACTING FACTORS. Genes & Development 24, 2093-2104.

Richter, R., Behringer, C., Zourelidou, M., and Schwechheimer, C. (2013). Convergence of auxin and gibberellin signaling on the regulation of the GATA transcription factors *GNC* and *GNL* in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences 110, 13192-13197.

Robson, F., Costa, M.M.R., Hepworth, S.R., Vizir, I., Piñeiro, M., Reeves, P.H., Putterill, J., and Coupland, G. (2001). Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. The Plant Journal 28, 619-631.

Scazzocchio, C. (2000). The fungal GATA factors. Current Opinion in Microbiology 3, 126-131.

Shoichet, S.A., Malik, T.H., Rothman, J.H., and Shivdasani, R.A. (2000). Action of the *Caenorhabditis elegans* GATA factor END-1 in Xenopus suggests that similar mechanisms initiate endoderm development in ecdysozoa and vertebrates. Proceedings of the National Academy of Sciences 97, 406-4081.

Smith, Z.R., and Long, J.A. (2010). Control of *Arabidopsis* apical-basal embryo polarity by antagonistic transcription factors. Nature 464, 423-427.

Song, C., and Galbraith, D.W. (2006). AtSAP18, an orthologue of human SAP18 is involved in the regulation of salt stress and mediates transcriptional repression in Arabidopsis. Plant Molecular Biology 60, 241-257.

Starich, M.R., Wilkström, M., Arst Jr, H.N., Clore, G.M., and Gronenborn, A.M. (1998). The solution structure of a fungal AREA protein-DNA complex: An alternative binding mode for the basic carboxyl tail of GATA factors. Journal of Molecular Biology 277, 605-620.

Szemenyei, H., Hannon, M., and Long, J.A. (2008). TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. Science 319, 1384-1386.

Tiwari, S.B., Hagen, G., and Guilfoyle, T.J. (2004). Aux/IAA proteins contain a potent transcriptional repression domain. The Plant Cell 16, 533-543.

Tiwari, S.B., Wang, X., Hagen, G., and Guilfoyle, T.J. (2001). AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. The Plant Cell 13, 2809-2822.

Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M., and Orkin, S.H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell 90, 109-119.

Turki-Judeh, W., and Courey, A.J. (2012). Groucho: A corepressor with instructive roles in development. Current Topics in Developmental Biology 98, 65-96.

Wang, Z., Cao, H., Sun, Y., Li, X., Chen, F., Carles, A., Li, Y., Ding, M., Zhang, C., Deng, X., Soppe, W.J.J., and Liu, Y. (2013). Arabidopsis paired amphipathic helix proteins SNL1 and SNL2 redundantly regulate primary seed dormancy via abscisic acid-ethylene antagonism mediated by histone deacetylation. The Plant Cell 25, 149-166.

Whipple, C.J., Hall, D.H., DeBlasio, S., Taguchi-Shiobara, F., Schmidt, R.J., and Jackson, D.P. (2010). A conserved mechanism of bract suppression in the grass family. The Plant Cell 22, 565-578.

Zhang, X., Zhou, Y., Ding, L., Wu, Z., Liu, R., and Meyerowitz, E.M. (2013). Transcription repressor HANABA TARANU controls flower development by integrating the actions of multiple hormones, floral organ specification genes, and GATA3 family genes in *Arabidopsis*. The Plant Cell 25, 83-101.

Zhao, Y., Medrano, L., Ohashi, K., Fletcher, J.C., Yu, H., Sakai, H., and Meyerowitz, E.M. (2004). HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in *Arabidopsis*. The Plant Cell 16, 2586-2600.

Zhu, Z., Xu, F., Zhang, Y., Cheng, Y.T., Wiermer, M., Li, X., and Zhang, Y. (2010). *Arabidopsis* resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. Proceedings of the National Academy of Sciences 107, 13960-13965.

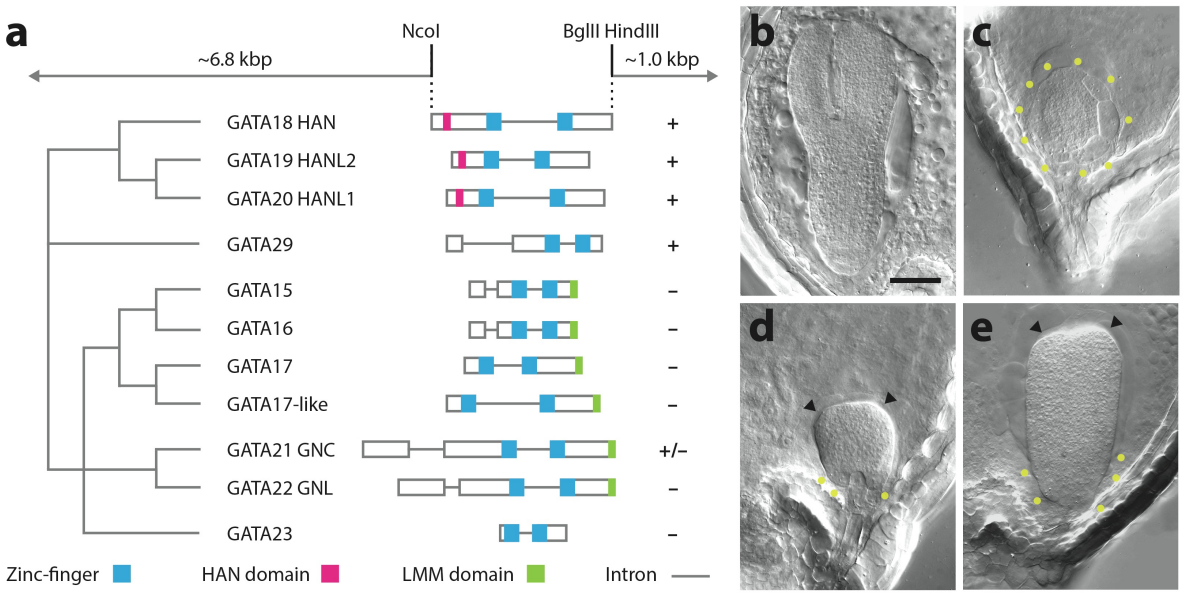

Figure 3.1. Molecular complementation of *han hanl1 hanl2* mutants with B-class GATA factors. (a) Overview of proposed phylogenetic relationships within the Arabidopsis B-class GATA factors (left; after Reyes et al., 2004; Behringer et al., 2014; most branches are not well supported), their gene structure including conserved domains (middle) and their ability to rescue triple mutant embryos (right, indicated as "+" or "-"); all genes were expressed under the control of HAN regulatory sequences (outlined on top). (b–e) Phenotypic classes used for scoring molecular complementation; (b) normal embryo at the torpedo stage, (c) *hanl hanl1 hanl2* embryo with a strong phenotype, and (d, e) representative examples of intermediate phenotypes. Enlarged surface cells marked with dots and rudimentary cotyledon primordia, the defining hallmarks of intermediate phenotypes, with arrowheads; scale bar, 50 µm.

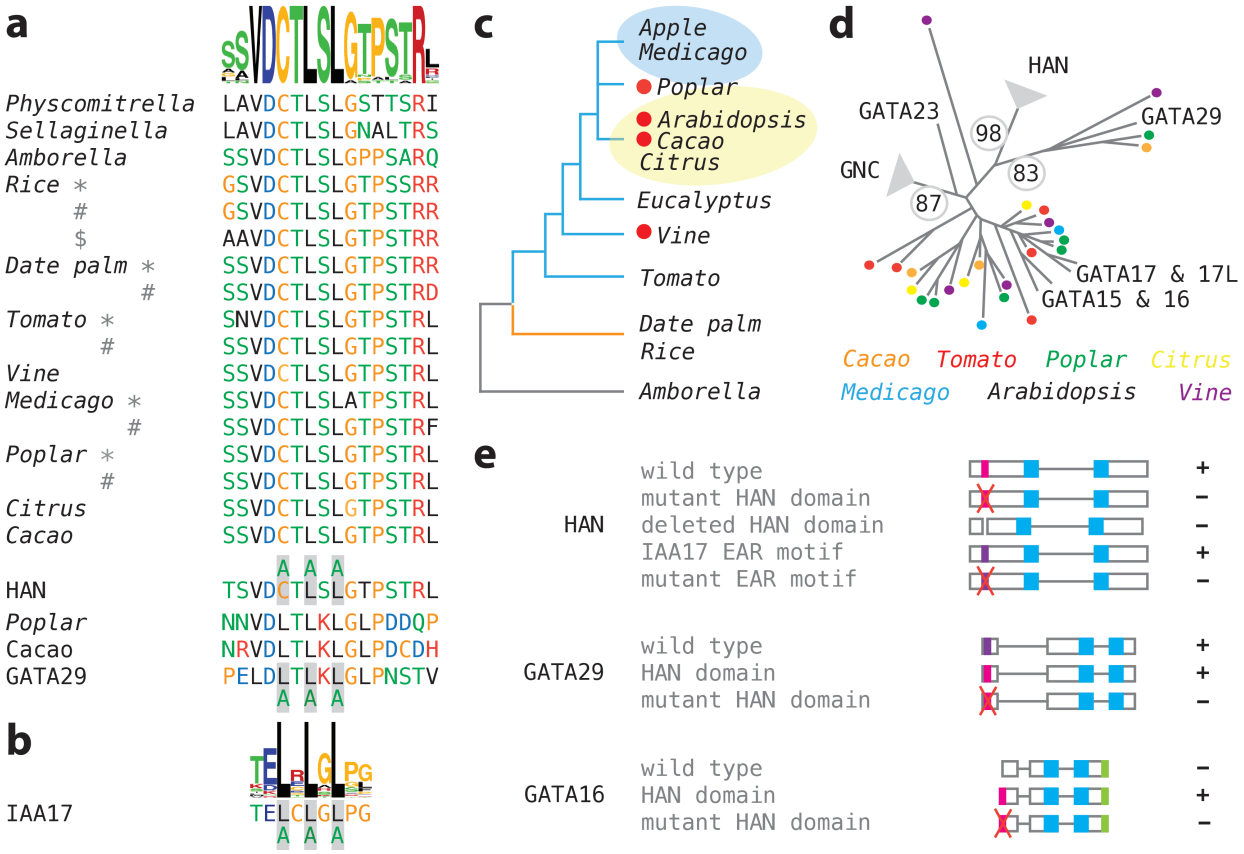

Figure 3.2. Functional characterization of the HAN domain and the GATA29 EAR motif.

(a) HAN domain of representative land plants with a consensus sequence shown on top; accession numbers: Physcomitrella XP_001752308, Sellaginella XP_002963788, Amborella XP_006878595, rice * NP_001056415, # NP_001172493, $ EEE64727, date plam * XP_008800719, # XP_008805393, tomato * XP_004233721, # XP_004233970, vine XP_010649139, Medicago * XP_003629440, # XP_003612004, poplar * XP_002310237, # XP_006383192, Citrus XP_006466800, cacao XP_007046928; three substitutions introduced into Arabidopsis HAN (C^{69}>A, L^{71}>R, L^{73}>A) highlighted by grey boxes; the N-terminus of GATA29 genes from poplar, cacao, and Arabidopsis is aligned below to show the similarity between their consensus EAR motif and the HAN domain; three substitutions introduced into Arabidopsis GATA29 (L^{7}>A, L^{9}>A, L^{11}>A) highlighted by grey boxes.

(b) EAR motif consensus sequence of the Arabidopsis IAA gene family (after Tiwari et al., 2004); the IAA17 motif (residues 12-20) used in chimeric proteins shown below, with three substitutions (L^{14}>A, L^{16}>A, L^{18}>A) shown to abolish function highlighted by grey boxes.

(c) Distribution of recognizable GATA29 homologs (marked by red dots) within the flowering plants; Amborella is at the base of the angiosperms; the monocot and eudicot clades are shown in orange and blue, respectively (no basal dicots are included); tomato represents the asterids, and vine is considered sister to the eurosids; the eurosids I are highlighted in blue and the eurosids II in yellow – poplar has been placed in either group and its position is still debated.
(d) Un-rooted, simplified tree of B-class GATA factors from representative eudicots, generated from an alignment of the full-length protein sequences; Arabidopsis GATA factors are listed by name, for the other proteins only the species of origin is indicated (colored dots); only three clades have robust bootstrap support (numbers in circles): HAN, GNC (represented as triangles, individual genes not shown), and GATA29; support for all other clades is poor (bootstrap values of 50 and lower), with exception of some branches at the tips resulting from recent duplication events; very similar topologies and support were obtained with the amino acid or the DNA sequence of the zinc finger domain only; see methods for accession numbers and Appendix B, Fig. S3.1, S3.2, S3.3 for detailed representations of this as well as alternative trees.
(e) Functional analysis of the HAN domain and the EAR motif of GATA29 by molecular complementation; schematic representation of mutant and chimeric HAN, GATA29, and GATA16 variants and their ability to complement *han hanl1 hanl2* mutants (noted as "+" or "-"); the zinc finger and the LMM domain highlighted in blue and yellow, respectively; the HAN domain is shown in pink, and the EAR motif of GATA29 or IAA17 in purple.

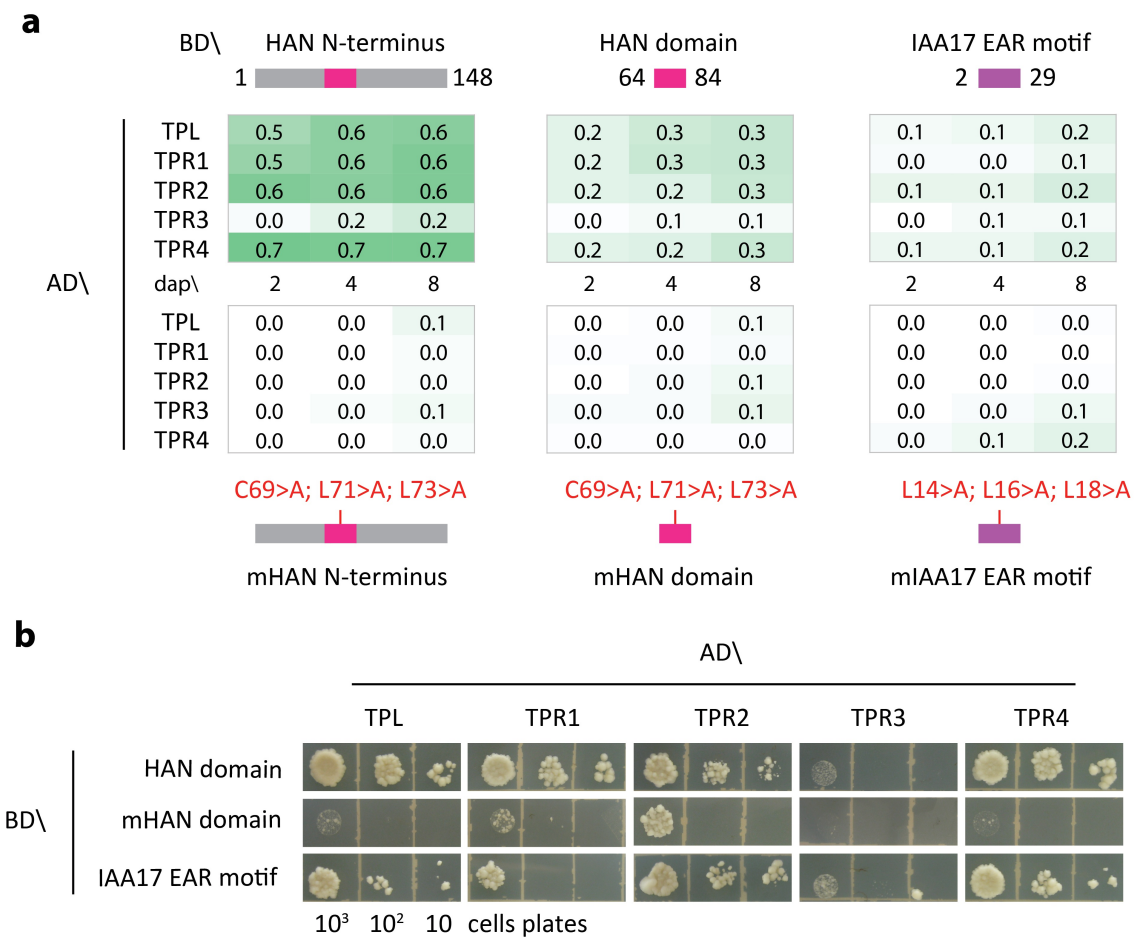

Figure 3.3. Protein-protein interactions between HAN and TPL corepressors.

(a) Interactions of TPL proteins fused to a VP16 transcriptional activation domain with variants of HAN fused to a Gal4 DNA binding domain in yeast. Numbers represent relative confluence of yeast colonies after 2, 4, and 8 days (see methods for details). Wild type segments that were tested for interaction with TPL are outlined above the table (HAN N-terminus, HAN domain, EAR motif of IAA17; numbers refer to the amino acid sequence) and mutant segments tested as control below the table (substituted positions indicated in red). (b) Images of representative interaction tests on 25mM –HIS 3-AT showing wild type and mutant forms of the HAN domain and the IAA17 domain, respectively, for comparison.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

Plants display a diversity of forms and structures, all of which are derived from the activity of meristems set at the tips of branches. These meristems are established during embryogenesis, a process beginning with the single-celled zygote and through a series of formative cell divisions, giving rise to a simplified version of the plant body: two meristems set at either pole interconnected by a central vascular cylinder, and a radial arrangement of primary tissue. How is this spatial arrangement decided upon, and what does the molecular network responsible for regulating this process look like? In this dissertation I have outlined my efforts towards elucidating the mechanisms establishing embryonic polarity in *Arabidopsis thaliana*. Previous work in my lab identified a GATA transcription factor with an intriguing effect on embryogenesis; the loss of HANABA TARANU (HAN) results in embryos that express genes normally confined to the hypophysis and suspensor in cells of the lower tier, while genes expressed in the lower tier become limited to the upper tier. This effect was interpreted as an apical shift in the proembryo boundary, across which the root meristem normally originates from; consistent with this view, *han* mutants accumulate enlarged suspensor-like cells in their lower tier.

In chapter 2, I aimed to identify why *han* mutants often recover from their defects to form complete seedlings. HAN is closely related to two other B-class GATA factors, HAN-LIKE1 (HANL1) and HANL2, and all three proteins share a conserved sequence N-terminal to the zinc

finger, called the HAN domain. My hypothesis was that the recovery of *han* mutants might be attributable to these factors. To determine the functional equivalency of these genes, I made rescue constructs in which the coding sequence of HANL1/2 was placed under a HAN regulatory sequence, and expressed each of them in *han* mutants. I found that HANL1 and HANL2 were biochemically equivalent to HAN in embryos. I next constructed fluorescent reporter genes using sequence upstream of HANL1/2 driving a triple Venus tagged to a nuclear localization signal, and performed *in situ* hybridizations using probes for non-conserved regions of these proteins, to determine their expression. I found that both genes were below detection prior to the globular stage, yet afterwards, their expression closely mimics HAN transcription; all three genes are broadly expressed in globular embryos, with strongest expression in central cells directly adjacent to the hypophysis, and thereafter in the provasculature, and shoot and root meristem.

To determine the function of HAN-like genes, mutant alleles were generated by the *Arabidopsis* TILLING facility and analyzed by DIC microscopy of immature seed. I found that the loss of *hanl1* and *hanl2* separately, or in combination, has only a minimal effect on early embryogenesis, as the embryos produce marginally fewer suspensor cells; and combining *han* with *hanl1* or *hanl2* results in embryos that look very similar to *han* mutants, though they too have fewer suspensor cells. Combining mutations in all three genes however, produces embryos with striking, novel defects. Triple mutants form oblong blimp-shaped embryos that display no anatomical signatures of organ primordia, a shoot or root apical meristem, or vascular cylinder, and they subsequently arrest. The failure of these embryos to recover indicates that the HAN-like genes are functionally redundant with HAN and necessary for root formation and subsequent development in *han* mutants.

Due to the absence of visible hallmarks of development, I resorted to characterizing these embryos using a panel of fluorescent reporter genes for key patterning processes: the early establishment of polarity, formation of the meristems, initiation of lateral organs, and vascular tissue differentiation. I found that triple mutants were fairly normal in regards to the establishment of embryonic polarity. Like *han*, these mutants express ML1, WUS, and PLT – regulators of epidermis formation, shoot meristem maintenance, and root QC specification– in their respective domains; indicating that triple mutants retain molecular markers of polarity. At the same time, markers for other basic elements of the body plan never become expressed in triple mutant embryos. These included; CLV3, TMO, and FIL, which function to maintain the shoot apical meristem, control cell divisions, and specify leaf polarity; indicating that these structures are incompletely specified. Strikingly, markers normally expressed below the proembryo boundary – the hypophysis and its daughters, the lens-shaped precursor cell of the QC and the precursor of the root columnella – displayed a much greater expansion than was observed in *han* single mutants. This included WOX8 and SUC3, normally expressed in the basal daughter of the hypophysis, and WOX5, normally expressed in the apical daughter, which become expressed in all surface or sub-surface cells, respectively, to surround triple mutants. I interpret this effect as a partial radialization of the apical-basal pattern.

As no other mutants affect patterns of gene expression in a similar way to *han* in the embryo, I looked to potentially analogous changes in mutants affecting leaf polarity. In these cases, the loss of adaxial or abaxial determinants expressed at one leaf surface, results in the expansion of the determinant normally expressed at the other. As a consequence, mutant leaves develop as rod-like organs lacking a lamina and taking on the appearance of only one or the other surface; they lose the boundary which normally separates the two polar domains and

become radialized (Kerstetter et al. 2001; Emery et al. 2003; Eshed et al., 2004). Viewing the triple mutant phenotype in this light led me to elaborate our model of HAN function in embryo patterning. While the boundary between proembryo and suspensor is apically shifted in *han* single mutants, the additional loss of *hanl1* and *hanl2* apparently prevents embryos from maintaining it, and it becomes lost completely. Normally, the boundary would separate vascular precursors from precursors of the root QC and root columnella along the apical basal axis of the root; but as these basal gene expression domains expand into all sub-surface or surface cells, it seems that they are no longer delimited in the apical-basal dimension. Intriguingly, these gene expression domains still recognize a radial pattern; QC and distal root markers expand just around the periphery but not into the central, vascular domain of embryos.

In chapter 3, I describe efforts towards clarifying potential functional overlap between HAN family genes and other B-class GATA factors, as well as determining the structural basis of HAN activity. To assess functional equivalency with HAN I used the same constructs used for assessing equivalency of HANL1/2 with HAN; coding sequences for each of the B-class GATA factors, as well as chimeric and mutant variants, were placed under the HAN regulatory sequence, and transformed into triple mutant plants. Rescue was determined as a reduction in the number of triple mutant phenotypes produced, and I found that only GATA29 and GNC displayed activity in triple mutants. GATA29 displayed equivalent function, while GNC expression resulted in embryos with variable intermediate phenotypes somewhat reminiscent of *han* single mutants; suggesting that it is only partially able to substitute for HAN in embryos.

Constructs in which the HAN domain is mutated or deleted are inactive, indicating that it is required for the normal functioning of HAN. Interestingly, the HAN domain displays a somewhat similar arrangement of hydrophobic amino acids as found in an LxLxL EAR motif,

although it does not conform to the EAR consensus. In contrast, a canonical EAR motif can be found at the N-terminus of GATA29. EAR motif sequences are recognized by members of the TOPLESS family of transcriptional co-repressors (Szemenyei et al., 2008), and an interaction between HAN and TPL would indicate that the HAN domain acts as a repressor domain; HAN was previously identified as one of many transcription factors binding TPL in a large-scale interaction study (Causier et al., 2012). To confirm this finding, I assayed the N-terminal half of HAN containing the HAN domain; the HAN domain alone; and mutant variants of each domain fused to a GAL4 DNA binding domain for interaction with TPL family members fused to a transcriptional activation domain in yeast. We found that HAN interacts with nearly all TPL family members and that the HAN domain is necessary and sufficient for mediating this interaction. These results indicate that the HAN domain indeed functions similarly to an EAR motif. GATA29 is not closely related to HAN family members; thus, the equivalent biochemical activity of GATA29 and HAN in our molecular complementation assay may be due to convergent function. In summary, I found that there are two requirements for HAN function; a B-class zinc finger (a number of mutant alleles are substitutions of conserved residues in this domain) and a repressor domain, suggesting that HAN functions as a transcriptional repressor.

Looking forward, I am interested in the possible significance of other domains found in B-class GATA proteins, in particular the type of zinc finger. The plant GATA genes have been subdivided based on the structure of their zinc fingers, yet it is unknown how these differences affect DNA binding properties. To answer this question, I would replace the B-class zinc finger of a HAN rescue construct with an A-, C-, or D-class zinc finger and assess complementation in *han* mutants (in the same manner previously described). Our expectation is that these constructs will not be able to rescue *han* mutants given that sequence differences between zinc finger

classes are concentrated in the α-helix region, where the zinc finger binds the major groove of DNA, and are believed to influence DNA binding specificities (Reyes et al., 2004). Another short, conserved domain of all B-class GATA factors with the exception of HAN family genes is the hydrophobic LLM domain, found at the C-terminus; the domain is less conserved in GATA23 (Behringer et al., 2014). Mutating the LLM domain of GNC or GNL in overexpression constructs suggests that it is important for function (Behringer et al., 2014). I would assess the contribution of this domain to the function of B-class GATA factors by fusing it to the HAN rescue construct and performing complementation assays. This will determine whether the absence of an LLM domain is necessary for HAN function in embryos, and alternatively, whether the domain can confer any particular biochemical activity to HAN.

It still remains to be determined whether the HAN domain functions as a repressor domain in plants. I would address this by fusing the HAN domain to an activating transcription factor. MYB23 is positive regulator of trichome development (Kirik et al., 2001; Oppenheimer et al., 1991), and when it is fused to a strong, synthetic EAR motif, termed SRDX, fewer or no trichomes are formed (Hiratsu et al., 2003). In a similar fashion, I would overexpress MYB23 fused to the HAN domain in plants and assess trichome density. By comparing chimeric MYB genes with a HAN domain, SRDX domain, or the EAR motif of IAA17, we may be able to quantify the activity of the HAN domain. We could complement these studies by assessing protein-protein interactions between HAN and TPL family members, using the transient expression of chimeric proteins in leaves and co-precipitation experiments. Additionally, combining *han* alleles with dominant-negative *tpl-1* would allow us to assess genetic interactions.

Perhaps the most exciting way forward however, would be to identify factors that regulate HAN or the targets of HAN. The best context in which to approach this might be the shoot or root apical meristem. HAN family genes are expressed above the initials of the root meristem and within the organ boundaries of the shoot meristem. HAN function is needed for normal inflorescence and flower meristem development, as the loss of *han* results in small flattened shoot meristems, fewer floral organs, and fused stamens and carpels (Zhao et al., 2004). It was also found that inducibly overexpressing HAN causes root swelling, perturbs the arrangement of cells in files and diminishes the root cap; eventually causing arrest (Zhao et al., 2004). These results imply that HAN functions in the apical meristems; tissues from these more readily available sources could be used for protein binding experiments and transcriptional profiling. Additionally, the HAN family could be targeted using inducible RNAi (Wielopolska et al., 2005), which would provide a way to overcome embryonic lethality, and uncover the presumptive function of HAN and HAN-like genes in seedlings or adult plants.

References

Behringer, C., Bastakis, E., Ranftl, Q.L., Mayer, K.F.X., and Schwechheimer, C. (2014). Functional diversification within the family of B-GATA transcription factors through the leucine-leucine-methionine domain. Plant Physiology 166, 293-305.

Causier, B., Ashworth, M., Guo, W., and Davies, B. (2012). The TOPLESS interactome: A framework for gene repression in *Arabidopsis*. Plant Physiology 158, 423-438.

Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F., an Bowman, J.L. (2003). Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. Current Biology 13, 1768-1774.

Eshed, Y., Izhaki, A., Baum, S.F., Floyd, S.F., Floyd, S.K., and Bowman, J.L. (2004). Asymmetric leaf development and blade expansion in *Arabidopsis* are mediated by KANADI and YABBY activities. Development 131, 2997-3006.

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

Kerstetter, R.A., Bollman, K., Taylor, R.A., Bomblies, K., and Peothig, R.S. (2001). KANADI regulates organ polarity in *Arabidopsis*. Nature 411, 706-709.

Kirik, V., Schnittger, A., Radchuk, V., Adler, K., Hülskamp, M., and Bäumlein, H. (2001). Ectopic expression of the *Arabidopsis AtMYB23* gene induces differentiation of trichome cells. Developmental Biology 235, 366-377.

Oppenheimer, D.G., Herman, P.L., Sivakumaran, S., Esch, J., and Marks, M.D. (1991). A *myb* gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. Cell 67, 483-493.

Reyes, J.C., Muro-Pastor, M.I., and Florencio, F.J. (2004). The GATA family of transcription factors in *Arabidopsis* and rice. Plant Physiology 134, 1718-1732.

Szemenyei, H., Hannon, M., and Long, J.A. (2008). TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. Science 319, 1384-1386.

Wielopolska, A., Townley, H., Moore, I., Waterhouse, P., and Helliwell, C. (2005). A high-throughput inducible RNAi vector for plants. Plant Biotechnology Journal 3, 583-590.

Zhao, Y., Medrano, L., Ohashi, K., Fletcher, J.C., Yu, H., Sakai, H., and Meyerowitz, E.M. (2004). HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in *Arabidopsis*. The Plant Cell 16, 2586-2600.

APPENDIX A

CHAPTER 2 SUPPLEMENTARY

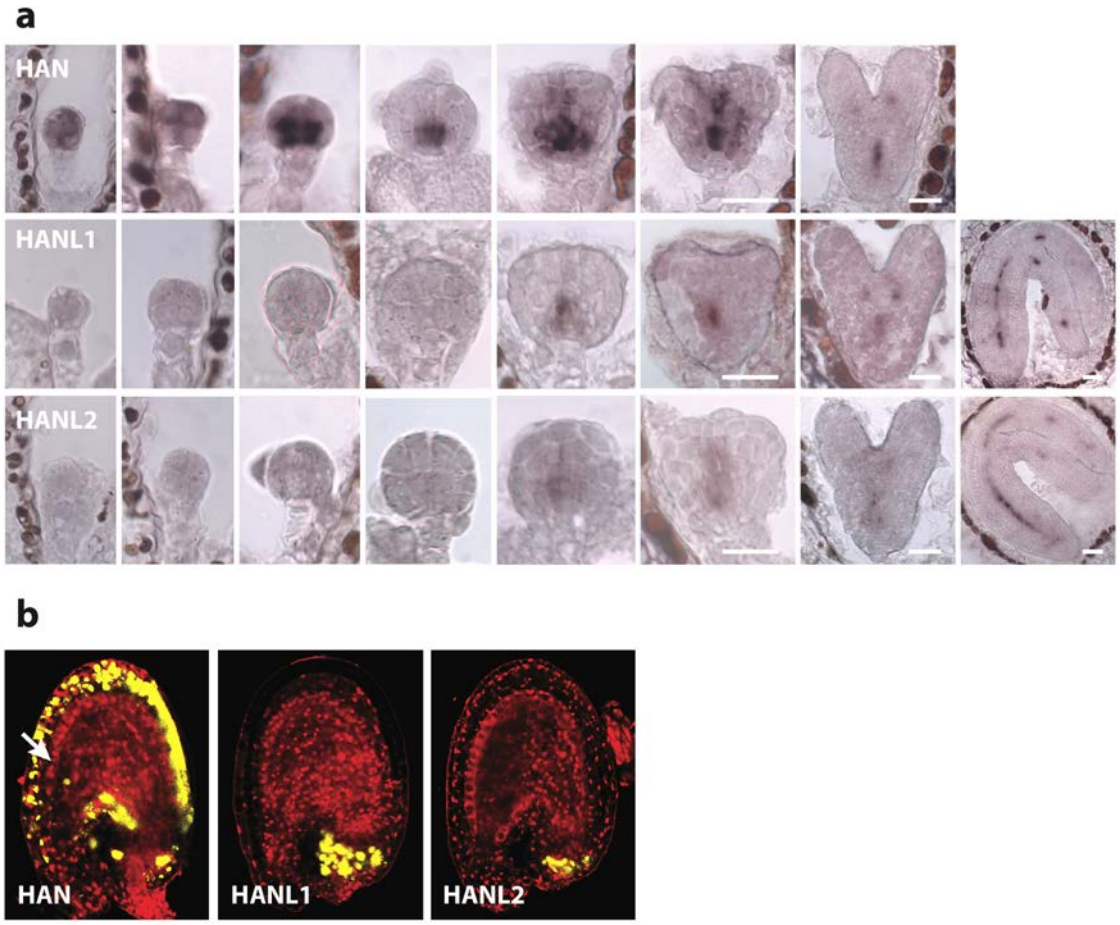

Figure. S2.1. Expression of HAN family genes in embryos. HAN transcripts are detected throughout the embryo proper at the 2-,4-, and 16-cell stage (a, upper row), before being refined to central cells adjacent to the hypophysis, then to the provasculature. The expression of HANL1 (middle row) and HANL2 (bottom row) however are not detectable prior to the globular stage; when they mimic HAN expression. HANL1/2 appeared to be weaker than HAN in *in situ* hybridizations. (b) The expression of HAN is evident at the 1-cell stage, viewed through the seed coat (only HAN expression is visible; see arrowhead). Scale bar is 20 μm

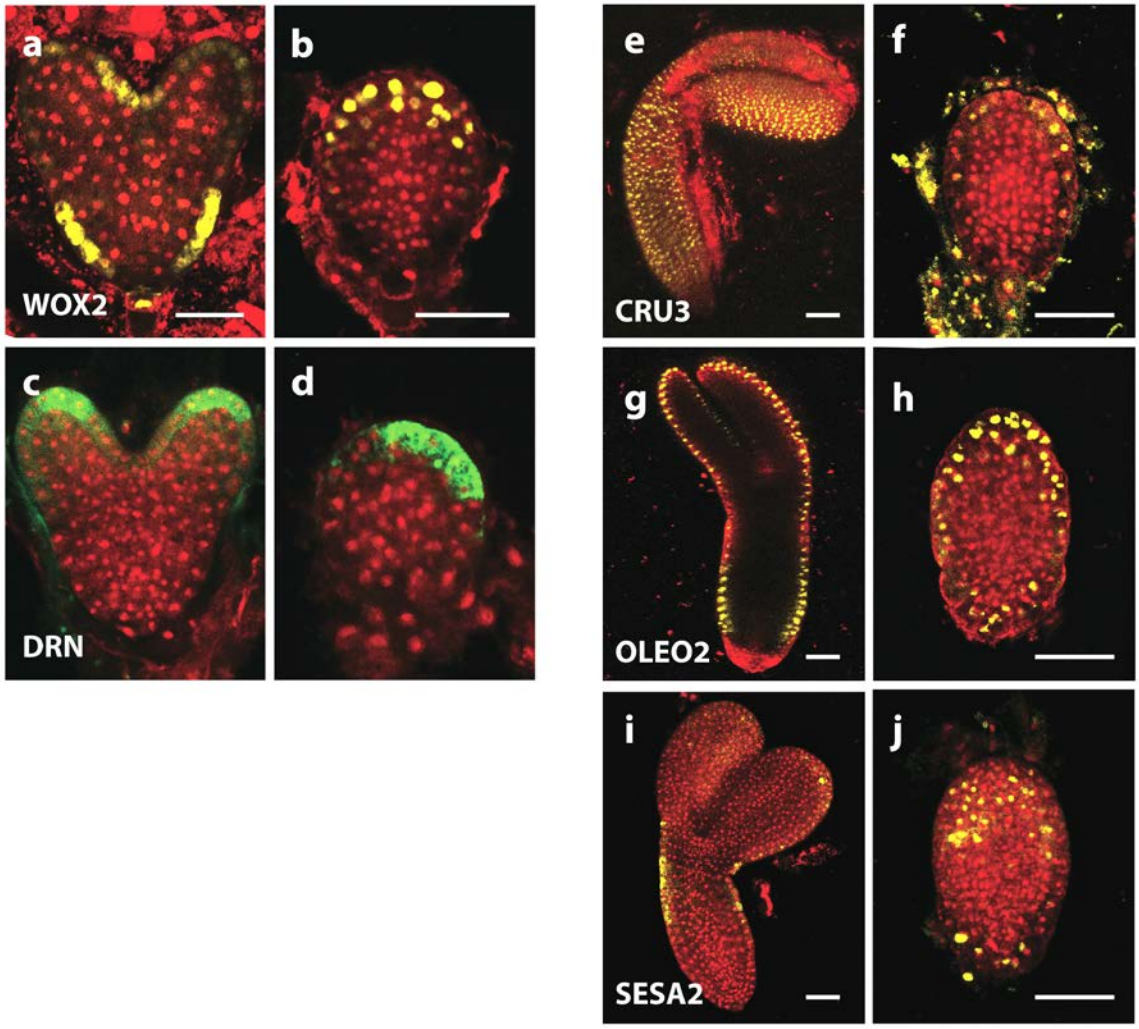

Figure. S2.2. Expression of genes regulating early polarity and maturation. Regulators of shoot development, *pWOX2::nYFP* (Breuninger et al., 2008), and lateral organ initiation, *pDRN::erGFP* (Chandler et al., 2011), are expressed in similar overall patterns in wild type (a,c) and triple mutant (b,d) torpedo stage embryos. Additionally, markers for seed storage proteins and lipid accumulation, part of the maturation process, CRU3::Venus-NLS, OLEO2::Venus-NLS, and SESA2::Venus-NLS, and active in triple mutant embryos (f,h,j), as they are in wild type (e,g,i) at the bent cotyledon stage.

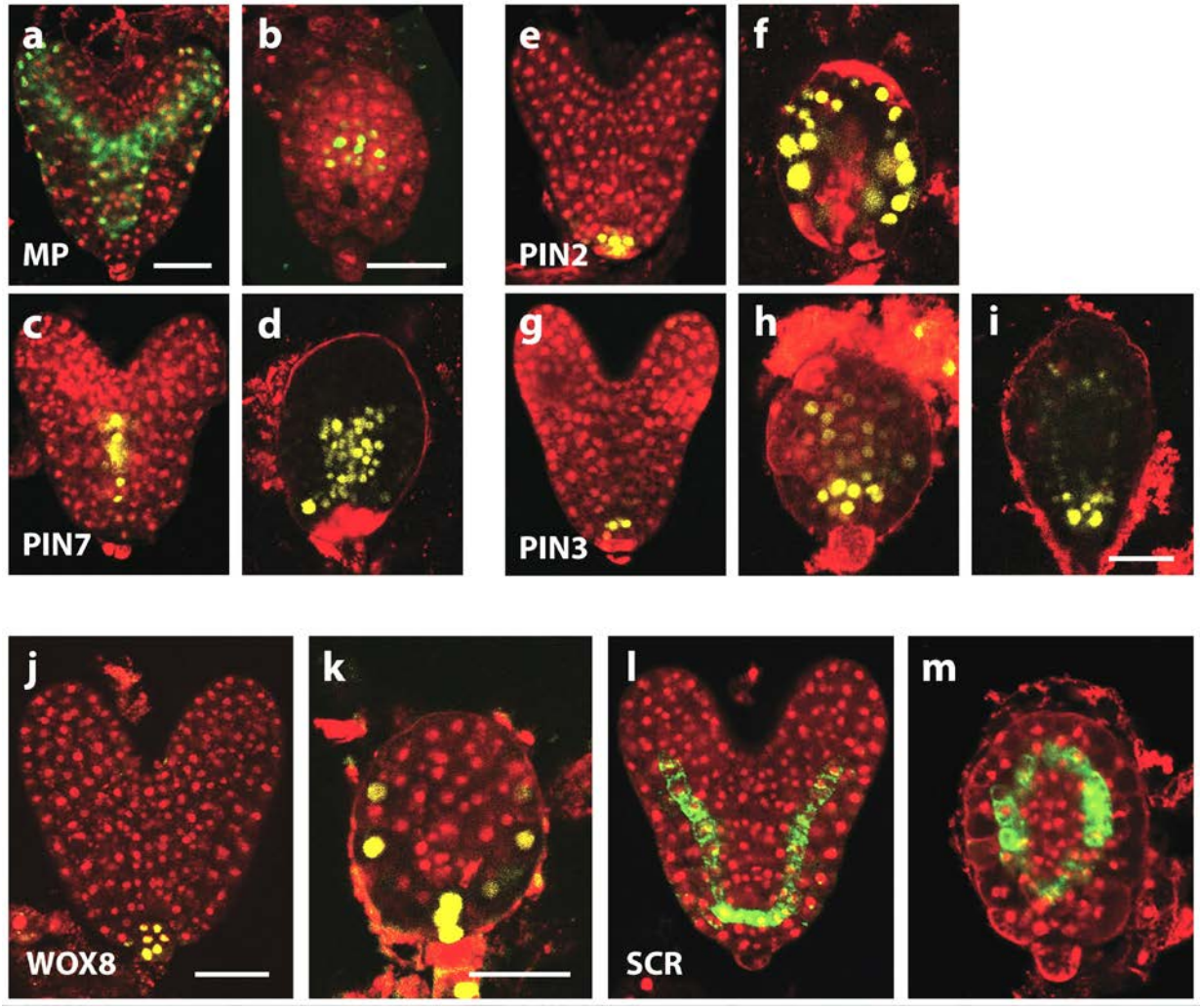

Figure S2.3. Auxin transporters and patterning genes are expressed in rings. The auxin transporter, PIN7::Venus-NLS, and regulator of root formation, pMP::MP-GFP (Krogan et al., 2012), are expressed in a generally similar pattern in wild-type (a,c) and triple mutant embryos (b,d) at the torpedo stage. In constrast, auxin efflux carriers with uncharacterized roles embryogenesis are expanded in triple mutant embryos; the expression of PIN2::Venus-NLS is strongly activated in enlarged surface cells (e,f), while PIN3::Venus-NLS displays more variable expression, some triple mutants express it broadly, while others display a ring pattern (h,i), compared to wild type (g). Similarly, the earliest marker of the basal cell lineage, *pWOX8::nYFP* (Breuninger et al., 2008), and a regulator of QC fate, *pSCR:GFP* (Wysocka-Diller et al., 2000), are upregulated in surface cells (j,k), and sub-surface cells (l,m) in triple mutants at the torpedo stage.

Table S2.1. HANL1/2 rescue constructs. The rescue activity of HANL1 and HANL2 was tested in the T1 generation of transgenic plants with the genotype *han-16/+ han1-56 hanl2-79* by counting the number of blimp-shaped arrest mutants. Description of the approach is outlined in the materials and methods section of chapter 2.

	wild-type (%)	**han (%)**	**triple (%)**	**n =**
HANL1::HANL1	1383 (88.7)	38 (2.4)	138 (8.9)	1559
HANL2::HANL2	994 (85.2)	46 (3.9)	127 (10.9)	1167

Table S2.2. Primers for complementation assays, reporter gene construction, and *in situ* hybridizations.

Purpose	Primer sequence	Notes
HANL1 swap	CAATTTCAccATGGGATACCAAACAAACTCTAA	~1.2 kbp
	gacttgctaagcttaagATCtATTCCAAGACATAAAAGAAGAAGCC	NcoI, (BglII) HindIII
HANL2 swap	ATATATAGccATGGGTTTCTCAATGTTCTTCTC	~0.8 kbp
	ctatgaaaaagcttaagatctCATCGTAAAATCGTGGACAAGG	NcoI, (BglII) HindIII
HB8 reporter	GTGTGCCTTATCgCgaGGGAC	~2.1 kbp
	GACTATTATTGaaTCCTCCTCC**CAT**C	NruI, BamHI
PIN1 reporter	CCGACTCCGTAACGTCGcGaGTGG	~4.3 kbp
	CGCgGatcc**CAT**cttttgttcgccgg	NruI, BamHI
PIN2 reporter	TTGGAAAGTAGTGAAGgGGCCcAGAAGG	~1.4 kbp
	CATCGTACATGTCTTTcCCGGgGAT**C**	ApaI, SmaI
PIN3 reporter	CAAATATTGCTTTTCGcGaCCCTTCGG	~4.6 kbp
	CCGTAAGCGAGGATCATGGatcCGTAGAG	NruI, BamHI
PIN4 reporter	GAGTGAGAAAATGtCGcGATGTGAGC	~0.8 kbp
	CCACCACTGTACGGATCCGTAG	NruI, BamHI
PIN7 reporter	CTGAGGCCGTTGCGTCGCgaTCTC	~1.9 kbp
	ccaggatcc**CAT**ATTGTTGTTCGCCG	NruI, BamHI

HAN 5'-F probe	TACACTTAGGGTTTTCAAACCAG
HAN 5'-T7R probe	*GAGATAATACGACTCACTATAGGA*AGATGCAAGAAGAAGCACCAG
HAN 3'-F probe	GGTTCAAACCGACCAGTACG
HAN 3'-T7R probe	*GAGATAATACGACTCACTATAGGA*ATAAACATCAACATCGTCTCTCG
HANL1 5'-F probe	TGAATTTGACCAAGACTGG
HANL1 5'-T7R probe	*GAGATAATACGACTCACTATAGGA*AGAAGTCCCAGTAGAGAATTC
HANL1 3'-F probe	CCTTCTTGTACGAAGAAGTC
HANL1 3'-T7R probe	*GAGATAATACGACTCACTATAGGA*GTAGTCTTCAATTTCATGATGG
HANL2 5'-F probe	AGAATTTAACGAAGGCTAAGAC

HANL2 5'-T7R probe *GAGATAATACGACTCACTATAGGA*AGTTCCAACTTTGGATCATC

HANL2 3'-F probe CTTGGAACCGTTCAAGAAG

HANL2 3'-T7R probe *GAGATAATACGACTCACTATAGGA*ACTTCTCTTTCTCCTTCTCTC

Small letters indicate nucleotides that are mismatched to the target sequence. Restriction sites are underlined, the start and stop codons of open reading frames shown in bold, and the T7 promoter sequences in cursive. In the case of PIN3 and PIN4, the BamHI site is positioned in frame after codon 18 and 25 of the open reading frame, respectively.

APPENDIX B

CHAPTER 3 SUPPLEMENTARY

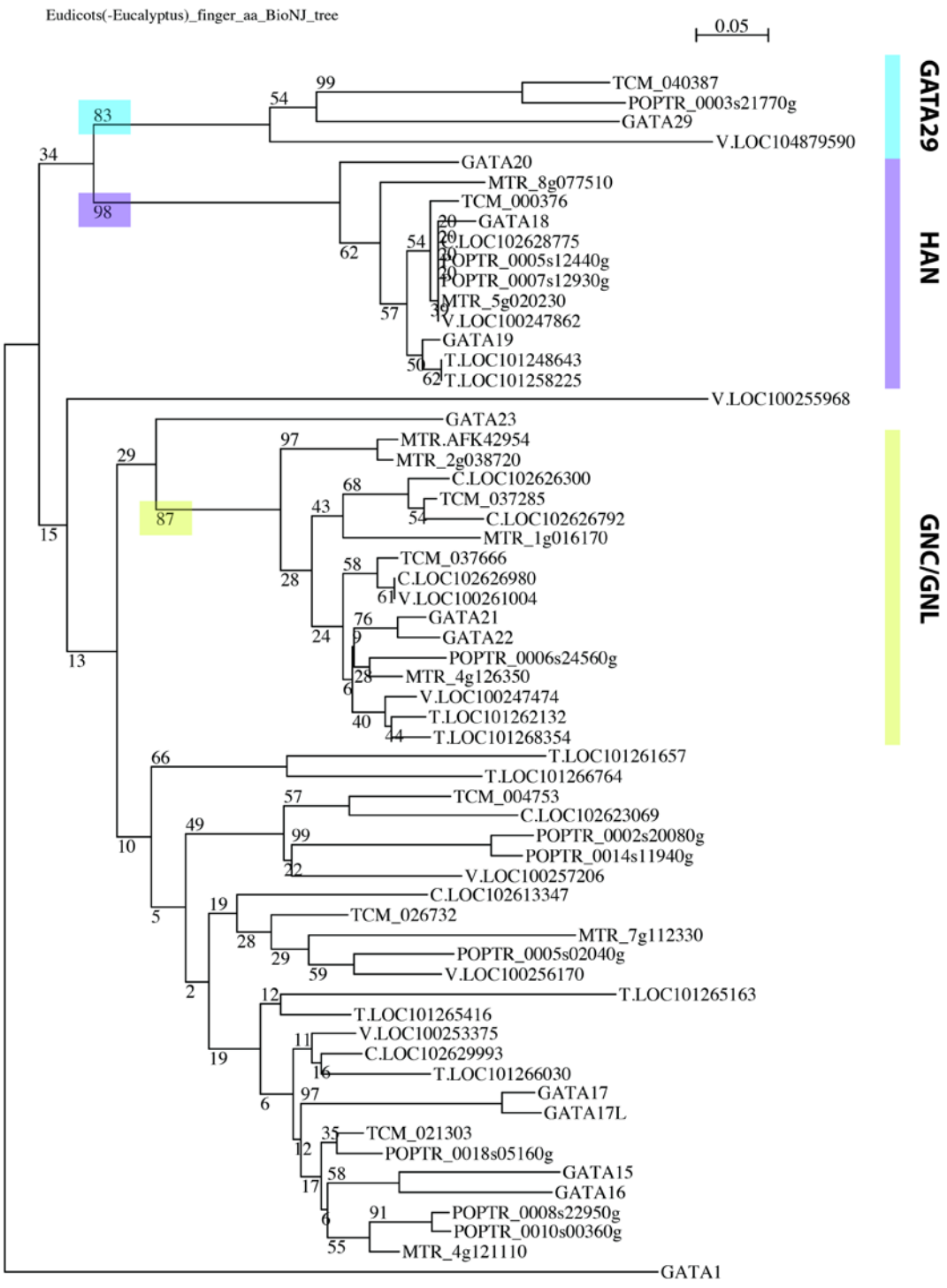

Figure S3.1. Phylogeny of B-class GATA genes using the complete amino acid sequence. Details including the species used and entries are found in the materials and methods section of chapter 3. HAN, GATA29, and GNC clades are highlighted.

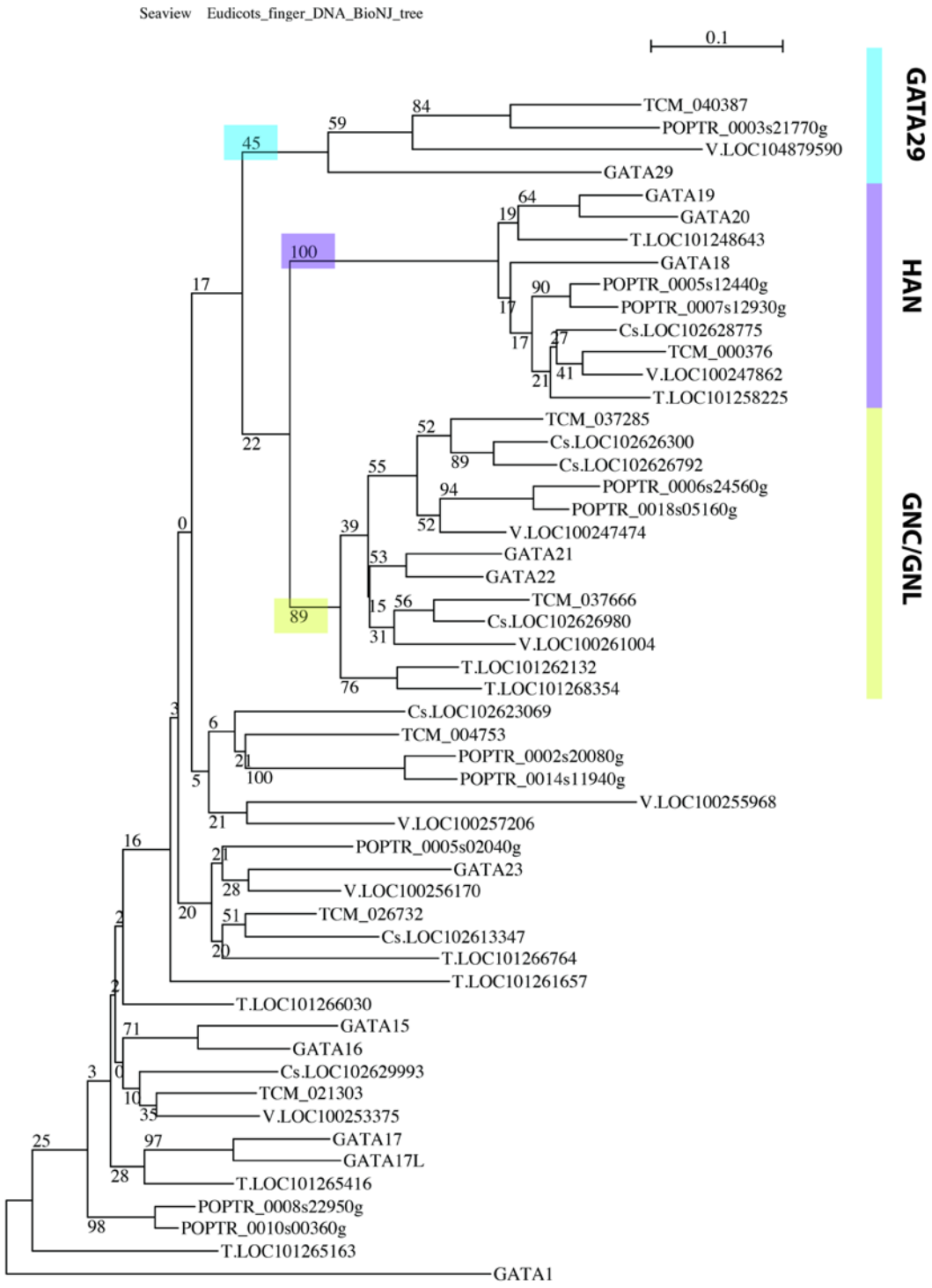

Figure S3.2. Phylogeny of B-class GATA genes using the complete DNA sequence. Details including the species used and entries are found in the materials and methods section of chapter 3. HAN, GATA29, and GNC clades are highlighted.

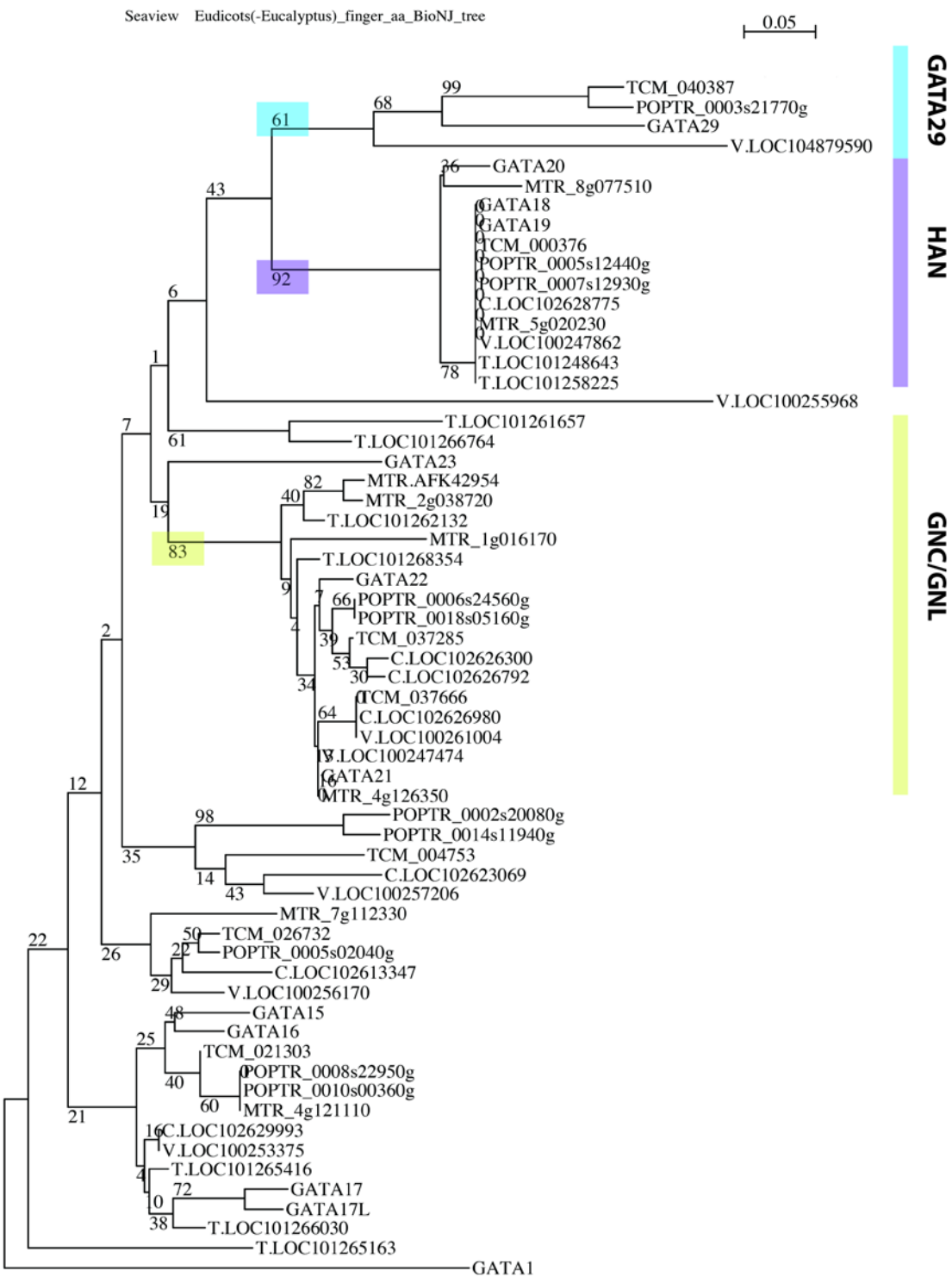

Figure S3.3. Phylogeny of B-class GATA genes using the zinc finger amino acid sequence. Details including the species used and entries are found in the materials and methods section of chapter 3. HAN, GATA29, and GNC clades are highlighted.

Table S3.1. Complementation assays testing the functional equivalency of HAN and B-class GATA factors, and mutant and chimeric variants of HAN, GATA29, and GATA16. Methodology is described in the materials and methods section of chapter 3.

Construct	wild-type (%)	intermediate (%)	triple (%)	n =
HAN::HANL1	1154 (88.4)	1 (0.1)	150 (11.5)	1305
HAN::HANL2	1652 (87.4)	7 (0.4)	232 (12.3)	1891
HAN::GATA 29	734 (87.5)	11 (1.3)	94 (11.2)	839
HAN::GATA21/GNC	993 (83.0)	67 (5.6)	137 (11.4)	1197
HAN::GATA15	782 (75.5)	1 (0.1)	253 (24.4)	1036
HAN::GATA16	989 (73.2)	18 (1.3)	344 (25.5)	1351
HAN::GATA17	990 (73.7)	4 (0.3)	350 (26.0)	1344
HAN GATA17-like	1179 (77.1)	0 (0.0)	350 (22.9)	1529
HAN::GATA22/GNL	950 (74.6)	1 (0.1)	323 (25.4)	1274
HAN::GATA23	1250 (77.6)	12 (0.7)	348 (21.6)	1610
HAN::GATA 1	672 (75.9)	9 (1.0)	204 (23.1)	885
HAN::HAN	1497 (91.1)	0 (0.0)	146 (8.9)	1643
HAN::HAN_mHAN domain	1048 (73.1)	1 (0.1)	385 (26.8)	1434
HAN::HAN_dHAN domain	1023 (75.0)	0 (0.0)	341 (25.0)	1364
HAN::GATA 29_HAN domain	170 (87.2)	3 (1.5)	22 (11.3)	195
HAN::GATA 29_mHAN domain	534 (76.7)	18 (2.6)	144 (20.7)	696
HAN::GATA16_HAN domain	326 (83.6)	30 (7.7)	34 (8.7)	390
HAN::GATA16_mHAN domain	391 (77.9)	1 (0.2)	110 (21.9)	502
HAN::HAN_EAR	933 (90.4)	6 (0.6)	93 (9.0)	1032
HAN::HAN_mEAR	200 (73.5)	3 (1.1)	69 (25.4)	272

Table S3.2. Primers for complementation assays.

Construct	Primer sequence	Notes
HANL1	CAATTTCAccATGGGATACCAAACAAACTCTAA	~1.2 kbp
	gacttgctaagcttaagATCtATTCCAAGACATAAAAGAAGAAGCC	NcoI, (BglII) HindIII
HANL2	ATATATAGccATGGGTTTCTCAATGTTCTTCTC	~0.8 kbp
	ctatgaaaaagcttaagatctCATCGTAAAATCGTGGACAAGG	NcoI, (BglII) HindIII
GATA15	GAAGCTGGccATGgTAGATCCCAC	~0.6 kbp
	AACTTAAGATCtAGAAAATTCGTTAAGC	NcoI, BglII
GATA16	GACGATCTAACAccATGgTAGATC	~0.6 kbp
	CCAGAAGAAAAAAAAGCttTAAGCG	NcoI, HindIII
GATA17	AAAAAAccATGgCTGAGGGATCAGAAGATAC	~0.8 kbp
	AGAGAAAagCTTAGGCGTAAACAGAGC	NcoI, HindIII
GATA17-like	TTCTTTTGTccATGGCAATGACG	~1.0 kbp
	GGTTAAGAGgacAATTgTCAACCACAAG	NcoI, BoxI
GATA21 (GNC)	CATTTGTTTATCTCATCcATGGATTC	~1.7 kbp
	GTGATGATaaGcTtAACCGTGAACCATTC	NcoI, HindIII
GATA22 (GNL)	ATCTCTATCGGccATGGGTTC	~1.5 kbp
	AAATACTgAcAATAgTCACCCGTGAACC	NcoI, BoxI
GATA23	ccaccaccATGgATCCAAGGAAGCTACTA	~0.5 kbp
	caccacagatCTTAGGCCAAAACAGAGCTAC	NcoI, BglII
GATA29	GATTTACAGACcATGGAAATGGAATC	~1.1 kbp
	CCTAATCAagcTTTATCCACAATCTTTCC	NcoI, HindIII

Small letters indicate nucleotides that are mismatched to the target sequence. Restriction sites are underlined, and the start and stop codons (forward and reverse complementary, respectively) of open reading frames shown in bold.

Table S3. Sequences for mutating/deleting/changing GATA29/HAN/IAA17 domains for swaps

>pCN-HAN (coding sequence of modified HAN cDNA in capitals, with start and stop codons in bold; NcoI, BglII & HindIII cloning sites underlined; EcoRI & SacI sites marking the beginning of the vector backbone in capital and underlined)

catgccaaccacagggttcccctcgggatcaaagtactttgatccaacccctccgctgctatagtgcagtcggcttctgacgttcagtgcagccgtcttctg
aaaacgacatgtcgcacaagtcctaagttacgcgacaggctgccgccctgccctttcctggcgttttcttgtcgcgtgttttagtcgcataaagtagaatact
tgcgactagaaccggagacattacgccatgaacaagagcgccgccgctggcctgctgggctatgcccgcgtcagcaccgacgaccaggacttgacca
accaacgggccgaactgcacgcggccggctgcaccaagctgttttccgagaagatcaccggcaccaggcgcgaccgcccggagctggccaggatg
cttgaccacctacgccctggcgacgttgtgacagtgaccaggctagaccgcctggcccgcagcacccgcgacctactggacattgccgagcgcatcca
ggaggccggcgcgggcctgcgtagcctggcagagccgtgggccgacaccaccacgccggccggccgcatggtgttgaccgtgttcgccggcattg
ccgagttcgagcgttccctaatcatcgaccgcacccggagcgggcgcgaggccgccaaggcccgaggcgtgaagtttggcccccgccctaccctcac
cccggcacagatcgcgcacgcccgcgagctgatcgaccaggaaggccgcaccgtgaaagaggcggctgcactgcttggcgtgcatcgctcgaccct
gtaccgcgcacttgagcgcagcgaggaagtgacgcccaccgaggccaggcggcgcggtgccttccgtgaggacgcattgaccgaggccgacgccc
tggcggccgccgagaatgaacgccaagaggaacaagcatgaaaccgcaccaggacggccaggacgaaccgtttttcattaccgaagagatcgaggc
ggagatgatcgcggccgggtacgtgttcgagccgcccgcgcacgtctcaaccgtgcggctgcatgaaatcctggccggtttgtctgatgccaagctggc
ggcctggccggccagcttggccgctgaagaaaccgagcgccgccgtctaaaaaggtgatgtgtatttgagtaaaacagcttgcgtcatgcggtcgctgc
gtatatgatgcgatgagtaaataaacaaatacgcaaggggaacgcatgaaggttatcgctgtacttaaccagaaaggcgggtcaggcaagacgaccatc
gcaacccatctagcccgcgccctgcaactcgccgggggccgatgttctgttagtcgattccgatccccagggcagtgcccgcgattgggcggccgtgcg
ggaagatcaaccgctaaccgttgtcggcatcgaccgcccgacgattgaccgcgacgtgaaggccatcggccggcgcgacttcgtagtgatcgacgga
gcgccccaggcggcggacttggctgtgtccgcgatcaaggcagccgacttcgtgctgattccggtgcagccaagcccttacgacatatgggccaccgc
cgacctggtggagctggttaagcagcgcattgaggtcacggatggaaggctacaagcggcctttgtcgtgtcgcgggcgatcaaaggcacgcgcatcg
gcggtgaggttgccgaggcgctggccgggtacgagctgcccattcttgagtcccgtatcacgcagcgcgtgagctacccaggcactgccgccgccgg
cacaaccgttcttgaatcagaacccgagggcgacgctgcccgcgaggtccaggcgctggccgctgaaattaaatcaaaactcatttgagttaatgaggta
aagagaaaatgagcaaaagcacaaacacgctaagtgccggccgtccgagcgcacgcagcagcaaggctgcaacgttggccagcctggcagacacg
ccagccatgaagcgggtcaactttcagttgccggcggaggatcacaccaagctgaagatgtacgcggtacgccaaggcaagaccattaccgagctgct

atctgaatacatcgcgcagctaccagagtaaatgagcaaatgaataaatgagtagatgaattttagcggctaaaggaggcggcatggaaaatcaagaac

aaccaggcaccgacgccgtggaatgccccatgtgtggaggaacgggcggttggccaggcgtaagcggctgggttgtctgccggccctgcaatggca

ctggaacccccaagcccgaggaatcggcgtgacggtcgcaaaccatccggcccggtacaaatcggcgcggcgctgggtgatgacctggtggagaag

ttgaaggccgcgcaggccgcccagcggcaacgcatcgaggcagaagcacgccccggtgaatcgtggcaagcggccgctgatcgaatccgcaaaga

atcccggcaaccgccggcagccggtgcgccgtcgattaggaagccgcccaagggcgacgagcaaccagatttttcgttccgatgctctatgacgtgg

gcacccgcgatagtcgcagcatcatggacgtggccgttttccgtctgtcgaagcgtgaccgacgagctggcgaggtgatccgctacgagcttccagac

gggcacgtagaggtttccgcagggccggccggcatggccagtgtgtgggattacgacctggtactgatggcggtttcccatctaaccgaatccatgaac

cgataccgggaagggaagggagacaagcccggccgcgtgttccgtccacacgttgcggacgtactcaagttctgccggcgagccgatggcggaaag

cagaaagacgacctggtagaaacctgcattcggttaaacaccacgcacgttgccatgcagcgtacgaagaaggccaagaacggccgcctggtgacgg

tatccgagggtgaagccttgattagccgctacaagatcgtaaagagcgaaaccgggcggccggagtacatcgagatcgagctagctgattggatgtacc

gcgagatcacagaaggcaagaacccggacgtgctgacggttcaccccgattacttttgatcgatcccggcatcggccgttttctctaccgcctggcacg

ccgcgccgcaggcaaggcagaagccagatggttgttcaagacgatctacgaacgcagtggcagcgccggagagttcaagaagttctgtttcaccgtgc

gcaagctgatcgggtcaaatgacctgccggagtacgatttgaaggaggaggcggggcaggctggcccgatcctagtcatgcgctaccgcaacctgatc

gagggcgaagcatccgccggttcctaatgtacggagcagatgctagggcaaattgccctagcaggggaaaaaggtcgaaaaggtctctttcctgtggat

agcacgtacattgggaacccaaagccgtacattgggaaccggaacccgtacattgggaacccaaagccgtacattgggaaccggtcacacatgtaagt

gactgatataaaagagaaaaaaggcgatttttccgcctaaaactctttaaaacttattaaaactcttaaaacccgcctggcctgtgcataactgtctggccag

cgcacagccgaagagctgcaaaaagcgcctacccttcggtcgctgcgctccctacgccccgccgcttcgcgtcggcctatcgcggccgctggccgctc

aaaaatggctggcctacggccaggcaatctaccagggcgcggacaagccgcgccgtcgccactcgaccgccggcgcccacatcaaggcaccctgc

ctcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagc

ccgtcagggcgcgtcagcgggtgttggcgggtgtcggggcgcagccatgacccagtcacgtagcgatagcggagtgtatactggcttaactatgcggc

atcagagcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgtaaggagaaaataccgcatcaggcgctcttccgcttcctcg

ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgca

ggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagc

atcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgt

tccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtc

gttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacg

acttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggcta

cactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggta

gcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacga

aaactcacgttaagggattttggtcatgcattctaggtactaaaacaattcatccagtaaaatataatattttattttctcccaatcaggcttgatccccagtaagt

caaaaaatagctcgacatactgttcttccccgatatcctccctgatcgaccggacgcagaaggcaatgtcataccacttgtccgccctgccgcttctcccaa

gatcaataaagccacttactttgccatctttcacaaagatgttgctgtctcccaggtcgccgtgggaaaagacaagttcctcttcgggcttttccgtctttaaaa

aatcatacagctcgcgcggatctttaaatggagtgtcttcttcccagttttcgcaatccacatcggccagatcgttattcagtaagtaatccaattcggctaag

cggctgtctaagctattcgtatagggacaatccgatatgtcgatggagtgaaagagcctgatgcactccgcatacagctcgataatcttttcagggctttgtt

catcttcatactcttccgagcaaaggacgccatcggcctcactcatgagcagattgctccagccatcatgccgttcaaagtgcaggacctttggaacaggc

agctttccttccagccatagcatcatgtccttttcccgttccacatcataggtggtccctttataccggctgtccgtcatttttaaatataggttttcattttctccca

ccagcttatataccttagcaggagacattccttccgtatcttttacgcagcggtatttttcgatcagttttttcaattccggtgatattctcattttagccatttattatt

tccttcctcttttctacagtatttaaagataccccaagaagctaattataacaagacgaactccaattcactgttccttgcattctaaaaccttaaataccagaaa

acagctttttcaaagttgttttcaaagttggcgtataacatagtatcgacggagccgattttgaaaccgcggtgatcacaggcagcaacgctctgtcatcgtta

caatcaacatgctaccctccgcgagatcatccgtgtttcaaacccggcagcttagttgccgttcttccgaatagcatcggtaacatgagcaaagtctgccgc

cttacaacggctctcccgctgacgccgtcccggactgatgggctgcctgtatcgagtggtgattttgtgccgagctgccggtcggggagctgttggctgg

ctggtggcaggatatattgtggtgtaaacaaattgacgcttagacaacttaataacacattgcggacgtttttaatgtactgaattaacgccgaattaattcgg

gggatctggattttagtactggattttggttttaggaattagaaattttattgatagaagtattttacaaatacaaatacatactaagggtttcttatatgctcaacac

atgagcgaaaccctataggaaccctaattcccttatctgggaactactcacacattattatggagaaactcgagtcaaatctcggtgacgggcaggaccgg

acggggcggtaccggcaggctgaagtccagctgccagaaacccacgtcatgccagttcccgtgcttgaagccggccgcccgcagcatgccgcgggg

ggcatatccgagcgcctcgtgcatgcgcacgctcgggtcgttgggcagcccgatgacagcgaccacgctcttgaagccctgtgcctccagggacttca

gcaggtgggtgtagagcgtggagcccagtcccgtccgctggtggcgggggggagacgtacacggtcgactcggccgtccagtcgtaggcgttgcgtg

ccttccagggacccgcgtaggcgatgccggcgacctcgccgtccacctcggcgacgagccagggatagcgctcccgcagacggacgaggtcgtcc

gtccactcctgcggttcctgcggctcggtacggaagttgaccgtgcttgtctcgatgtagtggttgacgatggtgcagaccgccggcatgtccgcctcggt

ggcacggcggatgtcggccgggcgtcgttctgggctcatggtagatcccctcgatcgagttgagagtgaatatgagactctaattggataccgagggga

atttatggaacgtcagtggagcatttttgacaagaaatatttgctagctgatagtgaccttaggcgacttttgaacgcgcaataatggtttctgacgtatgtgctt

agctcattaaactccagaaacccgcggctcagtggctccttcaacgttgcggttctgtcagttccaaacgtaaaacggcttgtcccgcgtcatcggcgggg

gtcataacgtgactcccttaattctccgctcatgatcgataacattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaagggcgatcg

gtgcgggcctcttcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgta

aaacgacggccagtgagcgcgcgtaatacgactcactatagggcgaattgggtaccatttaaatgggcccGAATTCgatgggagaagactgaag

attcgaaaatgtttgaagtgccaattcatatccaattatcattttacctttgaaattcttatgatccatctctgcattttccagtaccagctttttgacctttttagtac

ccaatagattgatagatacctatgtgggttttatatacgcaaaaataaaattggttttatgtgatttcaaagtttctagtggctattgcctaaccaatggtaaatgt

acacgtttaattgctcttttaaaaaacaccaatagactagttgtttacatagtagagattatacgatgaattgtgaggataatgcgaataggaaacatctaggtt

aattgtttggccattttgagagtggatatcgaaatatttgaagtttaaagtagttataataattttctatttagagaaaatcgacgattatccgaattaagatggaat

gtttccctctttatgcactaaaagctaatgggcaacaagaacaagaaaaaagttggaaatgttgtgttgtgttgaagacaaaagaatataagtggaagattca

aagtggagttatttaattgggaaacgttaagagcgaaacacatatataaagaaagaaacttgttggataaagatgatacaaatccaatatacatttgtgccct

agaagaattggtcaagctagcaaagatgtaccatggaccaaaagtgaccacctttgatcgtatttatatatgcataacttatgttttttgttatatcatgattgat

acctctcccaatcttggggcattcctctagttggtccaacaatttttcttttccactttgccctttcacttacgttggaccactcccacatttgtttctcttctttttt

gcattttcctatatttatctcctttgatttatgtgattatttcatccattacacaaatcatttaaatgtatatattacatttgaaaatgtttgaagtgcaaattcatatcca

actagagtaaagtcttactaacaccaaaatttttatatttgatcaatgttaagataatttgtgtttctaaattgttttagaaaagcaaaattggagctacaagtatac

tacagtcaaccatatgtatactagttgaccgaatgcaaaaactgtagctgaagtgttttaaaaatatacaaaacagttcataacatacgcatctttaggaaatgt

gaagagacgctttcgagaggacaacatgaatttacaggagacgtagacaaaaaaaaaagtcttaacatatgacagctggaatttgtactcttataattgttgt

tatttaattgcaacaaataaaaacttattttaaagtggttatacattaaattttgccatagactctgattcaatctaaaatactaaacgcattgaaactcaataagttt

aaaattattccttacattttgtgtgaaaactattactaatctagagactaactacctctgtatagacatcactattttcacattatttttgaatgtgtataacagttcta

aattgatatttatatttttcacattttctagtgtttatgatataatttgaaaagcagaaagcattcttagttagcccctataatgatttttttacatttttttatgaaggctac

catacaaacagaattttcaaacaaaaaaaatgaaaaaaataaagagaaagcagaatccatcaaaagggcagagatgagaaaacattccagtcgtgttttg

taacttttgaggaatgtcccgaatattttaattaatgtgaccaagacgactcctaaaccaacacacatactactatattctattttaattaaacagttaagaataaa

ttagttgttgtttaatgacatatcattagatcgctttaacatgcatgcatgcggagaattaaatctcgcgggcttctgtccacgtggcgaccctgattaacgaca

aacggattaattcaagtaaacgtcgtcaccaaaaaacacgatctctgtcttggaactcgcgacacttttttggggaacgttgtggttcccaccgaatccacc

gaagccaacatcatcatcatttgcatttgcatcatcccatcaaacatcgtgtagttttattgctctcgttctggactacgagccggtcttcttttgttttgcgtata

gccgtttggtttttccttttggactactaaaatactcttgtccacacaaacgtcaagatacttaaaacatctcaatatttgcatatgaactagttagtagtatacat

gtatatactcgtcggaagtgacatttttaagattgtgaaatattttgcgatatttgaatttaaaatagattaaaaacgatttaaatttcactgtaaaagcaaatccc

ggtgaataattacagaaagccgatcttgtaccgtagttaattacgtagttagtttatgggcattagagaaatgtttggtttattcaaatcactaaacaaccacgc

gccattatagctcgtgagagttttagaacgtgtgggtcgatgggagcgtattagtcacgggtggttgggtcgcgtgtgtgatatttagtagtagctttgcatgt

agtggcgaagcgtaacgctttgtgtctcggggtaatcacggttttatgtttcgagtttatcaaattactataaatatagattacgtacaataattaagagttataa

attgtataaatataaacgttcgacagagcatcttttggaaagtggtggtcgtcccattagtcgaactaagtttcggcaatctcaagacgatctgattttcacctttt

ttaatagtatatctattttcaaaaaaaaaattgcaaggaattgataggtaacttgcgattgcgaaacaactttaacataattttaaaaaaacttgagaaacaatga

gatagcgtagtaacaaaatagctaccattttcatgtatgtgttgatcaactgcatcgtcatttgaaagtttattggataggtctcgcttttcaggtccgagtgttta

gtatgatatttatcatatgtgtgttcgacaatgatacatatgtataagactaagcactaagcagattcgtccggattactcaacaaaacaaaataagattgagat

ttgagtctttggatcatttttctttttgggtcaaaacgttgatatttgtaagtttattgataataacaaaaagatataaattattgtttcttaaatagtcataaattacata

cagttcctattttatcccatttttatagttaaataattaattgtgctaaatgatttaccaattcaatatatatatatatatatatagtattttatatttaaattatgttactgtt

aatgaaatgttataaccaataacttttcatatatatatttcaaggatttttttttcaggtatgatgactacataatcagctacataatactacaaaacgatgaagtatt

ggaatgcaaaaataatcagctattgactgaccatgatttgttatatttttgaaactcaattgcatgttttgttttgccagcatatatacaaataacatatctatttcga

tgaactataaaaccccgtgattatccatggaaacatatatgtgtaccatatcattataatttggcgatagagaacaaaaaaaagttatcatgtggtattactatta

ggtataatgatttaagaggtgggtccaatagcttttccttttgtttcttttcactagttcaagatcaaatttatggggagtattattagttctttgacctttccttttaga

gagtccttctttgacccatcaagattcaagagcagaatctctctcttgctaagctaagaaaacatgaaaagtggaataaatatttttaaaagaaagaagagtat

agctcatacatgtgtatgcattggttaattgggggtatgttgctttgatgattgtaggaaggtacttcacttgaatagaggggattcttccctttatctttgggttc

ttcacaatctaatctattagcttttcttttcatgcagtatttttcttgttactctaaacctcttccaaatcttcttggatattctcatatcattatgattccaaattcgtttctc

tttattaaaaaccccaataagaattggattctcatggttttcgatacgttattcagccatacgacgactatgagtgattctagtgagactaattgaggatttttttttt

ctatttttacatgagaatactgaagatgtttaatttatctctgattaaatataatccattatacagcatattacgataagtaacaccgcctcgtttgtttttgtttagct

agtaaaccgaaattttgatcgaatacattatataaaaattcatgtaaaatactctacaatataatactctcttgtagtctgaatgtaattataaaaaaatgttttgtgt

agacattaggtaaaccatctagagtattagaaacaaaataatgtaagtaaacgtaaaactaagacgatttcatgactaaattattttgaaggtcgaaccaacc

aatcttaaatacacatatatgaagaaatatatgttttcgaacaaatggcaatcaagaaaatgctttttgtgtgttaatgaaatataaggttgacacatcaatatcg

cgttgatattctcattgtagattgttggttatgtagggagactaataatgcgaactagaaaaagatatgtgagtcaactaaggttagtgttgatatagcaactct

cttatgtttttaaatttgccataaaataggttgttagtccaaatagccgatataaacatcaatagaagagtaattaatcgttccatcggaacatccgatgaaattg

gaacgagacctatacagaagaaaacaaaagagcaactaatcatcttgatcacatgcccaatatgcatttaagtaaattggagagaatgatatagaaaaaca

agacattatctacaactacaagatatcaaagcactcggttaaccgcgtccctgtgaagtgttagttatcagtacatgaattgtgatttcagcatcttggtccttttt

ttttttggacaaccgtcatctttttcttgatcggtttttataacgtacaaacatatactatttcaacagttttcttttaatttttttttttttgaatatacactttctttctttatt

tgtttgattacaaaattttcaaaaaatagtgttcggcaaaaaaaaaaattagatcggccaatgcattatatcccctagcgtaagctctttttcaaccaatcgaga

gtccatgttttcaataaaaatagggttttattagtaacagtagttaaagcatttaacatacgtaatcccaaaaaaataaaataattctgtatagggacgaagact

agtattcgttgacgacaagagtagagtagactctagctagctgcaagttttgaagaaatctaataaccaaatctaagtcattttacttacatcatcatggatcat

cattaaccacatgaggctgagcagatatcgaacgtgtgattgcgatgtggagaagtaaaggacgaggccacgcgcaactgagcaggagcgcgtgaga

ttcagggttttagggcagagatacggaaagagacagaggttttatagcattatattgaaagagaacaaataaaaataaaaatatgagcctcgtgatcggag

atttttgcagtgatgactgtactcttcttcctcttctctccccccttgtcacgtgggctaaattttgaccacccttcttcaatgaccatcaacatcgcatcacactctc

tccttcttcttcttctacgtttttttattttctgttaaatttcctccacaaaatcgataatttcaatttttttcatacaatatcatagttctagtaaagtaaacattcgtgtgtga

gtgacataataattgcttcaaaaagaaataacgaattaataagagttgtgtgcttgatggtccaaaagtcgtatgtttttcaccatcacaataagatttaccatat

ctacttaaaccaaacttcaatctatgatatttggaagctagcaatgaatgattttgttttagcatatgttaccgagcctattcatctttagattgaaagttttgtgtcc

aacaactagcaaacccttcgctttttaaaccaagatcagtttttctacatagtttcgacttttttgttattttacttcccatcatattttgttttatttactactgacattttc

ataatgtacggtacattttcatatattagcatccttatctaaaagctaatcaactccataacgtacttgaaatagggatattttagttgaattattcaaaacaatcac

acagtgacacagactctatgtagaccaagcatgtgttatgtaattagaaaaaataaactgtgattaaaatatcttttaacttgttaacttggctgcctaaagaaa

cttgatactgtatcatcgttcatcgtattttgtaattaataaaacataaaaaagcaaggaaaacccatttgagtaaatgatatattaatattacttcgtaaaaagct

gcgtccccacatattctattgattttaaatactattcgtattaatataaacctgcatataagcgtacatatacatttctgattctatatataggtgtgagaccttcttct

ccattctacacttcactccaacgtgaaacctgcaaactgaacttatctcttcatattttctttagcttctcttaattttacacttagggttttcaaaccagtagagag

agagagaagtgtgtaagagcc**ATGG**CCCAGACTCCGTACACTACTTCAACGCAGGGGCAATATTGTCAT

TCTTGTGGAATGTTCCACCACCATAGCCAAAGCTGCTGCTACAACAACAACAACAACTCCAA

CGCCGGTTCTTACTCGATGGTCTTCTCCATGCAAAACGGTGGCGTTTTCGAGCAGAACGGTG

AGGACTATCATCACTCTTCCTCCCTCGTTGACTGCACTCTCTCTCTTGGAACTCCTTCTACGA

GGCTTTGTGAGGAAGATGAGAAACGTAGACGCTCTACTTCATCTGGTGCTTCTTCTTGCATCT

CCAACTTTTGGGACTTGATTCACACCAAAAACAACAACTCCAAAACGGCACCGTACAATAAC

GTTCCTTCTTTCTCCGCTAACAAGCCAAGTCGCGGTTGTTCCGGTGGTGGTGGTGGCGGAGG

AGGCGGTGGCGTAGGTGACTCTCTTCTCGCTAGACGCTGTGCCAACTGTGACACTACTTCTA

CTCCACTATGGAGGAATGGTCCTAGAGGCCCTAAGTCCCTATGCAACGCATGCGGCATTCGT

TTCAAGAAGGAAGAGAGAAGAACTACTGCGGCTTCAGGAAACACCGTCGTCGGAGCTGCAC

CGGTTCAAACCGACCAGTACGGGCATCACAACTCTGGCTACAATAATTACCATGCTGCCACT

AATAACAACAATAATAATGGTACTCCGTGGGCTCATCACCACTCGACGCAGAGGGTTCCGTG

TAATTATCCGGCAAATGAGATCAGGTTCATGGATGATTACGGCAGTGGAGTAGCAAACAAC

GTTGAATCCGACGGTGCTCACGGCGGTGTTCCGTTCCTTTCTTGGAGGCTTAATGTAGCGGAT

AGGGCAAGTCTTGTCCATGACTTTACCAGATCT**TAA**gcttagaatagaaaacgacgttagattctcttaaaatttccatgt

ggtttaattatgtctaattaaaaaaaaaccaagaaaatatttctgaaaaaaatagaagtgtatatcgagagacgatgttgatgtttattgatagtgacggtctaa

cttatggtttatctataaagtaccgttctggtcaatttttaaaatcattttctaatatttattttaaaaatataaccttctttttatggaaatcttctcttctctagtctataa

aataaacactatagttagattcttctattgattacacagtgatcgtcataaaacattacagatgtaatcaaaacattagaaagaaaaacaaaaaatgaaagatg

tgcataagagaaaacacacagaccaaaaatatagacataggatataactcttttaatgtatatagtttagagttgaacggagaagtttcagttccacctgagt

ccctgaatttcacaactgagactttaataataataatattcttactattggttaataaaaaatagagaaaataaaataaaaattagggtgtttggtagtgaagatt

gaaatctgatcgtcgaattggggagagattagattctgatcgtaccaatgtgggttttggtactgttgtgttttcagttttggatcagagagcattgtagtgattt

aatgaacatatatttgtcgtattatgctctgcttcttcctcctttatattttgaccgtcgttttgtcctttactcactcctctcttcttgacctttttgttcgtgtcaatatgtg

ctctcaagaatattgtttttcttcctttttaattacttacttttccgatttttttattactgaaccgattctgcaacatgttttgtactagtaaccagctttttaatccatctct

ctctgcacgatcatataaaatatacgttatttttttttcaattcaggcgtcgaccgGAGCTCgacgtcagcttggcactggccgtcgttttacaacgtcgtg

actgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttc

ccaacagttgcgcagcctgaatggcgaatgctagagcagcttgagcttggatcagattgtcgtttcccgccttcagtttaaactatcagtgtttgacaggata

tattggcgggtaaacctaagagaaaagagcgtttattagaataatcggatatttaaaagggcgtgaaaaggtttatccgttcgtccatttgtatgtg

>GATA15 (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI, BglII & HindIII cloning sites underlined)

cc**ATG**GTAGATCCCACCGAGAAAgtgagtgatctgattccctcttctcgctctgatatctagcgatcggattgattaagcttaacgcgct

actacttctgtgtttagGTAATCGATTCAGAATCAATGGAAAGCAAACTCACATCAGTAGATGCGATCG

AAGAACACAGCAGCAGTAGCAGTAATGAAGCTATCAGCAACGAGAAGAAGAGTTGTGCCAT

TTGTGGTACCAGCAAAACCCCTCTTTGGCGAGGCGGTCCTGCCGGTCCCAAGgttagtctcttccccgat

cttaatttgacctaattgttatctagttcgatctgatccgattgctttttgtgtagTCGCTTTGTAACGCATGCGGGATCAGAAACA

GAAAGAAAAGAAGAACACTGATCTCAAATAGATCAGAAGATAAGAAGAAGAAGAGTCATA

ACAGAAACCCGAAGTTTGGTGACTCGTTGAAGCAGCGATTAATGGAATTGGGGAGAGAAGT

GATGATGCAGCGATCAACGGCTGAGAATCAACGGCGGAATAAGCTTGGCGAAGAAGAGCA

AGCCGCCGTGTTACTCATGGCTCTCTCTTATGCTTCTTCCGTTTATGCT**TAA**cgaattttctagatcttaag
ctt

>GATA16 (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI & HindIII cloning sites underlined)

cc**ATG**GTAGATCACAGTGAAAAGgtaaaaatgatttgcttctcttctgatctctatttttttgattctgacgacgatgatctaataataagctg
atcgctcttctgtgtttgtttcagGTCTTATTGGTTGATTCAGAAACCATGAAAACAAGAGCTGAAGATATG
ATCGAACAGAACAACACTAGTGTTAACGACAAGAAGAAGACTTGTGCTGATTGTGGAACCA
GTAAAACTCCTCTTTGGCGTGGTGGTCCTGTTGGTCCAAAGgttcatacataaaccctaatttcactctgattttgggg
aaaaaatcgaggttgatcatttgatcttgtctctgatctctgattctgatatgttttggttttggttctgtgtagTCGTTGTGTAACGCGTGTGG
GATCAGAAACAGAAAGAAGAGAAGAGGAGGAACAGAAGATAATAAGAAATTAAAGAAATC
GAGTTCTGGCGGCGGAAACCGTAAATTTGGTGAATCGTTAAAACAGAGTTTGATGGATTTGG
GGATAAGGAAGAGATCAACGGTGGAGAAGCAACGACAGAAGCTTGGTGAAGAAGAACAAG
CCGCTGTGTTACTCATGGCTCTTTCTTATGGCTCTGTTTACGCT**TAA**agctt

>GATA17 (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI & HindIII cloning sites underlined)

cc**ATG**GCTGAGGGATCAGAAGATACGAAAACAAAACTCGACTCTGCCGGAGAGTTATCAGAT
GTTGATAACGAGAACTGCAGTAGCAGCGGAAGTGGCGGTGGTAGTAGTTCTGGTGATACCA
AAAGGACTTGCGTTGATTGCGGAACTATCCGAACTCCTCTTTGGCGTGGTGGTCCTGCCGGA
CCAAAGgtgagatttttaactaacccttcatttcaatcttcttatagtctattgggttgttgttagattccgatctagtagatctaatcaatggatcgcgaaat
agcatctatattatgcatgatcctattttaggttttatcgtttttcaaatttatttaggtgatgaatcaatttagtgaccggaaagatctgatctcatattgctctgaag
ttcaaatttttcttatttctttggtttgtctttgctatgtagTCATTGTGCAATGCTTGTGGGATCAAGAGTAGGAAGAAG
AGACAAGCAGCACTTGGTATGAGATCAGAGGAGAAGAAGAAGAACAGAAAAAGCAATTGC
AATAATGATCTAAACCTCGACCATCGAAACGCCAAGAAATACAAAATCAACATAGTTGATG
ATGGCAAGATCGACATCGATGATGATCCGAAAATTTGCAATAACAAGCGTAGTAGTAGTAG

TAGCAGTAACAAAGGAGTGAGTAAGTTTTTGGATTTAGGGTTTAAAGTACCGGTGATGAAG
AGATCAGCGGTTGAGAAGAAGAGGTTATGGAGAAAACTCGGTGAGGAAGAAAGAGCTGCT
GTGCTTCTCATGGCTCTCTCTTGTAGCTCTGTTTACGCC**TAA**gctt

>GATA17-like (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI & HindIII cloning sites underlined; a small insertion / cloning artifact is present between the stop codon and the HindIII site)

cc**ATG**GCAATGACGGAAGAAACCAAGACGACGAAGCTCGAATCGGCGGGAGATTCGTCCGA
TGTAGACAACGGGAACTGTAGTAGCAGTGGAAGTGGAGGTGATACGAAGAAGACATGCGTT
GATTGCGGAACAAGCAGGACTCCTCTTTGGCGTGGTGGACCTGCTGGCCCTAAAgtaaggatcatga
tctctctctaccattttcttttaaacccttaaaaatccaagctttaaggtctcatgtgttagatctagtcaaataggtacttgaatccacttaaaatgtaggttgatct
tgtgtgcttagctctgttgctagtgggttttgtctcattgatgtgatcatgtgtatacatagtctaggtatgcgaggtttgtgtcaatctgtggttgatatttaaatttg
ttatgtgatctctatgttttgttttgattctgtgtttgaagccttattactaattgatcaggttaagattattgttacactgtttttatgatgccttttttactaagagatggt
tatgtctctttaattttagatctgtttgatgattgttcactatttgttctctggtctgatcatggattgactttgtttgcagTCATTGTGCAATGCGTGT
GGTATCAAGAGCAGAAAGAAGAGACAAGCAGCTCTTGGAATTAGACAAGATGATATTAAAA
TCAAGAGCAAAAGTAATAACAATCTTGGTCTTGAGAGCCGAAATGTCAAGACCGGAAAAGG
CGAACCAGTTAATGTTAAGATTGCCAAATGTGAACCGGGAATTGTCAAGATTGCCAAAGGT
GAACCAGGAAATGTCAAGAACAAAATCAAGAGAGATCCTGAGAATTCCAGTAGCAGCAATA
ATAACAAGAAGAACGTGAAAAGGGTTGGTAGATTTTTGGATTTTGGGTTTAAGGTACCGGCG
ATGAAGAGATCGGCTGTGGAGAAGAAGAGGCTATGGAGGAAACTTGGGGAAGAGGAACGA
GCCGCAGTACTTCTCATGGCTCTTTCTTGTGGT**TGA**gacctaaagctt

>GATA21/GNC (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI & HindIII cloning sites underlined)

cc**ATG**GATTCAAATTTTCATTACTCGATAGATCTTAACGAAGATCAAAACCATCACGAACAA
CCCTTTTTCTATCCTCTTGGATCCTCTTCCTCGCTTCATCATCATCATCATCATCATCATC

AAGTCCCTTCTAATTCTTCATCTTCTTCTTCGTCCATTTCATCGCTCTCCTCTTACCTCCCTTTC

TTGATCAACTCTCAAGAAGATCAACATGTTGCCTACAACAACACTTATCACGCTGATCATCT

CCATCTTTCTCAACCCCTCAAGgtaaatttacaattacatatcaaccttaagtgtgtgtgtataaatgttattcttgtacgatcgttctttttag

tttagcagttaaaatcatgcctgatcatgtaaaattcatacaataatattagtttctgaatttgtatgtcgagaattctatatagacgaaatttaatttatcaaataaa

aaggagggatcttaagaagccatatcattatattattatgtacaattccatctgttgttgattataaaccttgatcggtgatatatatataccagGCCAAGA

TGTTTGTGGCTAACGGTGGATCATCAGCATGCGATCACATGGTGCCAAAGAAGGAGACAAG

ACTGAAACTAACGATAAGGAAAAAAGATCACGAAGACCAACCCCATCCTCTTCATCAAAAC

CCGACAAAACCCGATTCAGACTCCGACAAGTGGTTGATGTCCCCAAAGATGCGGTTGATCAA

GAAAACAATCACCAACAATAAACAGCTCATTGATCAGACTAATAATAATAATCATAAAGAA

AGTGATCACTACCCTTTGAATCATAAGACTAATTTCGACGAGGATCACCATGAAGATCTTAA

TTTCAAGAACGTCTTGACCAGGAAGACCACGGCCGCGACCACCGAGAATCGCTACAATACA

ATCAACGAGAACGGTTATAGTAATAACAATGGCGTGATTAGGGTTTGTTCGGATTGTAACAC

CACCAAGACTCCTCTTTGGCGAAGTGGACCTCGAGGTCCCAAGgtagattaatcacacttacttcaatcgaaattg

tcttgatatatgtcatactatctatatctacctagtttagttcacatttttaaaattaatagttatatttcaagaatgtagatatttaatttatttaaacctgatacatccttt

aaacattaatactttaatcaatgatactcataagtccaacattctatgatttttttttgtcatctacagttcgcatactaattttcatttatcattgcagTCTCTTT

GTAACGCATGTGGTATACGGCAAAGAAAGGCAAGGCGAGCCGCTATGGCCGCGGCCGCTGC

AGCCGGCGACCAAGAGGTGGCGGTAGCGCCCCGAGTGCAACAATTACCGCTGAAAAAGAAG

TTGCAAAATAAAAAAAAGAGATCAAACGGAGGGGAAAAATACAATCACTCTCCTCCAATGG

TGGCCAAGGCCAAAAAGTGCAAGATCAAAGAGGAAGAGGAGAAGGAAATGGAAGCGGAAA

CGGTTGCCGGAGATTCAGAGATCAGCAAATCTACAACTTCTTCTAATTCTTCGATTTCGTCAA

ACAAATTTTGCTTCGATGATTTGACAATAATGTTGAGCAAAAGCTCAGCTTATCAACAAGTG

TTCCCACAAGATGAGAAGGAGGCTGCTGTTTTGCTCATGGCTCTGTCGTATGGAATGGTTCA

CGGT**TAA**gctt

>GATA22/GNL (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI & HindIII cloning sites underlined)

cc**ATG**GGTTCCAATTTTCATTACACAATAGATCTCAATGAAGATCAAAACCATCAGCCTTTTT

TCGCTTCTCTTGGATCCTCTCTTCATCATCATCTACAACAACAACAACAACAACAACAACATT

TTCATCACCAAGCTTCTTCTAATCCCTCTTCTTTGATGTCACCGTCTCTTTCCTACTTTCCTTTC

TTGATAAACTCTCGCCAAGATCAAGTATATGTTGGGTACAACAATAACACTTTTCATGATGT

TCTTGATACCCATATCTCCCAACCTCTCGAGgtataatttcaaatgaatttcatgtcaatatattaactcttgatccttatatatatg

gtttatattcatatggaaatttcagACCAAGAACTTTGTATCTGATGGTGGTTCATCATCAAGTGATCAAATG

GTGCCCAAGAAGGAGACACGACTAAAATTGACGATAAAGAAGAAAGATAATCATCAAGAC

CAAACCGATCTTCCTCAATCCCCAATAAAAGACATGACAGGAACTAACTCGCTCAAGTGGAT

ATCTTCGAAGGTGAGATTAATGAAGAAGAAAAAGGCGATTATTACCACCAGCGACAGCAGC

AAACAACACACTAATAACGACCAATCCTCAAACCTAAGCAATTCGGAAAGACAGAATGGTT

ATAACAACGATTGCGTGATTAGGATTTGCTCCGATTGTAACACAACCAAGACTCCTCTTTGG

AGAAGTGGTCCGAGAGGTCCCAAGgtattaatcacacgtctgttaattacacatatatgcgtaagctttttttggttttaatgttaaaagt

agatgcctaagtaatgtcatatatactcatattgattattagaagctcacgaagaatatgtctttttatagtactaattaatagcaaatgaaaagtaaatctacaaa

tctctaactggtaaaaacaatttcatgtagcattataagttggtttaaacagattaatctttgagatcttctaaaattcatgttacaagataattaactaaattatcat

gttttttcttgttttgttattacagTCTCTTTGTAACGCTTGTGGAATAAGGCAAAGGAAGGCCAGGCGGGCCG

CTATGGCCACGGCAACCGCAACCGCAGTCTCTGGCGTATCCCCACCGGTCATGAAGAAGAA

GATGCAAAACAAGAACAAGATATCAAATGGAGTTTATAAAATCTTATCTCCTTTGCCCCTAA

AGGTAAACACGTGTAAGAGAATGATCACACTAGAGGAGACCGCATTAGCCGAGGATTTGGA

GACCCAGAGCAACTCCACGATGTTATCATCTTCAGACAATATCTATTTCGATGATCTAGCATT

ACTGTTGAGCAAAAGTTCAGCTTATCAGCAAGTTTTCCCTCAAGATGAGAAGGAGGCTGCCA

TTTTACTAATGGCTCTATCGCACGGAATGGTTCACGGG**TGA**ctaagctt

>GATA23 (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI, BglII & HindIII cloning sites underlined)

cc**ATG**GATCCAAGGAAGCTACTATCTTGTTCATCCTCTTACGTGTCAGTGAGAATGAAAGAA

GAGAAGGGGACAATTAGGTGTTGCAGTGAGTGTAAGACCACCAAGACACCAATGTGGAGAG

GTGGACCAACTGGTCCTAAGgtctcttcttctacctttaattactatattcataactttgtttgatcttaagataattcatcaagtgttcttaagtt
gtttattttgatttgtggtgggatttgcagTCACTTTGCAATGCATGTGGAATTAGACACAGAAAACAGAGACGA
TCAGAGTTATTGGGTATTCATATTATTCGCAGCCACAAAAGCTTAGCCTCCAAGAAGATAAA
CCTATTATCATCATCACACGGTGGCGTGGCGGTGAAGAAACGAAGGAGTCTAAAGGAGGAA
GAACAAGCTGCTTTGTGTCTATTGTTATTGTCTTGTAGCTCTGTTTTGGCC**TAA**gatcttaagctt

>GATA29 (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI & HindIII cloning sites underlined)

cc**ATG**GAGCCGGAACTGGACTTAACACTGAAGTTGGGTTTGCCAAATTCGACTGTTGAGACA
CATCTGACCTTAAGCCCTCCCACCACCACCACTGATCAGgtcttatttcatttatttattccttattattattatttttttcactat
aatttattgatatgagcatgaataaatatgaaaaatcttcatctgatgaatcaatcctagagagatctattgataccaccatctcttctttcatttattttgcttttgaa
ataattggactaattaatttcatgcttatcaaagtttttttaaattctttggactattattaaattattaaaaatctctcttaaaaaacgctaatcccactgtctttttaaag
gtaaaaacactaattggaggaaaaaaacattggaaccttatcatcatcttctttttacaattaattcctacatacggacaaatattacataaaatatgcattaatta
ctagtgtaacaaacaaatatagGGAACCAACGTCGTGGACGGCGGAGAAGTTATAAACCACCGTCGGGGT
CTTTTGGGAGATGATGAAGTGATCCACAACGAACCTACAAGAAACAACGTGGAGTTCAACA
TCAGAATCTACAATTACGTCTTTCAACAGTTTGTTGGTGCCCCCAACACCTTGAACTTTGCTC
CTTATCCGATGCCTCCGTCTCCGGCACCGGCACCTGAGACTCCTCCAGTGAGTGATGAGTAT
GTTCTGATTGATGTTCCTGCTAGGAGAGCTCGTCGTAACAACTCTACGGTGATGACAAACTC
TTGGAAGGAAAACGCAACCCCGAAGCGGATTCGCGGTTGTGGTGGATTCTGCGGGGGGAGG
ATTGAAGGGATGAAGAAGTGTACGAACATGAATTGTAACGCACTTAACACTCCTATGTGGC
GAAGAGGTCCTCTTGGTCCTAAGgtaattagtctccgaattttcaactaataaatcatgtttaaatattaattatgctgcttaataattattat
taatttttaaattggtaaatgtgtttcagAGTTTATGCAATGCTTGTGGGATCAAGTTCAGGAAGGAGGAAGAG
AGGAAAGCCAAAAGAAATGTAGTGATTGTATTAGATGAC**TAA**gctt

>HAN variant with mutant HAN domain (coding sequence in capitals, with start and stop codons in bold; NcoI, PflMI & HindIII cloning sites underlined)

cc**ATG**GCCCAGACTCCGTACACTACTTCAACGCAGGGGCAATATTGTCATTCTTGTGGAATGT

TCCACCACCATAGCCAAAGCTGCTGCTACAACAACAACAACAACTCCAACGCCGGTTCTTAC

TCGATGGTCTTCTCCATGCAAAACGGTGGCGTTTTCGAGCAGAACGGTGAGGACTATCATCA

CTCTTCCTCCCTCGTTGACGCTACTAGATCTGCTGGAACTCCTTCTACGAGGCTTTGTGAGGA

AGATGAGAAACGTAGACGCTCTACTTCATCTGGTGCTTCTTCTTGCATCTCCAACTTTTGGGA

CTTGATTCACACCAAAAACAACAACTCCAAAACGGCACCGTACAATAACGTTCCTTCTTTCT

CCGCTAACAAGCCAAGTCGCGGTTGTTCCGGTGGTGGTGGTGGCGGAGGAGGCGGTGGCGT

AGGTGACTCTCTTCTCGCTAGACGCTGTGCCAACTGTGACACTACTTCTACTCCACTATGGAG

GAATGGTCCTAGAGGCCCTAAGTCCCTATGCAACGCATGCGGCATTCGTTTCAAGAAGGAAG

AGAGAAGAACTACTGCGGCTTCAGGAAACACCGTCGTCGGAGCTGCACCGGTTCAAACCGA

CCAGTACGGGCATCACAACTCTGGCTACAATAATTACCATGCTGCCACTAATAACAACAATA

ATAATGGTACTCCGTGGGCTCATCACCACTCGACGCAGAGGGTTCCGTGTAATTATCCGGCA

AATGAGATCAGGTTCATGGATGATTACGGCAGTGGAGTAGCAAACAACGTTGAATCCGACG

GTGCTCACGGCGGTGTTCCGTTCCTTTCTTGGAGGCTTAATGTAGCGGATAGGGCAAGTCTT

GTCCATGACTTTACCAGATCT**TAA**gctt

>HAN variant with deleted HAN domain (coding sequence in capitals, with start and stop codons in bold; NcoI, PflMI & HindIII cloning sites underlined)

cc**ATG**GCCCAGACTCCGTACACTACTTCAACGCAGGGGCAATATTGTCATTCTTGTGGAATGT

TCCACCACCATAGCCAAAGCTGCTGCTACAACAACAACAACAACTCCAACGCCGGTTCTTAC

TCGATGGTCTTCTCCATGCAAAACGGTGGCGTTTTCGAGCAGAACGGTGAGGACTATCATCA

CTCTTCCTCCCTCGAATTCTGTGAGGAAGATGAGAAACGTAGACGCTCTACTTCATCTGGTG

CTTCTTCTTGCATCTCCAACTTTTGGGACTTGATTCACACCAAAAACAACAACTCCAAAACG

GCACCGTACAATAACGTTCCTTCTTTCTCCGCTAACAAGCCAAGTCGCGGTTGTTCCGGTGGT

GGTGGTGGCGGAGGAGGCGGTGGCGTAGGTGACTCTCTTCTCGCTAGACGCTGTGCCAACTG

TGACACTACTTCTACTCCACTATGGAGGAATGGTCCTAGAGGCCCTAAGTCCCTATGCAACG

CATGCGGCATTCGTTTCAAGAAGGAAGAGAGAAGAACTACTGCGGCTTCAGGAAACACCGT
CGTCGGAGCTGCACCGGTTCAAACCGACCAGTACGGGCATCACAACTCTGGCTACAATAATT
ACCATGCTGCCACTAATAACAACAATAATAATGGTACTCCGTGGGCTCATCACCACTCGACG
CAGAGGGTTCCGTGTAATTATCCGGCAAATGAGATCAGGTTCATGGATGATTACGGCAGTGG
AGTAGCAAACAACGTTGAATCCGACGGTGCTCACGGCGGTGTTCCGTTCCTTTCTTGGAGGC
TTAATGTAGCGGATAGGGCAAGTCTTGTCCATGACTTTACCAGATCT**TAA**gctt

>HAN variant with IAA17 EAR motif (coding sequence in capitals, with start and stop codons in bold; NcoI, PflMI & HindIII cloning sites underlined)

cc**ATG**GGCAGTGTCGAGCTGAATCTGAGGGAGACTGAGCTGTGTCTTGGTCTTCCCGGGGGA
GATACAGTGGCTCCGGTAACCGGAAAACGTAGACGCTCTACTTCATCTGGTGCTTCTTCTTG
CATCTCCAACTTTTGGGACTTGATTCACACCAAAAACAACAACTCCAAAACGGCACCGTACA
ATAACGTTCCTTCTTTCTCCGCTAACAAGCCAAGTCGCGGTTGTTCCGGTGGTGGTGGTGGCG
GAGGAGGCGGTGGCGTAGGTGACTCTCTTCTCGCTAGACGCTGTGCCAACTGTGACACTACT
TCTACTCCACTATGGAGGAATGGTCCTAGAGGCCCTAAGTCCCTATGCAACGCATGCGGCAT
TCGTTTCAAGAAGGAAGAGAGAAGAACTACTGCGGCTTCAGGAAACACCGTCGTCGGAGCT
GCACCGGTTCAAACCGACCAGTACGGGCATCACAACTCTGGCTACAATAATTACCATGCTGC
CACTAATAACAACAATAATAATGGTACTCCGTGGGCTCATCACCACTCGACGCAGAGGGTTC
CGTGTAATTATCCGGCAAATGAGATCAGGTTCATGGATGATTACGGCAGTGGAGTAGCAAA
CAACGTTGAATCCGACGGTGCTCACGGCGGTGTTCCGTTCCTTTCTTGGAGGCTTAATGTAGC
GGATAGGGCAAGTCTTGTCCATGACTTTACCAGATCT**TAA**gctt

>HAN variant with mutant IAA17 EAR motif (coding sequence in capitals, with start and stop codons in bold; NcoI, PflMI & HindIII cloning sites underlined)

cc**ATG**GGCAGTGTCGAGCTGAATCTGAGGGAGACTGAGGCATGTGCAGGTGCTCCCGGTGGA
GATACAGTGGCTCCGGTAACCGGTAAACGTAGACGCTCTACTTCATCTGGTGCTTCTTCTTGC

ATCTCCAACTTTTGGGACTTGATTCACACCAAAAACAACAACTCCAAAACGGCACCGTACAA

TAACGTTCCTTCTTTCTCCGCTAACAAGCCAAGTCGCGGTTGTTCCGGTGGTGGTGGTGGCGG

AGGAGGCGGTGGCGTAGGTGACTCTCTTCTCGCTAGACGCTGTGCCAACTGTGACACTACTT

CTACTCCACTATGGAGGAATGGTCCTAGAGGCCCTAAGTCCCTATGCAACGCATGCGGCATT

CGTTTCAAGAAGGAAGAGAGAAGAACTACTGCGGCTTCAGGAAACACCGTCGTCGGAGCTG

CACCGGTTCAAACCGACCAGTACGGGCATCACAACTCTGGCTACAATAATTACCATGCTGCC

ACTAATAACAACAATAATAATGGTACTCCGTGGGCTCATCACCACTCGACGCAGAGGGTTCC

GTGTAATTATCCGGCAAATGAGATCAGGTTCATGGATGATTACGGCAGTGGAGTAGCAAAC

AACGTTGAATCCGACGGTGCTCACGGCGGTGTTCCGTTCCTTTCTTGGAGGCTTAATGTAGC

GGATAGGGCAAGTCTTGTCCATGACTTTACCAGATCT**TAA**gctt

>GATA16 with HAN domain (coding sequence in capitals, with start and stop codons in bold; NcoI, MluI & HindIII cloning sites underlined)

cc**ATG**GtagatcacTCCTCCCTCGTcGACTGCACTCTCTCTCTTGGAACTCCTTCTACGAGGCTTTG

TGAGGAAGATTCTAGAacCATGAAAACAAGAGCTGAAGATATGATCGAACAGAACAACACT

AGCGTTAACGACAAGAAGAAGACTTGTGCTGATTGTGGAACCAGTAAAACTCCTCTTTGGCG

TGGTGGTCCTGTTGGTCCAAAGTCGTTGTGTAACGCGTGTGGGATCAGAAACAGAAAGAAG

AGAAGAGGAGGAACAGAAGATAATAAGAAATTAAAGAAATCGAGTTCTGGCGGCGGAAAC

CGTAAATTTGGTGAATCGTTAAAACAGAGTTTGATGGATTTGGGGATAAGGAAGAGATCAA

CGGTGGAGAAGCAACGACAGAAGCTTGGTGAAGAAGAACAAGCCGCTGTGTTACTCATGGC

TCTTTCTTATGGCTCTGTTTACGCT**TAA**agctt

>GATA16 with mutant HAN domain (coding sequence in capitals, with start and stop codons in bold; NcoI, MluI & HindIII cloning sites underlined)

cc**ATG**GTAGATCACTCCTCCCTCGTTGACGCTACTAGTTCTGCTGGAACTCCTTCTACGAGGC

TTTGTGAGGAAGATTCTAGAACCATGAAAACAAGAGCTGAAGATATGATCGAACAGAACAA

CACTAGCGTTAACGACAAGAAGAAGACTTGTGCTGATTGTGGAACCAGTAAAACTCCTCTTT
GGCGTGGTGGTCCTGTTGGTCCAAAGTCGTTGTGTAACGCGTGTGGGATCAGAAACAGAAA
GAAGAGAAGAGGAGGAACAGAAGATAATAAGAAATTAAAGAAATCGAGTTCTGGCGGCGG
AAACCGTAAATTTGGTGAATCGTTAAAACAGAGTTTGATGGATTTGGGGATAAGGAAGAGA
TCAACGGTGGAGAAGCAACGACAGAAGCTTGGTGAAGAAGAACAAGCCGCTGTGTTACTCA
TGGCTCTTTCTTATGGCTCTGTTTACGCT**TAA**agctt

>GATA16 with IAA17 EAR motif (coding sequence in capitals, with start and stop codons in bold; NcoI, MluI & HindIII cloning sites underlined)

cc**ATG**GGCAGTGTCGAGCTGAATCTGAGGGAGACTGAGCTGTGTCTTGGTCTTCCCGGGGGA
GATACAGTGGCTCCGGTAACCGGATCTAGAGATTCAGAAACCATGAAAACAAGAGCTGAAG
ATATGATCGAACAGAACAACACTAGcGTTAACGACAAGAAGAAGACTTGTGCTGATTGTGG
AACCAGTAAAACTCCTCTTTGGCGTGGTGGTCCTGTTGGTCCAAAGTCGTTGTGTAACGCGT
GTGGGATCAGAAACAGAAAGAAGAGAAGAGGAGGAACAGAAGATAATAAGAAATTAAAGA
AATCGAGTTCTGGCGGCGGAAACCGTAAATTTGGTGAATCGTTAAAACAGAGTTTGATGGAT
TTGGGGATAAGGAAGAGATCAACGGTGGAGAAGCAACGACAGAAGCTTGGTGAAGAAGAA
CAAGCCGCTGTGTTACTCATGGCTCTTTCTTATGGCTCTGTTTACGCT**TAA**agctt

>GATA16 with mutant IAA17 EAR motif (coding sequence in capitals, with start and stop codons in bold; NcoI, MluI & HindIII cloning sites underlined)

cc**ATG**GGCAGTGTCGAGCTGAATCTGAGGGAGACTGAGGCATGTGCAGGTGCTCCCGGTGGA
GATACAGTGGCTCCGGTAACCGGTTCTAGAGATTCAGAAACCATGAAAACAAGAGCTGAAG
ATATGATCGAACAGAACAACACTAGCGTTAACGACAAGAAGAAGACTTGTGCTGATTGTGG
AACCAGTAAAACTCCTCTTTGGCGTGGTGGTCCTGTTGGTCCAAAGTCGTTGTGTAACGCGT
GTGGGATCAGAAACAGAAAGAAGAGAAGAGGAGGAACAGAAGATAATAAGAAATTAAAGA
AATCGAGTTCTGGCGGCGGAAACCGTAAATTTGGTGAATCGTTAAAACAGAGTTTGATGGAT

TTGGGGATAAGGAAGAGATCAACGGTGGAGAAGCAACGACAGAAGCTTGGTGAAGAAGAA

CAAGCCGCTGTGTTACTCATGGCTCTTTCTTATGGCTCTGTTTACGCT**TAA**agctt

>GATA29 with mutant EAR consensus motif (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI, BglII & HindIII cloning sites underlined)

cc**ATG**GAGCCGGAgCTCGACGCTACAGCAAAGGCTGGTTTGCCAAATTCGACTGTTGAGACA

CATCTGACCTTAAGCCCTCCCACCACCACCACTGATCAGgtcttatttcatttatttattccttattattattatttttttcactat

aatttattgatatgagcatgaataaatatgaaaaatcttcatctgatgaatcaatcctagagagatctattgataccaccatctcttctttcatttattttgcttttgaa

ataattggactaattaatttcatgcttatcaaagtttttttaaattctttggactattattaaattattaaaaatctctcttaaaaaacgctaatcccactgtctttttaaag

gtaaaaacactaattggaggaaaaaaacattggaaccttatcatcatcttctttttacaattaattcctacatacggacaaatattacataaaatatgcattaatta

ctagtgtaacaaacaaatatagGGAACCAACGTCGTGGACGGCGGAGAAGTTATAAACCACCGTCGGGGT

CTTTTGGGAGATGATGAAGTGATCCACAACGAACCTACAAGAAACAACGTGGAGTTCAACA

TCAGAATCTACAATTACGTCTTTCAACAGTTTGTTGGTGCCCCCAACACCTTGAACTTTGCTC

CTTATCCGATGCCTCCGTCTCCGGCACCGGCACCTGAGACTCCTCCAGTGAGTGATGAGTAT

GTTCTGATTGATGTTCCTGCTAGGAGAGCTCGTCGTAACAACTCTACGGTGATGACAAACTC

TTGGAAGGAAAACGCAACCCCGAAGCGGATTCGCGGTTGTGGTGGATTCTGCGGGGGGAGG

ATTGAAGGGATGAAGAAGTGTACGAACATGAATTGTAACGCACTTAACACTCCTATGTGGC

GAAGAGGTCCTCTTGGTCCTAAGgtaattagtctccgaattttcaactaataaatcatgtttaaatattaattatgctgcttaataattattat

taatttttaaattggtaaatgtgtttcagAGTTTATGCAATGCTTGTGGGATCAAGTTCAGGAAGGAGGAAGAG

AGGAAAGCCAAAAGAAATGTAGTGATTGTATTAGATGAC**TAA**gctt

>GATA29 with HAN domain (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI, BglII & HindIII cloning sites underlined)

cc**ATG**GAGCCGGAATCCTCCCTCGTcGACTGCACTCTCTCTCTTGGAACTCCTTCTACGAGGCT

TTGTGAGGAAGATACACATCTGACCTTAAGCCCTCCCACCACCACCACTGATCAGgtcttatttcattt

atttattccttattattattatttttttcactataatttattgatatgagcatgaataaatatgaaaaatcttcatctgatgaatcaatcctagagagatctattgatacca

ccatctcttctttcatttattttgcttttgaaataattggactaattaatttcatgcttatcaaagttttttaaattctttggactattattaaattattaaaaatctctcttaa

aaaacgctaatcccactgtctttttaaaggtaaaaacactaattggaggaaaaaaacattggaaccttatcatcatcttctttttacaattaattcctacatacgg

acaaatattacataaaatatgcattaattactagtgtaacaaacaaatatagGGAACCAACGTCGTGGACGGCGGAGAAGTTA

TAAACCACCGTCGGGGTCTTTTGGGAGATGATGAAGTGATCCACAACGAACCTACAAGAAA

CAACGTGGAGTTCAACATCAGAATCTACAATTACGTCTTTCAACAGTTTGTTGGTGCCCCCA

ACACCTTGAACTTTGCTCCTTATCCGATGCCTCCGTCTCCGGCACCGGCACCTGAGACTCCTC

CAGTGAGTGATGAGTATGTTCTGATTGATGTTCCTGCTAGGAGAGCTCGTCGTAACAACTCT

ACGGTGATGACAAACTCTTGGAAGGAAAACGCAACCCCGAAGCGGATTCGCGGTTGTGGTG

GATTCTGCGGGGGGAGGATTGAAGGGATGAAGAAGTGTACGAACATGAATTGTAACGCACT

TAACACTCCTATGTGGCGAAGAGGTCCTCTTGGTCCTAAGgtaattagtctccgaattttcaactaataaatcatgtttta

aatattaattatgctgcttaataattattattaatttttaaattggtaaatgtgtttcagAGTTTATGCAATGCTTGTGGGATCAAGTTC

AGGAAGGAGGAAGAGAGGAAAGCCAAAAGAAATGTAGTGATTGTATTAGATGAC**TAA**gctt

>GATA29 with mutant HAN domain (coding sequence in capitals, with start and stop codons in bold; introns in small letters; NcoI, BglII & HindIII cloning sites underlined)

cc**ATG**GAGCCGGAAtCCTCCCTCGTTGACGCTACTAGTTCTGCTGGAACTCCTTCTACGAGGCT

TTGTGAGGAAGATACACATCTGACCTTAAGCCCTCCCACCACCACCACTGATCAGgtcttatttcattt

atttattccttattattattatttttcactataatttattgatatgagcatgaataaatatgaaaaatcttcatctgatgaatcaatcctagagagatctattgatacca

ccatctcttctttcatttattttgcttttgaaataattggactaattaatttcatgcttatcaaagttttttaaattctttggactattattaaattattaaaaatctctcttaa

aaaacgctaatcccactgtctttttaaaggtaaaaacactaattggaggaaaaaaacattggaaccttatcatcatcttctttttacaattaattcctacatacgg

acaaatattacataaaatatgcattaattactagtgtaacaaacaaatatagGGAACCAACGTCGTGGACGGCGGAGAAGTTA

TAAACCACCGTCGGGGTCTTTTGGGAGATGATGAAGTGATCCACAACGAACCTACAAGAAA

CAACGTGGAGTTCAACATCAGAATCTACAATTACGTCTTTCAACAGTTTGTTGGTGCCCCCA

ACACCTTGAACTTTGCTCCTTATCCGATGCCTCCGTCTCCGGCACCGGCACCTGAGACTCCTC

CAGTGAGTGATGAGTATGTTCTGATTGATGTTCCTGCTAGGAGAGCTCGTCGTAACAACTCT

ACGGTGATGACAAACTCTTGGAAGGAAAACGCAACCCCGAAGCGGATTCGCGGTTGTGGTG

GATTCTGCGGGGGGAGGATTGAAGGGATGAAGAAGTGTACGAACATGAATTGTAACGCACT

TAACACTCCTATGTGGCGAAGAGGTCCTCTTGGTCCTAAGgtaattagtctccgaattttcaactaataaatcatgttta

aatattaattatgctgcttaataattattattaatttttaaattggtaaatgtgtttcagAGTTTATGCAATGCTTGTGGGATCAAGTTC

AGGAAGGAGGAAGAGAGGAAAGCCAAAAGAAATGTAGTGATTGTATTAGATGAC**TAA**gctt

>N-terminal HAN fragment (attB1 & B2 recombination sites bold, SalI & XbaI cloning sites underlined, HAN-derived coding sequence in capital letters)

acaagtttgtacaaaaaagcaggctctttaaaggaaccaattcagtcgacaagagccATGGCCCAGACTCCGTACACTACTTC

AACGCAGGGGCAATATTGTCATTCTTGTGGAATGTTCCACCACCATAGCCAAAGCTGCTGCT

ACAACAACAACAACAACTCCAACGCCGGTTCTTACTCGATGGTCTTCTCCATGCAAAACGGT

GGCGTTTTCGAGCAGAACGGTGAGGACTATCATCACTCTTCCTCCCTCGTTGACTGCACTCTC

TCTCTTGGAACTCCTTCTACGAGGCTTTGTGAGGAAGATGAGAAACGTAGACGCTCTACTTC

ATCTGGTGCTTCTTCTTGCATCTCCAACTTTTGGGACTTGATTCACACCAAAAACAACAACTC

CAAAACGGCACCGTACAATAACGTTCCTTCTTTCTCCGCTAACAAGCCAAGTCGCGGTTGTT

CCGGTGGTGGTGGTGGCGGAGGTGACTCTTAGtctag**acccagctttcttgtacaaagtgg**

>N-terminal HAN fragment with mutant HAN domain (attB1 & B2 recombination sites bold, SalI & XbaI cloning sites underlined, HAN-derived coding sequence in capital letters)

acaagtttgtacaaaaaagcaggctctttaaaggaaccaattcagtcgacaagagccATGCCCAGACTCCGTACACTACTTCA

ACGCAGGGGCAATATTGTCATTCTTGTGGAATGTTCCACCACCATAGCCAAAGCTGCTGCTA

CAACAACAACAACAACTCCAACGCCGGTTCTTACTCGATGGTCTTCTCCATGCAAAACGGTG

GCGTTTTCGAGCAGAACGGTGAGGACTATCATCACTCTTCCTCCCTCGTTGACGCTACTAGAT

CTGCTGGAACTCCTTCTACGAGGCTTTGTGAGGAAGATGAGAAACGTAGACGCTCTACTTCA

TCTGGTGCTTCTTCTTGCATCTCCAACTTTTGGGACTTGATTCACACCAAAAACAACAACTCC

AAAACGGCACCGTACAATAACGTTCCTTCTTTCTCCGCTAACAAGCCAAGTCGCGGTTGTTC

CGGTGGTGGTGGTGGCGGAGGTGACTCTTAGtctag**acccagctttcttgtacaaagtgg**

>HAN motif (attB1 & B2 recombination sites bold, NcoI & XbaI cloning sites underlined, HAN-derived coding sequence in capital letters)

acaagtttgtacaaaaaagcaggctctttaaaggaaccaattcagtcgagaaccATGGTAGATCACTCCTCCCTCGTCGACT GCACTCTCTCTCTTGGAACTCCTTCTACGAGGCTTTGTGAGGAAGATtctagatagctag**acccagctttctt gtacaaagtgg**

>Mutant HAN motif (attB1 & B2 recombination sites bold, NcoI & XbaI cloning sites underlined, HAN-derived coding sequence in capital letters)

acaagtttgtacaaaaaagcaggctctttaaaggaaccaattcagtcgagaaccATGGTAGATCACTCCTCCCTCGTTGACG CTACTAGTTCTGCTGGAACTCCTTCTACGAGGCTTTGTGAGGAAGATtctagatagctag**acccagctttctt gtacaaagtgg**

>IAA17 EAR motif (attB1 & B2 recombination sites bold, NcoI & XbaI cloning sites underlined, IAA17-derived coding sequence in capital letters)

acaagtttgtacaaaaaagcaggctctttaaaggaaccaattcagtcgagaaccATGGGCAGTGTCGAGCTGAATCTGAGG GAGACTGAGCTGTGTCTTGGTCTTCCCGGGGGAGATACAGTGGCTCCGGTAACCGGAtctagata gctag**acccagctttcttgtacaaagtgg**

>Mutant IAA17 EAR motif (attB1 & B2 recombination sites bold, NcoI & XbaI cloning sites underlined, IAA17-derived coding sequence in capital letters)

acaagtttgtacaaaaaagcaggctctttaaaggaaccaattcagtcgagaaccATGGGCAGTGTCGAGCTGAATCTGAGG GAGACTGAGGCATGTGCAGGTGCTCCCGGTGGAGATACAGTGGCTCCGGTAACCGGTtctagata gctag**acccagctttcttgtacaaagtgg**