A stylized illustration of a plant with a large, dark green leaf and a smaller, lighter green leaf. The plant is set against a background of a colorful, geometric pattern of triangles in shades of red, orange, and yellow. The text is overlaid on the dark green leaf.

Mechanistic dissection of plant embryo initiation

Tatyana Radoeva

Mechanistic Dissection of Plant Embryo Initiation

Tatyana Radoeva

Thesis committee

Promotor

Prof. Dr D. Weijers
Personal Chair of the Laboratory of Biochemistry
Wageningen University

Co-promotor

Prof. Dr S.C. de Vries
Professor of Biochemistry
Wageningen University

Other Members

Prof. Dr R. Immink, Wageningen University
Prof. Dr R. E. Koes, VU University Amsterdam
Prof. Dr C. M. Liu, Chinese Academy of Sciences, Beijing, China
Dr M. Bayer, Max Planck Institute, Tuebingen, Germany

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences.

Mechanistic Dissection of Plant Embryo Initiation

Tatyana Radoeva

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University

by the authority of the Rector Magnificus

Prof. Dr A. P. J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Tuesday 21 June 2016

at 8.30 a.m. in the Aula.

Tatyana Radoeva
Mechanistic Dissection of Plant Embryo Initiation,
184 pages.

PhD thesis, Wageningen University, Wageningen, NL (2016)
With references and summaries in Dutch and English

ISBN 978-94-6257-813-5
DOI 10.18174/380679

Table of Contents

Chapter 1	7
Introduction	
Chapter 2	33
A novel ARF/bHLH module regulates extra-embryonic identity during Arabidopsis embryogenesis	
Chapter 3	79
A genome-wide screen for embryo inducers	
Chapter 4	115
Optimization of suspensor-specific activation tagging screen	
Chapter 5	131
Molecular characterization of Arabidopsis GAL4/UAS enhancer trap lines identifies novel cell-type-specific promoters	
Chapter 6	161
General Discussion	
English Summary	173
Dutch Summary	175
Acknowledgements	177
Curriculum Vitae	179
Publications	180
Education Statement	181

Chapter 1

Introduction

Parts of this Chapter were published as:

Radoeva T. and Weijers D. (2014). A roadmap to embryo identity in plants. *Trends Plant Sci* 19, 709-716

Although plant embryogenesis is usually studied in the context of seed development, there are many alternative roads to embryo initiation. These include somatic embryogenesis in tissue culture and microspore embryogenesis, both widely used in breeding and crop propagation, but also include other modes of ectopic embryo initiation such as extra-embryonic (suspensor) embryogenesis. In the past decades several genes, mostly transcription factors, were identified that can induce embryogenesis in somatic cells. Because the genetic networks in which such regulators operate to promote embryogenesis are largely unknown, a key question is how their activity relates to zygotic and alternative embryo initiation. In this Chapter, the many roads to plant embryo initiation are discussed including a genetic framework for defining the mechanisms of plant embryogenesis regulators. Further, experimental approaches taken in this thesis to dissect the genetic basis and mechanisms underlying embryo initiation are outlined using the *Arabidopsis* suspensor as a model system. Finally, the mode of action of the plant hormone auxin is discussed.

Plant embryogenesis

Seeds form the means of reproduction and dispersal in many higher plants. In addition, seeds are an important part of human diet and animal feed. They contain an embryo, protected by a seed coat that ruptures upon germination. As seed and embryo are intimately connected, embryogenesis is often considered in the context of seed development. However, in evolution, embryos are a more ancient innovation than seeds. This earlier origin of the embryo explains why the ability to initiate embryogenesis in plants appears to be a property shared by many cells other than the zygote. It also means that to understand the process of embryo initiation and the molecular triggers underlying it, we should extend our view beyond zygotic embryogenesis and include alternative modes of embryo development. In the past two decades, molecular and genetic studies have shed light on embryo initiation in a variety of contexts. In this Chapter, the different types of plant embryogenesis are explored along with molecular evidence for transcriptional control of embryo initiation and finally the regulatory framework for this important developmental event is discussed.

The many roads to embryo formation

Embryos were presumably one of the first significant innovations acquired by the land plants during evolution and the presence of an embryo is crucial for plant reproduction. The first embryo-producing plants (embryophytes), derived from a group of green algae, have diversified into thousands of extinct and modern species (Brooker, 2011). Even though today the embryo is found inside the seed, paleobotanical evidence suggests that seeds are later innovation (Figure 1) (DiMichele et al., 1989; Rothwell and Erwin, 1987). Thus, zygotic embryogenesis is not a derived trait, but a manifestation of an ancient property. Nonetheless, at present, seed plants dominate the plant kingdom, which suggests that seeds offer a reproductive advantage.

Regardless of the origin of the embryo, embryogenesis establishes the body plan of the plant. In addition to zygote-derived embryogenesis, several other modes of embryogenesis have been described. Embryo formation can be induced in somatic cells, microspores, female-reproductive cells and extra-embryonic cells. The zygotic embryo is here used as a model for describing progression and morphogenetic events during embryogenesis (Figure 2) and a brief discussion of various modes of alternative embryogenesis is provided (Figure 3).

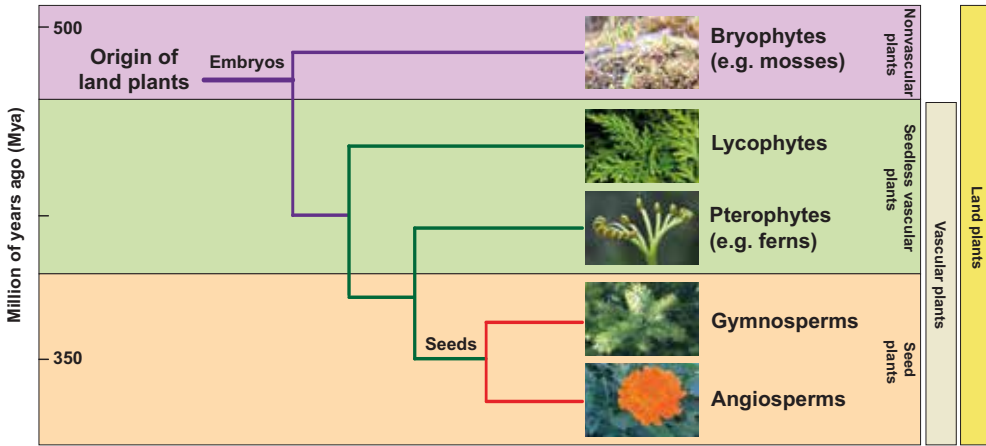


Figure 1. Evolutionary history of land plants.

Embryogenesis arose early in land plant evolution, while seeds evolved later. Representative images for each plant group are shown. Lycophyte (*Selaginella canaliculata*) image was reproduced from Wikimedia (http://commons.wikimedia.org/wiki/File:Selaginella_canaliculata.jpeg).

Zygotic embryogenesis

Zygotic embryogenesis is defined by its trigger, fertilization of the egg cell by a pollen-derived sperm cell. The early stages of embryo development are essentially the same in virtually all higher plants (Maheshwari, 1950; Raven et al., 1999; West and Harada, 1993). In many flowering plants (such as *Arabidopsis thaliana*), after fertilization, the highly polarized zygote elongates and divides asymmetrically, forming two different daughter cells: a smaller apical cell and a larger basal cell, which mark the apico-basal axis. The apical cell will undergo several rounds of coordinated cell divisions, changing its division plane and will give rise to most of the mature embryo (proembryo), whereas the basal cell will divide transversely, forming the extra-embryonic suspensor that provides mechanistic support and nutrients to the growing embryo (Mansfield and Briarty, 1991). It should be noted however that in a number of species, the polarity and asymmetry of zygote division deviates from this general pattern (Johri et al., 1992).

After three rounds of cell division (two longitudinal and one transverse) of the apical cell, the *Arabidopsis* embryo reaches the octant stage, where the upper and the lower tier of the proembryo are established (Mansfield and Briarty, 1991; Zhang and Laux, 2011) (Figure 2 and Figure 3A). Thereafter, all cells of the proembryo divide tangentially, resulting in the dermatogen stage. In the next stage, called globular, the uppermost suspensor cell is specified as hypophysis, which becomes part of the proembryo (Hamann et al., 1999). Development of the suspensor is

completed by this stage and it consists of a single file of six to nine cells. The following heart stage, where the proembryo is no longer spherical, is marked by the specification of the shoot apical meristem (SAM) cells and the cotyledons (Bard, 1994; Goldberg et al., 1994).

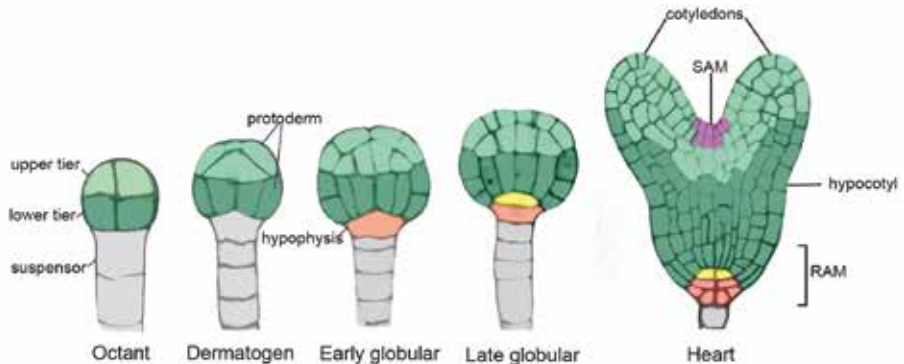


Figure 2. Arabidopsis zygotic embryogenesis.

Embryo development is shown from octant to heart stage. At octant stage, the upper and the lower tier of the proembryo are established, here represented in different colors. Later, when the embryo reaches globular stage, the uppermost suspensor (shown in grey) cell is specified as hypophysis (in orange). Next, the latter divides asymmetrically, forming apical lens-shaped cell (in yellow), the precursor of the QC (quiescent center) and basal cell (in orange), the progenitor of the columella stem cells. SAM, shoot apical meristem; RAM, root apical meristem.

During the progression of embryogenesis, the precursors of the major tissue types of the plant body are established (Lau et al., 2010; Laux and Jürgens, 1997; Mayer and Jürgens, 1998; Peris et al., 2010). Several recent reviews discuss the mechanisms underlying these processes, and their regulators in detail (Lau et al., 2010; Mayer and Jürgens, 1998; Moller and Weijers, 2009; Peris et al., 2010; De Smet et al., 2010; Wendrich and Weijers, 2013). Finally, the ovule containing the mature embryo shapes up to a seed, which remains dormant until beneficial conditions trigger germination and growth of the seedling (Bewley and Michael, 1985).

Somatic embryogenesis

Many somatic plant cells possess the capacity to induce an embryo and thereafter to regenerate a complete plant. This process, termed somatic embryogenesis, was first demonstrated and described in carrot (Reinert, 1959; Steward et al., 1958) and in the past decades in many other species. This system has been used for studying early regulatory events in plant zygotic embryogenesis (Zimmerman, 1993), but has also found wide application in plant breeding and *in vitro* propagation of crops (Brown and Thorpe, 1995).

Somatic embryogenesis can occur naturally, for example in *Kalanchoë* (mother of thousands) where somatic embryos are spontaneously formed on succulent leaves (Garces and Sinha, 2009) or *in vitro* after experimental induction, such as through treatment with hormones, notably auxin (de Vries et al., 1988; Reinert, 1959; Steward et al., 1958). Initiation can occur from various plant organs, and even from single cells such as leaf protoplasts (Luo Y. and U., 1997). It has been proposed that somatic embryo induction is a multistep process that involves prior installation of competence to respond to inductive signals (Dodeman et al., 1997; Zimmerman, 1993), but alternatively it may be directly induced in somatic cells. In either case, there needs to be a deviation from the normal developmental fate of the cell, that could be considered analogous to reprogramming somatic cells to induced pluripotent cells in animals (Takahashi, 2006) (Figure 3B).

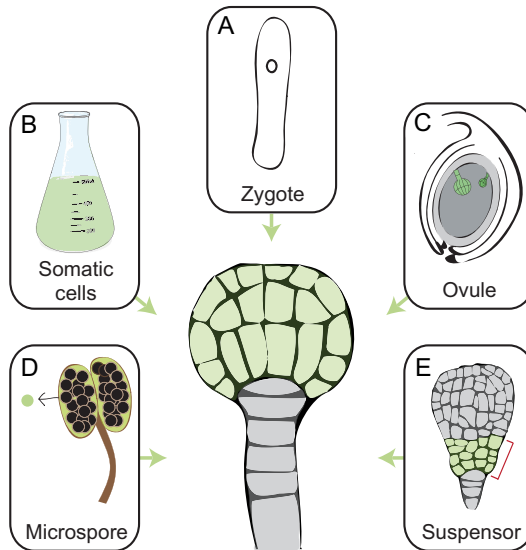


Figure 3. Different origins of plant embryos.

(A) Zygotic embryogenesis occurs after the fertilization of the ovule, giving rise to the zygote. The zygote then develops into an embryo (or a seed). (B) Somatic embryogenesis occurs naturally or after experimental induction, such as hormone treatment. Somatic embryos go through the same developmental stages as the zygotic embryos. (C) Apomictic embryogenesis, occurring in the seed primordium (ovule) is another road to embryo identity, known also as asexual reproduction. The generated embryos are genetically identical to the mother plant. (D) In microspore embryogenesis, microspores can also develop into embryos, when cultured under certain conditions. (E) Extra-embryonic (suspensor) embryogenesis occurs when suspensor cells switch identity to embryo cells.

Apomictic embryogenesis

The capability of somatic cells to initiate embryogenesis extends to cells in the seed primordium. The outcome is embryogenesis and subsequent seed development without fertilization, and is also known as asexual reproduction or apomixis. In addition to evading fertilization, apomixis often also skips meiosis and thus generates embryos that are genetically identical to the mother plant. Apomixis is thus a very interesting phenomenon, with great agronomic potential since it can fix hybrid genotypes (Spillane et al., 2004).

This asexual mode of reproduction is found throughout the plant kingdom (Asker and Lenn, 1992), and can be classified in two main types: sporophytic (adventitious embryony) or gametophytic (diplospory and apospory) (Barcaccia and Albertini, 2013; Bicknell and Koltunow, 2004; Koltunow et al., 1995) (Figure 3C). These different types of apomixis differ in the nature of the embryonic cells, and the step at which induction occurs. Following induction however, progression is highly similar to zygotic embryogenesis (Naumova, 1993). Loci associated with apomixis have been identified in several species (Catanach et al., 2006; Hojsgaard et al., 2011) but no causal genes have yet been identified, and thus the underlying triggers are unknown.

Microspore embryogenesis

Another efficient way of inducing embryo formation is from microspores (precursors of the pollen grains; Figure 3D). Microspore embryogenesis (androgenesis) is a process, where microspores or pollen grains develop into haploid or doubled haploid embryos when cultured under certain conditions (Touraev et al., 1997). This switch from microspores toward embryogenesis is widely used in plant propagation and plant breeding.

Upon applying an inductive signal such as for example a heat shock, some microspores deviate from the normal, gamete-producing pathway, and enter embryogenesis. This shift is accompanied by morphological changes as well as biochemical changes (Mordhorst et al., 1997; Segui-Simarro and Nuez, 2008). Although cases have been reported in which *Brassica* microspores follow a zygotic-like pattern of division from the very beginning (Supena et al., 2008; Zaki and Dickinson, 1991), microspore embryogenesis often involves several initial rounds of randomly oriented cell divisions inside the exine (the outer coat of a pollen grain or a spore). In such cases it is not clear exactly when the switch from microspore to embryo identity happens; therefore the initial stages are referred to as sporophytic growth. Later, a well-defined protoderm can be recognized and is considered to be a marker for

embryo formation. At this stage, compact structures with protoderm are interpreted as embryo-like structure, which will give rise to embryos with all major tissues found in zygotic embryos (reviewed in Soriano et al., 2013).

It is unclear whether the capacity to undergo microspore embryogenesis generates reproductive advantage, or whether analogous processes occur in nature. Alternatively, given the derivation of embryogenesis from spore-like evolutionary precursors (Taylor et al., 2009), microspore embryogenesis may be a remnant of a basic developmental potential. Although the precise mechanisms for induction are not known, it was recently shown that chromatin regulation contributes to its control (Li et al., 2014).

Extra-embryonic (suspensor) embryogenesis

In the above, embryogenesis in various cell types is described, more or less remote from the fertilized egg cell. In all but a few cases, these are experimentally induced types (e.g. somatic and microspore embryogenesis), or genetically derived states (e.g. *Kalanchoe* leaf margin embryos and apomixis). Another, less-studied mode of embryogenesis is unique in the sense that it occurs within the confines of the fertilization product, and may represent an ancient “backup” system. In most plant species, the embryo is subtended by a suspensor, a file of extra-embryonic cells (Maheshwari, 1950). While in some species the suspensor is limited to a few cells, in others it can contain thousands of cells (Yeung and Clutter, 1979; Yeung and Meinke, 1993). In a number of species, such as *Nicotiana rustica*, *Capsella* and *Arabidopsis*, these extra-embryonic cells have been shown to be capable of converting to embryo identity (Yeung and Meinke, 1993).

In *Arabidopsis*, the suspensor is determined after the first asymmetric division of the zygote and it is completed by the globular stage of the embryo development (Figure 2), hence it is the first terminally differentiated structure produced during embryogenesis (Mansfield and Briarty, 1991). The general view is that the suspensor provides nutrients and growth regulators during early stages of development of the embryo proper (Yeung and Meinke, 1993) as well as to fix the embryo in the micropylar cavity within the seed.

In several species, suspensor cells develop secondary embryos to form poly-embryonic progeny (Lakshmanan and Ambegaokar, 1984). In classical experiments, Haccius demonstrated that the suspensor-derived embryony could be induced by treatments that kill or incapacitate the original embryo (Haccius, 1955). This suggests that suspensor cells have an innate capacity to form embryos, and that this activity is suppressed by the presence of a live and functional embryo. Indeed, *Arabidopsis*

mutations that impair embryo viability, such as *sus* and *twn* (Schwartz et al., 1994; Vernon and Meinke, 1994) likewise induce suspensor-derived embryogenesis, which could also be induced by expressing a toxin in pro-embryo cells (Weijers et al., 2003). Thus, suspensor cells are a reservoir of cells that will develop embryos upon loss of inhibiting influence by the pro-embryo, and can be considered a population of “latent stem cells”. The mechanism of embryo suppression and the nature of an embryo-derived signal have long been mysterious. It was recently shown that the plant hormone auxin acts cell-autonomously in maintaining suspensor cell identity. Upon inhibition of auxin response, suspensor cells lose their identity and convert to embryonic cell fate, resulting in the expression of embryo-specific genes and leading to the formation of conjoined twin seedlings (Rademacher et al., 2012). This event represents an alternative, yet predictable path to embryogenesis from a defined population of cells (Figure 3E).

Transcriptional control of embryo identity

Genetic studies in the past decades have identified genes and loci required for initiation and progression of embryogenesis (e.g. Tzafrir et al., 2004) or apomixis (Catanach et al., 2006; Vijverberg et al., 2010). Nevertheless, whether such genes act specifically in embryo induction or are required for general aspects of plant cell viability remains to be shown. In parallel, molecular approaches have identified genes that are activated during the various embryo initiation processes. Several of these were later shown to be able to trigger embryogenesis. Here the most notable factors are discussed (listed in Table 1), mostly encoding transcription factors.

LEAFY COTYLEDON (*LEC*) was first identified based on its mutant phenotype. *lec* mutant cotyledons have leaf-like characteristics, which suggests that *LEC1* is involved in retaining juvenility (Meinke, 1992; Meinke et al., 1994). Later, it was shown that *LEC1* and other transcription factors that are functionally related, such as *LEC2* and *FUS3* (*FUSCA3*), can induce embryo development (*LEC2*) or late embryo properties (*FUS3*) in somatic cells (Braybrook and Harada, 2008; Gaj et al., 2005; Lotan et al., 1998). *LEC1* encodes the HAP3 subunit of the CCAAT box-binding factor, while both *LEC2* and *FUS3* are B3 domain transcription factors (Luerssen et al., 1998; Meinke, 1992; Meinke et al., 1994; West et al., 1994).

Members of the AP2 transcription factor family, including BABY BOOM (*BBM*/*AIL2*/*PLT4*), play an essential role in cell proliferation and embryogenesis. *BBM* is preferentially expressed in developing embryos and seeds and when overexpressed it induces somatic embryos from leaf and cotyledon margins and from the shoot

apical meristem (Boutillier et al., 2002). A close homolog of BBM, *AINTEGUMENTA-like 5 (AIL5)*, also known as *EMBRYOMAKER (EMK)* or *PLETHORA5 (PLT5)*, is expressed in developing and mature embryos, but not during germination. Just like BBM, *AIL5* is able to induce embryo formation when ectopically overexpressed (Tsuwamoto et al., 2010). In contrast, *PLT1/AIL3* and *PLT2/AIL4*, other closely related members of the AP2 family of transcription factors, whose expression can be found in the embryo as well as post-embryonically, could ectopically induce root formation (Aida et al., 2004; Galinha et al., 2007).

Although all genes discussed above are expressed in early embryos or their precursors, embryogenesis can also be induced by factors that are not normally expressed there. The homeobox gene *WUSCHEL (WUS)* is expressed in the organizing center of the shoot meristem and is required to keep adjacent stem cells in an undifferentiated state (Mayer and Jürgens, 1998; Schoof et al., 2000). In plants overexpressing *WUS*, formation of ectopic stem cells but also of somatic embryos is induced, indicating that stem cell and embryo programs may share overlaps (Zuo et al., 2002).

Although some regulators can directly induce (aspects of) embryo identity in somatic tissue, other factors act in a more indirect manner. The latter would increase the capacity to induce somatic embryos in response to other triggers rather than induce embryogenesis directly. The MADS domain transcription factor *AGL15 (AGAMOUS-like 15)* is preferentially expressed in developing embryos (Heck et al., 1995), and ectopic expression enhanced somatic embryo initiation from the shoot meristem in liquid seedling cultures (Harding et al., 2003). Likewise, RKD (RWP-RK DOMAIN-CONTAINING) transcription factors are expressed in the egg cell (*RKD1* and *RKD2*; Koszegi et al., 2011) or early embryo (*RKD4*; (Waki et al., 2011)), and when misexpressed induce callus formation (*RKD1*, *RKD2*; Koszegi et al., 2011) or cell populations that converted to embryos upon RKD removal (*RKD4*; Waki et al., 2011). While most embryogenesis inducers identified so far encode transcription factors, other components have been identified, such as the *SERK1 (SOMATIC EMBRYOGENESIS RECEPTOR KINASE1)* gene, which encodes a leucine-rich repeat (LRR) transmembrane receptor-like kinase (RLK) (Hecht et al., 2001). *SERK1* is expressed in developing embryos, as well as in cells acquiring embryonic competence during somatic embryogenesis (Kwaaitaal and de Vries, 2007; Schmidt et al., 1997). *SERK1* overexpression was shown to increase somatic embryogenesis in culture (Hecht et al., 2001).

Table 1. Genes that can induce ectopic embryogenesis.

AGIa	Gene name	Gene family	Expression pattern	Function	References
AT1G21970	LEC1	CCAAT-box binding	Early embryo	Embryo development; maturation	(Lotan et al., 1998)
AT1G28300	LEC2	B3 domain	Embryo and seedling	Embryo development; auxin biosynthesis	(Stone et al., 2001)
AT3G26790	FUS3	B3 domain	Seed maturation stage	Embryogenesis, positive regulation of abscisic acid biosynthesis	(Braybrook and Harada, 2008; Luerssen H et al., 1998)
AT5G17430	AIL2/BBM/PLT4	AP2/EREBP	Globular-late embryo	Cell proliferation, embryogenesis	(Boutillier et al., 2002)
AT5G57390	AIL5/EMK/PLT5	AP2/EREBP	Developing and mature embryos	Germination, cell proliferation, seed maturation	(Tsuwamoto et al., 2010)
AT3G20840	PLT1	AP2/EREBP	Basal embryo region, root stem cell niche	Root formation/development, stem cell maintenance	(Aida et al., 2004)
AT1G51190	PLT2	AP2/EREBP	Basal embryo region, root stem cell niche	Root formation/development, stem cell maintenance, embryogenesis	(Aida et al., 2004)
AT5G13790	AGL15	MADS-box	Embryo	Embryo development	(Harding et al., 2003; Heck et al., 1995)
AT2G17950	WUS	Homeodomain	Shoot stem cell niche	Stem cell maintenance	(Laux T et al., 1996; Zuo et al., 2002)
AT1G18790	RKD1	RWP-RK	Egg cell and egg cell apparatus	Egg cell specification and differentiation	(Koszegi et al., 2011)
AT1G74480	RKD2	RWP-RK	Egg cell and egg cell apparatus	Egg cell specification and differentiation	(Koszegi et al., 2011)
AT5G53040	RKD4	RWP-RK	Early embryos	Embryo development	(Waki et al., 2011)
AT1G71830	SERK1	LRR-RLK	Ovule, embryo until heart stage, vascular tissue	Embryo development, co-receptor	(Hecht et al., 2001)
AT2G25170	PKL (<i>pk1</i> mutant)	CHD3	Ubiquitous	Embryogenesis, cell proliferation, chromatin modification	(Ogas et al., 1999)

a. *Arabidopsis thaliana* gene index

Embryogenesis is not only subject to positive regulation but is also actively repressed, and several factors mediating this repression have been isolated. *PICKLE* (*PKL*) encodes a Chromatin-Helicase-DNA binding 3/4 (CHD3/4) like- chromatin remodeling factor and was shown to repress embryonic identity during germination (Ogas et al., 1999). Likewise, mutations in multiprotein chromatin regulatory complexes Polycomb Repressive Complex 1 and 2 (PRC1, 2), as well as the B3 transcription factors HIGH-LEVEL EXPRESSION OF SUGAR-INDUCIBLE GENE1 (*VAL1/HSI2*) and *VAL2/HSL1* cause retention of embryonic identity in seedlings (Bouyer et al., 2011; Bratzel et al., 2010; Tsukagoshi et al., 2007; Yang et al., 2013). The likely function of these factors is to allow phase transition from embryonic to post-embryonic by suppressing the embryonic state.

In summary, embryo identity is subject to both positive and negative regulation, and a host of factors has been identified that regulate either aspect. A key question is whether these factors constitute a coherent framework.

Converging paths or back to basic?

The identification of a range of genes that are able to trigger embryo initiation (Table 1) urges a fundamental question to be addressed. Are these factors all part of a large gene regulatory network that converges on a common set of “embryo” genes, or does the embryo initiation response reflect regression to a cellular “default” state? At present, it is impossible to distinguish definitely between these two possibilities, but each model makes several predictions. If embryo-induction genes are part of a larger network, one would expect that there is some degree of functional interdependence. Unless all act in parallel to control the same set of genes, there should be epistatic interactions and/or protein-protein interactions among the embryo induction factors.

Another important consideration when asking whether embryo-inducing genes are part of a coherent network concerns what precise cellular processes are entailed by conversion to embryo identity. If somatic-to-embryo conversion is not a single, transdifferentiation step (Fig. 4A), embryo-inducing genes may not be expected to act in a simple network. Conceptually, this transformation may additionally involve de-differentiation and competence installment (Fig. 4B-D). From misexpression studies using *RKD4* (Waki et al., 2011), as well as from experimentally induced somatic embryogenesis (de Vries et al., 1988; Zimmerman, 1993), it seems likely that additional steps are involved. In addition, because ubiquitous expression of embryo inducers does not convert all cells to embryos (Boutillier et al., 2002; Lotan

et al., 1998), competence may normally be limited to specific regions, and dedicated gene activity may be necessary to promote competence.

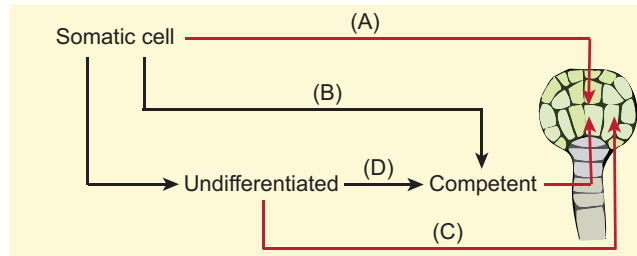


Figure 4. Concepts in embryo initiation.

Somatic cells can be reprogrammed to embryo identity in one of four ways: (A) direct transdifferentiation; (B) installation of competence followed by embryo identity specification; (C) de-differentiation, followed by embryo identity specification; (D) de-differentiation followed by installation of competence and embryo identity specification.

Because most embryo inducer genes (Table 1) encode transcription factors, a rigorous test of the network hypothesis would be the analysis of target genes, as these should reveal whether there is direct regulation of one by the other. A second testable prediction is that the transformation from somatic to embryonic as induced by each of these factors should entail the regulation of a common set of genes before the final establishment of embryo identity.

Several transcriptomics datasets on misexpression lines and mutants of embryo regulators are available (e.g. Passarinho et al., 2008; Stone et al., 2001; Waki et al., 2011; Zheng et al., 2009), and for a few there are also genome-wide protein-DNA binding profiles (Wang and Perry, 2013). It should be noted that the tissues used for transcript profiling are different in all cases because datasets were not generated with the purpose to compare different regulators. Several excellent reviews are available that describe the interactions between key regulators of embryogenesis and the transition from somatic to embryonic cell state (Braybrook and Harada, 2008; Braybrook et al., 2006; Horstman et al., 2014; Ikeuchi et al., 2013; Meinke, 1992; Wang and Perry, 2013). Here, the connections that are relevant to embryo initiation are reviewed.

From comparing transcriptional targets of LEC2, AGL15, FUS3, PKL and BBM, it appears that there are several interconnections (Figure 5), but perhaps not an overwhelming number. LEC2, AGL15 and FUS3 all bind and activate the *Aux/IAA30* gene (Braybrook et al., 2006; Wang and Perry, 2013; Zheng et al., 2009), a negative regulator of auxin response (Sato and Yamamoto, 2008). This finding is consistent with the role of auxin in suppressing embryogenesis in suspensor cells (Rademacher

et al., 2012). Furthermore, both LEC2 and FUS3 seem to activate AGL15 (Wang and Perry, 2013; Wang et al., 2004), thus constituting a feed-forward loop. However, a simple interpretation in which auxin activity is suppressed is at odds with the finding that LEC2 also activates the YUC2 and YUC4 biosynthesis genes (Stone et al., 2008).

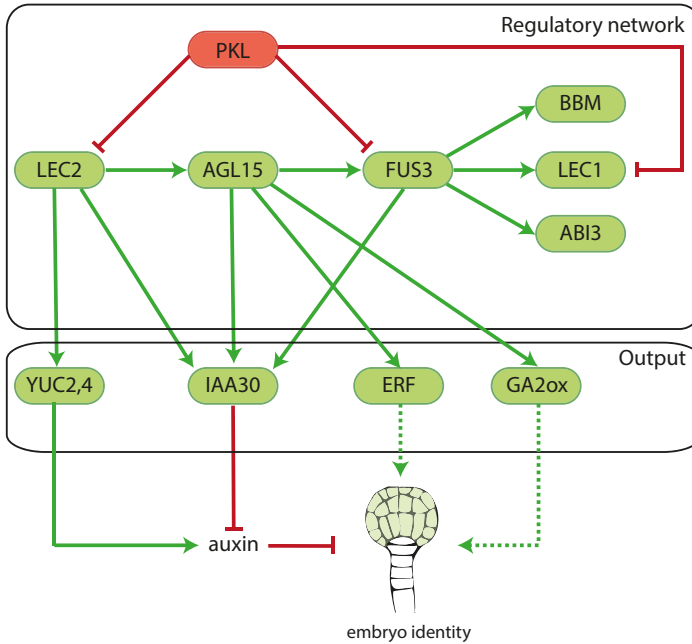


Figure 5: Regulatory interactions in the control of embryo initiation.

Interactions (protein-promoter or transcriptional regulation) are shown between regulators of embryo initiation. Green arrows indicate transcriptional activation, whereas red lines indicate transcriptional repression. Genes in green boxes are positive regulators of embryo initiation while those in red control embryogenesis negatively.

Nonetheless, there is some convergence in the transcriptional network surrounding embryo inducers because also LEC1, LIL1 (LEC1-LIKE), ABI3, BBM, and VAL1 and 2 are among the putative FUS3 targets (Wang and Perry, 2013). Finally, expression of LEC1, LEC2 and FUS3 depends on PKL activity (Dean Rider et al., 2003). In stark contrast, transcript profiling in BBM-overexpression lines did not recover any of the other components known to induce embryogenesis (Passarinho et al., 2008). Rather surprisingly, there is very little overlap in the transcriptional output of the different regulators. It should be noted that chromatin immunoprecipitation (ChIP) experiments usually identify many binding sites (hundreds to thousands) for transcription factors, and transcript profiling in mutants or overexpression lines can also recover indirect effects. Thus, although several potential convergence

points between embryo regulators are suggested from genome-wide approaches, these are not necessarily statistically significant and will need to be rigorously tested to establish relevance. Importantly, the functions of most of these regulators are not limited to embryo induction, and regulatory interactions may not occur in the context of embryo induction. A way forward would be to systematically examine the effect of embryo inducers on transcriptomes and embryo initiation in a single developmental system. Given its predictability, the recently developed suspensor model (Rademacher et al., 2012) could offer an attractive model system for such studies.

An alternative hypothesis to that of a core embryo regulatory network is that, when any of these genes is misexpressed, cellular homeostasis is disturbed in such a way that cells revert to a default state. This interpretation could rationalize why structurally diverse proteins all have the ability to induce embryogenesis, and is also in agreement with the omnipresence of competence for embryo induction. A prediction following from this interpretation is that the embryo inducers need not have functional connections. Another is that many more factors may be able to induce embryogenesis, given that disturbing homeostasis is likely a more widespread activity than triggering a specific genetic program. Again, analysis of embryo-inducing potential in a single developmental system would help to address this question. A future direction could be to systematically test the ability of genes for embryo-inducing activity in the same developmental context in a genome-wide fashion. Such directions are important to explore because they would have significant implications for the transferability of embryo induction to crop as part of for example improvement strategies for propagation of hybrids.

It is expected that knowledge derived from the activity of developmental switches will give insight into the underlying competence of plant cells to form embryos. In addition, this should help to define how, during evolution, this capacity has been restricted to the zygote, and how it has been adopted in alternative programs such as apomixis. Finally, such knowledge should not only help in understanding the fundamental principles of plant embryo induction, but would also provide an ideal starting point for engineering embryogenesis for crop propagation.

The suspensor as an experimental model for embryo initiation

As described above, embryogenesis can be induced in a variety of cell types, and by a number of different triggers. A key unanswered question however remains what genetic programs (reprogramming) underlies these cell fate changes, and also if they

are conserved. In this thesis, the *Arabidopsis* suspensor is used as a model to study the transformation of non-embryonic cells towards embryonic identity. The suspensor in *Arabidopsis* consists of only a few cells, yet these can initiate a second embryo. Here two complementary approaches are taken to exploit the simplicity and predictability of this system to dissect the mechanism underlying embryo initiation. The first builds on a previous observation that auxin response inhibition in suspensor cells triggers embryo transformation (Rademacher et al., 2012). The second takes an unbiased approach towards identifying genes that can convert suspensor cells to embryo identity. Each of these approaches will be introduced in the following.

Auxin signaling in the suspensor

Inhibition of auxin response causes loss of suspensor and gain of embryo identity (Rademacher et al., 2012), suggesting that this hormone suppresses embryo identity. Auxin has been studied for many decades and its action is understood in detail. Here, auxin action will be briefly summarized with emphasis on embryo development. For more details, we refer to more recent reviews on auxin action (Dinesh et al., 2015; Wang and Estelle, 2014) or auxin function in embryogenesis (Smit and Weijers, 2015).

The plant hormone auxin is a chemically simple molecule, but it has been recognized as an essential regulator of most of the patterning steps during *Arabidopsis* embryogenesis. Auxin action involves biosynthesis (Figure 6A), transport (Figure 6B) and response (Figure 6C). Auxin is an indole-derived phytohormone synthesized via two major pathways: a Tryptophan (Trp)-dependent and a Trp-independent pathway (Tao et al., 2008). The Trp-dependent pathway is better understood and it branches in multiple routes. TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS (TAA) and YUCCA (YUC) are the biosynthetic enzymes that function sequentially in one of these routes (Stepanova et al., 2008; Tao et al., 2008). Mutation in either of these leads to similar phenotypes, inferring that both are required to supply enough auxin for correct pattern formation (Moller and Weijers, 2009).

Auxin is a mobile signaling molecule and can be transported in two ways: passively through vascular tissue and actively via transporter proteins, known as influx and efflux carriers. AUX/LIKE AUX (AUX/LAX) proteins function as influx carriers and pump auxin into the cells, whereas the PIN-FORMED proteins (PINs) along with the ABCB/PGP proteins function as efflux carriers and export auxin from the cell (Petrášek and Friml, 2009). These transporters establish auxin gradients or local maxima required to instruct cell division, expansion and differentiation (Friml, 2010).

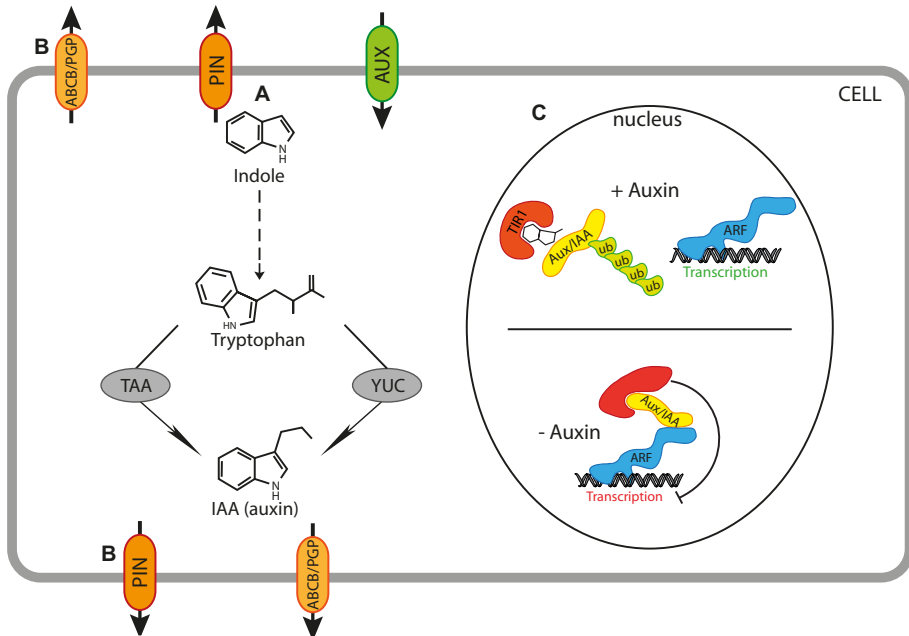


Figure 6: Simplified overview of auxin action.

(A) Tryptophan (Trp)-dependent biosynthesis of auxin branches in TAA and YUC-dependent routes. (B) Active auxin transport in and out of the cell is executed by influx (AUXs) and efflux (PINs; ABCB/PGPs) carrier proteins. (C) Auxin response is concentration dependent: upon low auxin concentration, ARFs are inhibited by interacting Aux/IAA proteins and released from inhibition upon high auxin concentration.

Auxin initiates signaling by binding to its nuclear receptor TRANSPORT INHIBITOR RESISTANT 1/AUXIN F-BOX (TIR1/AFB), a subunit of the SKP1-CULLIN1-F-BOX (SCF^{TIR1/AFB}) ubiquitin ligase complex. This binding enhances the affinity of SCF^{TIR1/AFB} for its substrate, the Auxin/Indole-3-Acetic Acid proteins (Aux/IAs), a family of transcription inhibitors (Wang and Estelle, 2014). Thus, auxin is commonly referred to as “molecular glue” that brings the Aux/IAs and the auxin receptor together. Once bound to the receptor complex, Aux/IAA proteins are ubiquitinated and degraded (Tan et al., 2007). At low auxin concentrations, Aux/IAs bind to and inhibit the activity of auxin response transcriptional regulators, the AUXIN RESPONSE FACTORS (ARFs), DNA-binding transcription factors that regulate the expression of auxin-responsive genes. When auxin is present at high concentrations in the cell, the Aux/IAs are degraded and the ARFs are free to perform their function as transcriptional activators or repressors (Ulmasov et al., 1999). Although the auxin response pathway is quite short and simple, it controls myriad cellular processes ranging from cell division and cell expansion to changes in developmental cell fate (Dubrovsky et al., 2008; Hardtke and Berleth, 1998; Rademacher et al., 2012; Weijers et al., 2006). The specificity of this rather simple response is obtained

through selection of target genes by the ARF and Aux/IAA proteins. In *Arabidopsis*, there are 23 ARFs and 29 Aux/IAA proteins (Remington et al., 2004). A cellular expression map of the ARFs reveals that every cell type has its own subset of expressed ARFs (Rademacher et al., 2011). Although the role of ARFs and their target genes in some cell types is well-known, for instance the role of ARF5/MP in embryonic root initiation and formation (Berleth and Jürgens, 1993; Schlereth et al., 2010), vascular tissue establishment (De Rybel et al., 2013), flower development (Yamaguchi et al., 2013), lateral root development (De Rybel et al., 2010), their role in maintaining suspensor identity is still novel. The suspensor expresses several ARF and Aux/IAA proteins (Rademacher et al., 2011, 2012). Importantly, suspensor to embryo transformation can be induced either by ectopic expression of a non-degradable mutant version of the IAA12/BODENLOS (*bdl*) protein, or by stabilization of the IAA10 protein that is normally expressed in suspensor cells (Rademacher et al., 2012). Thus, auxin acts to suppress embryo identity through genetic regulation. Therefore, we can expect to learn more about the molecular mechanisms of this activity by identifying the genes regulated during this process.

Activation tagging in *Arabidopsis*

Previous approaches, discussed in earlier sections of this chapter, have used misexpression of known genes, suspected to regulate embryo development, to determine if these genes can induce embryogenesis. All of these relied on prior isolation of the gene, and suspicion of a role in embryogenesis. Arguably however, there will likely be many more regulators of embryo initiation, whose function is not currently suspected. Ideally, we would like to systematically test the ability of every gene to induce embryogenesis. One method towards this goal is activation tagging. This strategy has been used for over a decade to identify novel genes that are not readily identified by loss-of-function screening, and have aided the dissection of genetic pathways. While the conventional loss-of-function mutagenesis have some limitations in the identification of the function of redundant genes or genes that are absolutely required in early embryo or gametophyte development, activation tagging has advantages that can circumvent these limitations (Weigel et al., 2000).

Activation tagging systems use either a T-DNA or a transposable element carrying tandem copies of a transcriptional enhancer sequence – e.g. from the constitutively active cauliflower mosaic virus (CaMV) 35S promoter - that are randomly inserted into the host genome (Marsch-Martinez et al., 2002; Suzuki et al., 2001; Weigel et al., 2000). These enhancer sequences can trigger transcriptional activation of

adjacent genes and subsequently lead to a dominant gain-of-function phenotype. In *Arabidopsis*, several genes have been successfully identified using this strategy, for instance *FLOWERING LOCUS T* (FT; Kardailsky et al., 1999), the patanin-like gene *STURDY* (Huang et al., 2001), auxin biosynthesis gene *YUCCA* genes (Zhao et al., 2001) and two *GA-2 oxidase* family genes (Hedden and Phillips, 2000). In addition, activation tagging has been used as an approach to isolate suppressor mutants of known mutant phenotypes (Neff et al., 1999; Weigel et al., 2000). However, integration of 35S enhancers in the host genome changes the level of gene expression, but may not alter its spatial pattern (Weigel et al., 2000). Thus phenotypes caused by altering spatiotemporal expression patterns will therefore not be identified.

Recently, an alternative, tissue-specific activation tagging screen was reported. This screen used the two-component GAL4/UAS system and allows tissue-specific activation of tagged genes (Waki et al., 2013). Plants that express a synthetic transcriptional activator GAL4-VP16 in a tissue-specific manner are used as host to randomly insert T-DNA carrying 5 tandem copies of GAL4-dependent upstream activation sequence (UAS) (Brand and Perrimon, 1993; Waki et al., 2013). Insertion of this T-DNA next to a gene may cause its activation leading to dominant phenotype. One can customize this tagging system by selecting a GAL4 enhancer trap line with a desired pattern of expression from a large collection (Haseloff, 1999). Thus, this approach should allow to systematically screen for genes that are capable of inducing embryogenesis in non-embryogenic cells.

Scope of the thesis

Embryogenesis in both plants and animals begins with a single cell, the zygote and ends with the formation of a mature embryo. However in plants, embryo formation can be initiated from a variety of cells and does not strictly depend on the zygote. In this thesis, emphasis is put on the initial step of embryo formation and its genetic regulation. On one hand, we use a uniform system, the suspensor-derived embryogenesis, to test the ability of already known “embryo inducers” to initiate embryo formation. On the other hand, we employ a genome-wide activation tagging screen to identify novel embryogenic regulators. The results obtained in this study provide one of the first systematic insights into the molecular bases for embryo initiation in plants.

In **Chapter 2**, we describe the results of genome-wide transcriptional profiling upon suspensor-specific auxin response inhibition that was performed to identify

genes regulated during suspensor-derived embryogenesis. We identified an auxin-dependent bHLH transcription factor that likely plays a role in suppressing embryo identity in the suspensor.

In **Chapter 3**, instead of characterizing embryo initiation as induced by a single regulator, we explore to what degree known embryogenesis inducers can trigger embryogenesis in suspensor cells. We find that only a specific set of regulators can do so. Based on this finding, we establish a genome-wide activation tagging screen towards identification of more genes that are able to induce suspensor-derived twin embryos.

Based on our findings, we next re-designed and optimized suspensor-specific activation tagging, the results of which are described in **Chapter 4**. In addition to setting up this new generation screen, we provide the results of a pilot “proof of concept” screen identifying a fertile, genetically heritable and dominant mutant in which suspensors develop a second embryo.

Our activation tagging screen relies on the use of a suspensor-specific *GAL4/UAS* enhancer trap line. This line and others have become very common tools in driving local expression of genes in *Arabidopsis*, but also as well-established cell and tissue identity markers. However, despite their abundant use, these *GAL4/UAS* enhancer trap lines are poorly characterized. **Chapter 5** describes the detailed characterization of the expression pattern and transgene insertion of a selection of 21 *GAL4/UAS* enhancer trap lines.

Chapter 6 summarizes and discusses the most important insights obtained in this thesis and provides directions for future research.

References

- Aida, M., Beis, D., Heidstra, R., Willemsen, V., Blilou, I., Galinha, C., Nussaume, L., Noh, Y.S., Amasino, R. and Scheres, B. (2004). The PLETHORA genes mediate patterning of the Arabidopsis root stem cell niche. *Cell* 119, 109-120.
- Asker, S.E. and Lenn, J. (1992). *Apomixis in plants* (Boca Raton: CRC Press).
- Barcaccia, G. and Albertini, E. (2013). Apomixis in plant reproduction: a novel perspective on an old dilemma. *Plant Reprod* 26, 159-179.
- Bard, J. (1994). *Embryos : color atlas of development* (London: Wolfe).
- Berleth, T. and Jürgens, G. (1993). The role of the MONOPTEROS gene in organising the basal body region of the Arabidopsis embryos. *Development* 118, 575-587.
- Bewley, J. and Michael, D.B. (1985). *Seeds : physiology of development and germination* (New York: Plenum Press).
- Bicknell, R.A. and Koltunow, A.M. (2004). Understanding apomixis: recent advances and remaining conundrums. *Plant Cell* 16, 228-245.
- Boutillier, K., Offringa, R., Sharma, V.K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.-M., van Lammeren, A.A.M., Miki, B.L.A., Custers, J. B. and van Lookeren Campagne, M.M. (2002). Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. *Plant Cell* 14, 1737-1749.
- Bouyer, D., Roudier, F., Heese, M., Andersen, E.D., Gey, D., Nowack, M.K., Goodrich, J., Renou, J.P., Grini, P.E., Colot, V. and Schnittger, A. (2011). Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. *PLoS Genet* 7(3): e1002014.
- Brand, A.H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401-415.
- Bratzel, F., López-Torrejón, G., Koch, M., Del Pozo, J.C. and Calonje, M. (2010). Keeping Cell Identity in Arabidopsis Requires PRC1 RING-Finger Homologs that Catalyze H2A Monoubiquitination. *Curr Biol* 20, 1853-1859.
- Braybrook, S.A. and Harada, J.J. (2008). LECs go crazy in embryo development. *Trends Plant Sci* 13, 624-630.
- Braybrook, S.A., Stone, S.L., Park, S., Bui, A.Q., Le, B.H., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2006). Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci USA* 103, 3468-3473.
- Brooker, R.J. (2011). *Biology* (New York: McGraw-Hill).
- Brown, D.C.W. and Thorpe, T.A. (1995). Crop improvement through tissue culture. *World J Microbiol and Biotechnol* 11, 409-415.
- Catanach, A.S., Erasmuson, S.K., Podivinsky, E., Jordan, B.R. and Bicknell, R. (2006). Deletion mapping of genetic regions associated with apomixis in Hieracium. *Proc Natl Acad Sci USA* 103, 18650-18655.
- De Rybel, B., Moller, B., Yoshida, S., Grabowicz, I., Barbier de Reuille, P., 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. *Dev Cell* 24, 426-437.
- De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W., Audenaert, D., Van Campenhout, J., Overvoorde, P., Jansen, L., Vanneste, S., 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. *Curr Biol* 20, 1697-1706.
- De Smet, I., Lau, S., Mayer, U. and Jurgens, G. (2010). Embryogenesis - the humble beginnings of plant life. *Plant J* 61, 959-970.

- de Vries, S.C., Booij, H., Meyerink, P., Huisman, G., Wilde, H.D., Thomas, T.L. and van Kammen, A. (1988). Acquisition of embryogenic potential in carrot cell-suspension cultures. *Planta* 176, 196-204.
- Dean Rider, S., Henderson, J.T., Jerome, R.E., Edenberg, H.J., Romero-Severson, J. and Ogas, J. (2003). Coordinate repression of regulators of embryonic identity by PICKLE during germination in *Arabidopsis*. *Plant J* 35, 33-43.
- DiMichele, W.A., Davis, J.I. and Olmstead, R.G. (1989). Origins of Heterospory and the Seed Habit: The Role of Heterochrony. *Taxon* 38, 1-11.
- Dinesh, D.C., Villalobos, L.I. and Abel, S. (2015). Structural Biology of Nuclear Auxin Action. *Trends Plant Sci* 21, 302-316.
- Dodeman, V., Ducreux G. and Kreis, M. (1997). Zygotic embryogenesis versus somatic embryogenesis. *J Exp Bot* 48, 1493-1509.
- Dubrovsky, J.G., Sauer, M., Napsucialy-Mendivil, S., Ivanchenko, M.G., Friml, J., Shishkova, S., Celenza, J. and Benková, E. (2008). Auxin acts as a local morphogenetic trigger to specify lateral root founder cells. *Proc Natl Acad Sci USA* 105, 8790-8794
- Friml, J. (2010). Subcellular trafficking of PIN auxin efflux carriers in auxin transport. *Eur J Cell Biol* 89, 231-235.
- Gaj, M.D., Zhang, S., Harada, J.J. and Lemaux, P.G. (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of *Arabidopsis*. *Planta* 222, 977-988.
- 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.
- Garces, H. and Sinha, N. (2009). The 'mother of thousands' (*Kalanchoe daigremontiana*): a plant model for asexual reproduction and CAM studies. *Cold Spring Harb Protoc* 2009 10: pdb emo133.
- Goldberg, R.B, de Paiva, G. and Yadegari, R. (1994). Plant embryogenesis: zygote to seed. *Science* 266, 605-614.
- Haccius, B. (1955). Experimentally Induced Twinning in Plants. *Nature* 176.
- 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.
- Harding, E.W., Tang, W., Nichols, K.W., Fernandez, D.E. and Perry, S.E. (2003). Expression and maintenance of embryogenic potential is enhanced through constitutive expression of AGAMOUS-Like 15. *Plant Physiol* 133, 653-663.
- Hardtke, C.S. and Berleth, T. (1998). The *Arabidopsis* gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO J* 17, 1405-1411.
- Haseloff, J. (1999). GFP variants for multispectral imaging of living cells. *Methods Cell Biol* 58, 139-151.
- Hecht, V., Vielle-Calzada, J.P., Hartog, M.V., Schmidt, E.D.L., Boutilier, K., Grossniklaus, U. and de Vries, S.C. (2001). The *Arabidopsis* Somatic Embryogenesis Receptor Kinase 1 Gene Is Expressed in Developing Ovules and Embryos and Enhances Embryogenic Competence in Culture. *Plant Physiol* 127, 803-816.
- Heck, G.R., Perry, S.E., Nichols, K.W. and Fernandez, D.E. (1995). AGL15, a MADS domain protein expressed in developing embryos. *Plant Cell* 7, 1271-1282.
- Hedden, P. and Phillips, A.L. (2000). Gibberellin metabolism: new insights revealed by the genes. *Trends Plant Sci* 5, 523-530.
- Hojsgaard, D.H., Martínez, E.J., Acuña, C.A., Quarin, C.L. and Pupilli, F. (2011). A molecular map of the apomixis-control locus in *Paspalum procurrens* and its comparative analysis with other species of *Paspalum*. *Theor Appl Genet* 123, 959-971.
- Horstman, A., Willemsen, V., Boutilier, K. and Heidstra, R. (2014). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends Plant Sci* 19, 146-157.
- Huang, S., Cerny, R.E., Bhat, D.S. and Brown, S.M. (2001). Cloning of an *Arabidopsis* Patatin-Like Gene, STURDY, by Activation T-DNA Tagging. *Plant Physiol* 125, 573-584.

- Ikeuchi, M., Sugimoto, K. and Iwase, A. (2013).** Plant Callus: Mechanisms of Induction and Repression. *Plant Cell* 25, 3159-3173.
- Johri, B.M., Ambegaokar, K.B. and Srivastava, P.S. (1992).** Comparative embryology of angiosperms (Berlin; New York: Springer-Verlag).
- Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K., Nguyen, J.T., Chory, J., Harrison, M.J. and Weigel, D. (1999).** Activation Tagging of the Floral Inducer FT. *Science* 286, 1962-1965.
- Koltunow, A.M., Bicknell, R.A. and Chaudhury, A.M. (1995).** Apomixis: Molecular Strategies for the Generation of Genetically Identical Seeds without Fertilization. *Plant Physiol* 108, 1345-1352.
- Köszegi, D., Johnston, A.J., Rutten, T., Czihal, A., Altschmied, L., Kumlehn, J., Wust, S.E., Kirioukhova, O., Gheyselinck, J., Grossniklaus, U. and Bäumllein, H. (2011).** Members of the RKD transcription factor family induce an egg cell-like gene expression program. *Plant J* 67, 280-291.
- Kwaaitaal, M.A. and de Vries, S.C. (2007).** The SERK1 gene is expressed in procambium and immature vascular cells. *J Exp Bot* 58, 2887-2896.
- Lakshmanan, K.K. and Ambegaokar, K.B. (1984).** Polyembryony. In *Embryology of Angiosperms*, B. Johri, ed. (Springer Berlin Heidelberg), pp. 445-474.
- Lau, S., Ehrismann, J.S., Schlereth, A., Takada, S., Mayer, U. and Jürgens, G. (2010).** Cell-cell communication in Arabidopsis early embryogenesis. *Eur J Cell Biol* 89, 225-230.
- Laux T., Mayer K.F., Berger J. and Jürgens G. (1996).** The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development* 122, 87-96.
- Laux, T. and Jürgens, G. (1997).** Embryogenesis: A New Start in Life. *Plant Cell* 9, 989-1000.
- Li, H., Soriano, M., Cordewener, J., Muiño, J.M., Riksen, T., Fukuoka, H., Angenent, G.C. and Boutilier, K. (2014).** The Histone Deacetylase Inhibitor Trichostatin A Promotes Totipotency in the Male Gametophyte. *Plant Cell* 26, 195-209.
- Lotan, T., Ohto, M., Yee, K.M., West, M.A., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1998).** Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93, 1195-1205.
- Luerssen, H., Kirik, V., Herrman, P. and Miséra, S. (1998).** FUSCA3 encodes a protein with a conserved VP1/AB13-like B3 domain which is of functional importance for the regulation of seed maturation in Arabidopsis thaliana. *Plant J* 15, 755-764.
- Luo, Y. and Koop, H.U. (1997).** Somatic embryogenesis in cultured immature zygotic embryos and leaf protoplasts of Arabidopsis thaliana ecotypes. *Planta* 202, 387-396.
- Maheshwari, P. (1950).** An introduction to the embryology of angiosperms (New York: McGraw-Hill).
- Mansfield, S.G. and Briarty, L.G. (1991).** Early embryogenesis in Arabidopsis thaliana. II. The developing embryo. *Can J Bot* 69, 461-476.
- Marsch-Martinez, N., Greco, R., Van Arkel, G., Herrera-Estrella, L. and Pereira, A. (2002).** Activation Tagging Using the En-I Maize Transposon System in Arabidopsis. *Plant Physiol* 129, 1544-1556.
- Mayer, U. and Jürgens, G. (1998).** Pattern formation in plant embryogenesis: A reassessment. *Semin Cell Dev Biol* 9, 187-193.
- Meinke, D.W. (1992).** A Homoeotic Mutant of Arabidopsis thaliana with Leafy Cotyledons. *Science* 258, 1647-1650.
- Meinke, D.W., Franzmann, L.H., Nickle, T.C. and Yeung, E.C. (1994).** Leafy Cotyledon Mutants of Arabidopsis. *Plant Cell* 6, 1049-1064.
- Möller, B. and Weijers, D. (2009).** Auxin control of embryo patterning. *Cold Spring Harb Perspect Biol* 1(5): a001545.
- Mordhorst, A.P., Toonen, M.A.J., de Vries, S.C. and Meinke, D. (1997).** Plant Embryogenesis. *Crit Rev Plant Sci* 16, 535-576.
- Naumova, T.N. (1993).** Apomixis in angiosperms: nucellar and integumentary embryony (Boca Raton, Fla.: CRC Press).

- Neff, M.M., Nguyen, S.M., Malancharuvil, E.J., Fujioka, S., Noguchi, T., Seto, H., Tsubuki, M., Honda, T., Takatsuto, S., Yoshida, S., and Chory, J. (1999). BAS1: A gene regulating brassinosteroid levels and light responsiveness in Arabidopsis. *Proc Natl Acad Sci USA* 96, 15316-15323.
- Ogas, J., Kaufmann, S., Henderson, J. and Somerville, C. (1999). PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in Arabidopsis. *Proc Natl Acad Sci USA* 96, 13839-13844.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M., Joosen, R., Lammers, M., Herdies, L., den Boer, B., van der Geest, L. and Boutilier, K. (2008). BABY BOOM target genes provide diverse entry points into cell proliferation and cell growth pathways. *Plant Mol Biol* 68, 225-237.
- Peris, C.I., Rademacher, E.H. and Weijers, D. (2010). Green beginnings - pattern formation in the early plant embryo. *Curr Top Dev Biol* 91, 1-27.
- Petrášek, J. and Friml, J. (2009). Auxin transport routes in plant development. *Development* 136, 2675-2688.
- Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Freire Rios, A., Borst, J.W., Lukowitz, W., Jurgens, G. and Weijers, D. (2012). Different auxin response machineries control distinct cell fates in the early plant embryo. *Dev Cell* 22, 211-222.
- Rademacher, E.H., Moller, B., Lokerse, A.S., Llavata-Peris, C.I., van den Berg, W. and Weijers, D. (2011). A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. *Plant J* 68, 597-606.
- Raven, P.H., Evert, R.F. and Eichhorn, S.E. (1999). *Biology of plants* (New York: W.H. Freeman : Worth Publishers).
- Reinert, J. (1959). Über die Kontrolle der Morphogenese und die Induktion von Adventivembryonen an Gewebekulturen aus Karotten. *Planta* 53, 318-333.
- 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 Physiol* 135, 1738-1752.
- Rothwell, G.W. and Erwin, D.M. (1987). Origin of Seed Plants: An Aneurophyte/Seed-Fern Link Elaborated. *Am J Bot* 74, 970-973.
- Sato, A. and Yamamoto, K.T. (2008). Overexpression of the non-canonical Aux/IAA genes causes auxin-related aberrant phenotypes in Arabidopsis. *Physiol Plant* 133, 397-405.
- Schlereth, A., Möller, B., Liu, W., Kientz, M., Flipse, J., Rademacher, E.H., Schmid, M., Jurgens, G. and Weijers, D. (2010). MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. *Nature* 464, 913-916.
- Schmidt, E.D., Guzzo, F., Toonen, M.A. and de Vries, S.C. (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* 124, 2049-2062.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F., 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.
- Schwartz, B.W., Yeung E.C. and Meinke D.W. (1994). Disruption of morphogenesis and transformation of the suspensor in abnormal suspensor mutants of Arabidopsis. *Development* 120, 3235.
- Segui-Simarro, J.M. and Nuez, F. (2008). How microspores transform into haploid embryos: changes associated with embryogenesis induction and microspore-derived embryogenesis. *Physiol Plant* 134, 1-12.
- Smit, M.E. and Weijers, D. (2015). The role of auxin signaling in early embryo pattern formation. *Curr Opin Plant Biol* 28, 99-105.
- Soriano, M., Li, H. and Boutilier, K. (2013). Microspore embryogenesis: establishment of embryo identity and pattern in culture. *Plant Reprod* 26, 181-196.
- Spillane, C., Curtis, M.D. and Grossniklaus, U. (2004). Apomixis technology development-virgin births in farmers' fields? *Nat Biotechnol* 22, 687-691.

- Stepanova, A.N., Robertson-Hoyt, J., Yun, J., Benavente, L.M., Xie, D.-Y., Doležal, K., 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.
- Steward, F.C., Mapes Marion O. and Joan, S. (1958). Growth and Organized Development of Cultured Cells. I. Growth and Division of Freely Suspended Cells. *Am J Bot* 45, 693-703.
- Stone, S.L., Braybrook, S.A., Paula, S.L., Kwong, L.W., Meuser, J., Pelletier, J., Hsieh, T.F., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2008). Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proc Natl Acad Sci USA* 105, 3151-3156.
- Stone, S.L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2001). LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci USA* 98, 11806-11811.
- Supena, E.D.J., Winarto, B., Riksen, T., Dubas, E., van Lammeren, A., Offringa, R., Boutillier, K. and Custers, J. (2008). Regeneration of zygotic-like microspore-derived embryos suggests an important role for the suspensor in early embryo patterning. *J Exp Bot* 59, 803-814.
- Suzuki, Y., Uemura, S., Saito, Y., Murofushi, N., Schmitz, G., Theres, K. and Yamaguchi, I. (2001). A novel transposon tagging element for obtaining gain-of-function mutants based on a self-stabilizing Ac derivative. *Plant Mol Biol* 45, 123-131.
- Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.
- 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.L., 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, 164-76.
- Taylor, E.L., Krings, M. and Taylor, T.N. (2009). *Paleobotany the biology and evolution of fossil plants* (Amsterdam; Boston: Academic Press).
- Touraev, A., Vicente, O. and Heberle-Bors, E. (1997). Initiation of microspore embryogenesis by stress. *Trends Plant Sci* 2, 297-302.
- Tsukagoshi, H., Morikami, A. and Nakamura, K. (2007). Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in Arabidopsis seedlings. *Proc Natl Acad Sci USA* 104, 2543-2547.
- Tsuwamoto, R., Yokoi, S. and Takahata, Y. (2010). Arabidopsis EMBRYOMAKER encoding an AP2 domain transcription factor plays a key role in developmental change from vegetative to embryonic phase. *Plant Mol Biol* 73, 481-492.
- Tzafirir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T.C., McElver, J., Aux, G., Patton, D. and Meinke, D. (2004). Identification of genes required for embryo development in Arabidopsis. *Plant Physiol* 135, 1206-1220.
- 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.
- Vernon, D. and Meinke, D. (1994). Embryogenic transformation of the suspensor in twin, a polyembryonic mutant of Arabidopsis. *Dev Biol* 165, 566-573.
- Vijverberg, K., Milanovic-Ivanovic, S., Bakx-Schotman, T. and van Dijk, P.J. (2010). Genetic fine-mapping of DIPLOSPOROUS in Taraxacum (dandelion; Asteraceae) indicates a duplicated DIP-gene. *BMC Plant Biol* 10, 154.
- Waki, T., Hiki, T., Watanabe, R., Hashimoto, T. and Nakajima, K. (2011). The Arabidopsis RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. *Curr Biol* 21, 1277-1281.

- Waki, T., Miyashima, S., Nakanishi, M., Ikeda, Y., Hashimoto, T. and Nakajima, K. (2013). A GAL4-based targeted activation tagging system in *Arabidopsis thaliana*. *The Plant journal : for cell and molecular biology* 73, 357-367.
- Wang, F. and Perry, S.E. (2013). Identification of direct targets of FUSCA3, a key regulator of *Arabidopsis* seed development. *Plant Physiol* 161, 1251-1264.
- Wang, H., Caruso, L.V., Downie, A.B. and Perry, S.E. (2004). The Embryo MADS Domain Protein AGAMOUS-Like 15 Directly Regulates Expression of a Gene Encoding an Enzyme Involved in Gibberellin Metabolism. *Plant Cell* 16, 1206-1219.
- Wang, R. and Estelle, M. (2014). Diversity and specificity: auxin perception and signaling through the TIR1/AFB pathway. *Curr Opin Plant Biol* 21, 51-58.
- Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E.J., Neff, M.M., Nguyen, J.T., Sato, S., Wang, Z.Y., Xia, Y., Dixon, R.A., Harrison, M.J., Lamb, C.J., Yanofsky, M.F. and Chory, J. (2000). Activation tagging in *Arabidopsis*. *Plant Physiol* 122, 1003-1013.
- Weijers, D., Schlereth, A., Ehrismann, J.S., Schwank, G., Kientz, M. and Jurgens, G. (2006). Auxin triggers transient local signaling for cell specification in *Arabidopsis* embryogenesis. *Dev Cell* 10, 265-270.
- Weijers, D., Van Hamburg, J.P., Van Rijn, E., Hooykaas, P.J. and Offringa, R. (2003). Diphtheria toxin-mediated cell ablation reveals interregional communication during *Arabidopsis* seed development. *Plant Physiol* 133, 1882-1892.
- Wendrich, J.R. and Weijers, D. (2013). The *Arabidopsis* embryo as a miniature morphogenesis model. *New Phytol* 199, 14-25.
- West, M. and Harada, J.J. (1993). Embryogenesis in Higher Plants: An Overview. *Plant Cell* 5, 1361-1369.
- West, M., Yee, K.M., Danao, J., Zimmerman, J.L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1994). LEAFY COTYLEDON1 Is an Essential Regulator of Late Embryogenesis and Cotyledon Identity in *Arabidopsis*. *Plant Cell* 6, 1731-1745.
- Yamaguchi, N., Wu, M.F., Winter, C.A., Berns, M., Markus, C., Nole-Wilson, S., Yamaguchi, A., Coupland, G., Krizek, B. and Wagner, D. (2013). A Molecular Framework for Auxin-Mediated Initiation of Flower Primordia. *Dev Cell* 24, 271-282.
- Yang, C., Bratzel, F., Hohmann, N., Koch, M., Turck, F. and Calonje, M. (2013). VAL- and AtBMI1-Mediated H2Aub Initiate the Switch from Embryonic to Postgerminative Growth in *Arabidopsis*. *Curr Biol* 23, 1324-1329.
- Yeung, E.C. and Clutter, M.E. (1979). Embryogeny of *Phaseolus coccineus*: the ultrastructure and development of the suspensor. *Can J Bot* 57, 120-136.
- Yeung, E.C. and Meinke, D.W. (1993). Embryogenesis in Angiosperms: Development of the Suspensor. *Plant Cell* 5, 1371-1381.
- Zaki, M.A.M. and Dickinson, H.G. (1991). Microspore-derived embryos in Brassica: the significance of division symmetry in pollen mitosis I to embryogenic development. *Sex Plant Reprod* 4, 48-55.
- Zhang, Z. and Laux, T. (2011). The asymmetric division of the *Arabidopsis* zygote: from cell polarity to an embryo axis. *Sex Plant Reprod* 24, 161-169.
- Zhao, Y., Christensen, S.K., Fankhauser, C., Cashman, J.R., Cohen, J.D., Weigel, D. and Chory, J. (2001). A Role for Flavin Monooxygenase-Like Enzymes in Auxin Biosynthesis. *Science* 291, 306-309.
- Zheng, Y., Ren, N., Wang, H., Stromberg, A.J. and Perry, S.E. (2009). Global Identification of Targets of the *Arabidopsis* MADS Domain Protein AGAMOUS-Like15. *Plant Cell* 21, 2563-2577.
- Zimmerman, J.L. (1993). Somatic Embryogenesis: A Model for Early Development in Higher Plants. *Plant Cell* 5, 1411-1423.
- Zuo, J., Niu, Q.W., Frugis, G. and Chua, N.H. (2002). The WUSCHEL gene promotes vegetative-to-embryonic transition in *Arabidopsis*. *Plant J* 30, 349-359.

Chapter 2

A novel ARF/bHLH module regulates extra-embryonic identity during *Arabidopsis* embryogenesis

Tatyana Radoeva^{1,*}, Annemarie S. Lokerse^{1,a,*}, Cristina I. Llavata-Peris^{1,b},
Jos R. Wendrich¹, Daoquan Xiang², Che-Yang Liao¹, Jan Willem Borst^{1,3},
Lieke Vlaar¹, Mark Boekschoten^{4,5}, Guido Hooiveld⁴, Raju Datla²,
and Dolf Weijers¹

1. Laboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703HA Wageningen, the Netherlands
 2. Plant Biotechnology Institute, National Research Council, Saskatoon, Saskatchewan, Canada S7N 0W9
 3. Microspectroscopy Center, Wageningen University, Dreijenlaan 3, 6703HA Wageningen, the Netherlands
 4. Division of Human Nutrition, Wageningen University, Dreijenlaan 2, 6703HA Wageningen, the Netherlands
 5. TI Food & Nutrition Wageningen, the Netherlands
- a. Present address: Rijk Zwaan Breeding B.V.
Eerste Kruisweg 9, Fijnaart, | The Netherlands
- b. Present address: Dyadic Nederland B.V., Nieuwe kanaal 7-s, 6709 PA, Wageningen, the Netherlands
- * These authors contributed equally

Land plants can reproduce sexually by developing an embryo from a fertilized egg cell. However, embryos can also be formed from other cell types in many plant species. A key question is thus how embryo identity in plants is controlled, and how this process is modified during non-zygotic embryogenesis. The *Arabidopsis* zygote divides to produce an embryonic lineage and an extra-embryonic suspensor. Yet, normally quiescent suspensor cells can develop a second embryo when the initial embryo is damaged, or when response to the signaling molecule auxin is locally blocked. Here we have used auxin-dependent suspensor embryogenesis as a model to determine transcriptome changes during embryonic reprogramming. We find that reprogramming is complex, and used large-scale *in vivo* expression analysis to isolate genes involved in cellular reprogramming. We identify an auxin-dependent bHLH transcription factor network that mediates the activity of this hormone in suppressing embryo identity in the suspensor.

Introduction

In many land plants, including *Arabidopsis thaliana*, zygotic embryogenesis begins with an asymmetric cell division, generating two cells with distinct fates. The small apical cell is the founder of the pro-embryo and will form the majority of the plant. The larger basal cell divides several times and gives rise to the suspensor, a filamentous support structure of which the uppermost cell generates part of the seedling root (Laux and Jürgens, 1997; Mayer and Jürgens, 1998). Hence, the suspensor is largely an extra-embryonic, yet zygote-derived structure. In contrast to the pro-embryo, the suspensor is already fully developed at globular stage and plays an important role in embryo development (Radoeva and Weijers, 2014; Schwartz et al., 1997). The suspensor provides mechanistic and nutritional support required for the growing embryo as well as a connection between the pro-embryo and the maternal endosperm (Kawashima and Goldberg, 2010; Raghavan, 2006).

Despite their mitotic quiescence under normal conditions, suspensor cells in several species have the potential to generate a new embryo (Lakshmanan and Ambegaokar, 1984). In *Arabidopsis*, mutations that impair growth of the proembryo (*raspberry* (*rsy*), *suspensor* (*sus*) and *twin* (*twi*)) can cause suspensor proliferation, eventually recapitulating embryogenesis and generating a new pro-embryo (Schwarz et al., 1994; Vernon and Meinke, 1994). Suspensor-derived embryogenesis can also be induced by experimental ablation of the pro-embryo through radiation or chemicals, or through genetic ablation (Haccius, 1955; Weijers et al., 2003). These observations reveal the developmental potential of the suspensor to undergo embryonic transformation and imply that its potential is limited by normal growth of the embryo proper. While regulatory mechanisms underlying this switch in fate are largely unknown, we recently reported that this involves the signaling molecule auxin (Rademacher et al., 2012). Components of auxin response are expressed in suspensor cells (Rademacher et al., 2012; Rademacher et al., 2011), and when these are inhibited, suspensor cells proliferate, express embryo markers, and develop a second embryo, which can ultimately give rise to twin-like seedlings (Rademacher et al., 2012). At present it is unclear how direct the involvement of auxin response in embryonic fate conversion is, but it provides a good entry point into studying the regulatory mechanisms underlying suspensor to embryo transformation (abbreviated S > E henceforth).

Auxin response is mediated by transcription factors of the AUXIN RESPONSE FACTOR (ARF) family, which are inhibited by interacting Aux/IAA proteins. In the presence of auxin, Aux/IAA proteins are ubiquitinated by the SCF-TIR1/AFB auxin receptor complex (reviewed in Wang and Estelle, 2014) and marked for degradation.

This releases ARFs to regulate transcription of primary target genes (Lokerse and Weijers, 2009; Wang and Estelle, 2014). While several target genes mediating auxin response in embryonic root initiation (Crawford et al., 2015; Schlereth et al., 2010) lateral root development (GATA23, De Rybel et al., 2010; LBD, Okushima et al., 2007) and flower development (LFY, Yamaguchi et al., 2013; ARR7/15, Zhao et al., 2010) have been isolated, their role in maintaining suspensor identity is not yet characterized. Here we use auxin-dependent, suspensor-derived embryogenesis (reviewed in Radoeva and Weijers, 2014) to identify molecular components that mediate this important cell fate transformation.

Through genome-wide transcriptional profiling upon local auxin response inhibition, we identify a set of genes that is regulated during suspensor-derived embryogenesis. We characterized in detail a set of bHLH transcription factors that mediate suspensor to embryo transformation. Previously, bHLH transcription factors have been identified as direct ARF target genes (De Rybel et al., 2013; Schlereth et al., 2010). Our work shows that an ARF-bHLH49 module is involved in suspensor development, suggesting that ARF/bHLH modules could very well be a more general principle in mediating the multitude of auxin-controlled plant developmental processes.

Results

Transcriptional analysis of suspensor reprogramming

Expression of the stabilized transcriptional auxin response inhibitor protein *iaa12/bodenlos* (*bdl*) in suspensor cells efficiently induces a switch from extra-embryonic to embryonic identity (Rademacher et al., 2012). We used this predictable, uniform response to identify genes whose expression changes during this fate transition. IAA12/BDL protein is normally degraded in response to auxin (Dharmasiri et al., 2005), but a P74S mutation in the BDL protein prevents degradation and leads to accumulation of this ARF inhibitor (Dharmasiri et al., 2005; Hamann et al., 2002). We expressed mutant *bdl* protein using the two-component GAL4/UAS system (Weijers et al., 2006). The GAL4 driver line M0171 is active in suspensor cells (Figure 1A) until heart stage, after which the GAL4 expression expands to include cells in the pro-embryo (Rademacher et al., 2012). By crossing homozygous M0171 and UAS-*bdl* plants, suspensor-to-embryo (S>E) transformation could be induced in most embryos at heart stage (83%; n=126). For whole-genome transcriptomic analysis, we first identified time points after pollination and after onset of M0171 expression, when the first aberrant divisions occurred. We found that by 72 hours,

about one third of suspensors showed a first aberrant division (Figure 1B,C; 31%; n = 149), and selected this as the first time point. We also included a 96-hour time point as by this time, fate conversion was completed (Figure 1D,E).

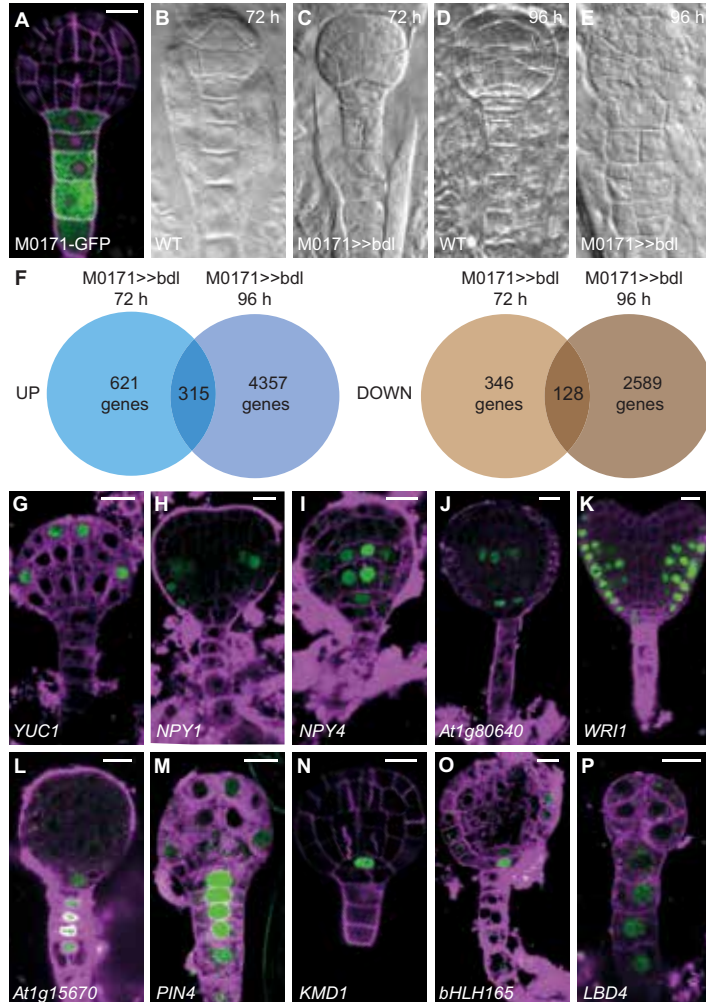


Figure 1: Transcriptional analysis of suspensor to embryo transformation.

(A) Expression of M0171-GFP in globular stage embryo. (B-E) Embryos from crosses of between M0171 and wild type (B and D) or M0171 and UAS-bdl (C and E) prepared 72 hours (B and C) or 96 hours post pollination (D and E). (F) Venn diagrams of genes misregulated in M0171 >> bdl embryos 72 hours- and 96 hours after pollination. (G-P) Expression of promoter-n3GFP reporters for *YUC1* (G), *NPY1* (H), *NPY4* (I), *At1g80640* (J), *WRI1* (K), *At1g15670* (L), *PIN4* (M), *KMD1* (N), *bHLH165* (O), *LBD4* (P) in globular stage (G, I, J, L, M, N, O and P) or heart stage (H and K) embryos. Scale bar represents 10 μ m in all panels.

M0171 plants were pollinated with UAS-*bdl* pollen (M0171 > > *bdl*), embryos were manually dissected after 72 and 96 hours and 300-400 individuals per biological replicate were pooled for RNA extraction. In parallel, M0171 plants were pollinated with Col-0 wild-type pollen to serve as isogenic wild-type controls (M0171 > > Col-0). Four replicates of each cross, and each time-point were hybridized to long-oligo microarrays (Xiang et al., 2011). After statistical analysis, we identified 621 and 359 genes (Figure 1F) that were robustly either up- or down-regulated in M0171 > > *bdl* 72 hours embryos compared to the control (>2-fold misregulated; FDR < 0.055; Supplemental Dataset 1). Analysis of functional categories did not reveal any enrichment among the differentially expressed genes when all were considered. When considering only the 100 most strongly affected genes however, there was an over-representation of transcription factors, auxin regulated genes as well as zygote-enriched genes. We therefore conclude that, despite the sampling of embryos with minimal morphological changes, the transcriptional reprogramming is complex, and likely encompasses secondary changes. We focused on transcription factors and signaling factors as potential primary regulators in follow-up studies.

Identification of differentially expressed regulatory genes

To explore the M0171 > > *bdl* transcriptome, we initially surveyed the expression of genes with known function in auxin response, embryo initiation or embryo development. As expected, *BDL/IAA12* is 10.6-fold upregulated and conversely *IAA10*, whose suspensor-specific expression (at earlier stages of embryogenesis) is lost in *pARF13::iaa10* background (Rademacher et al., 2012), is 3.5-fold downregulated in M0171 > > *bdl* dataset. Several other direct auxin response genes (*IAA19*, *20*, *26*, *30*; *GH3.3*, *3.4*, *3.5*, *3.17*; and several SAUR genes) were also strongly downregulated in M0171 > > *bdl* embryos (Table S1), which confirms that the transcriptional auxin response is indeed inhibited in these embryos.

Only a handful of genes were shown to induce embryogenesis when ectopically expressed (reviewed in Radoeva and Weijers, 2014). Of these, two are differentially expressed in M0171 > > *bdl* dataset. LEC1-LIKE (*LIL1*), also known as NF-YB6, is a subunit of NF-Y transcription factor complex and is able to induce embryo development in somatic cells when overexpressed (Kwong et al., 2003; Lotan et al., 1998). *LIL1* is 2.5-fold upregulated in our dataset. Another factor implicated in somatic embryogenesis that is 1.9-fold upregulated in our dataset is *FUSCA3* (*FUS3*), a B3 domain transcription factor that was shown to induce late embryo properties in somatic cells after ectopic expression (Gaj et al., 2005). Moreover, an important embryo marker, *SHOOT MERISTEMLESS* (*STM*) (Long et al., 1996) shown

to be expressed in proliferating suspensor cells (Rademacher et al., 2012) is 2.3-fold upregulated in M0171 >> bdl dataset, while the *CLAVATA3 (CLV3)* gene, expressed in the shoot apical meristem (Fletcher et al., 1999) was 45-fold upregulated. These examples confirm the overall validity of the generated microarray dataset.

We next selected 68 genes (Table S2) from the significantly misregulated genes to validate whether the microarray results report local expression changes in the embryo. This selection was based on the amplitude of misregulation (Fold-change M0171 >> bdl versus M0171 >> Col-0), the functional annotation (transcription factors, signaling components) and known auxin-responsiveness (using public array data). For these 68 genes, promoter-reporter lines were generated using a sensitive nuclear-localized 3xGFP (n3GFP). Patterns of expression were analyzed in wild-type embryos for all genes, but only 40 were expressed in the embryo and/or suspensor. Since our microarray analysis was aimed at identifying genes that are misregulated as part of S>E transformation we predict that, ideally, up-regulated genes are normally expressed in embryo cells and down-regulated genes in suspensor cells. While a few genes showed patterns of expression that were inconsistent with this prediction, we found several genes that conformed to these criteria (Figure 1G-P; Table S2). Notably, this subset of validated genes contained genes encoding basic Helix-Loop-Helix transcription factors (bHLHs). bHLHs are well-known regulators of cell identity in multicellular organisms (Murre et al., 1994), including plants (Feller et al., 2011). Importantly, other bHLHs were previously shown to mediate auxin-dependent development (Chandler et al., 2009; De Rybel et al., 2013; Schlereth et al., 2010). We next focused our analysis on these genes as potential regulators of S>E transformation.

Auxin-dependent expression of bHLH genes

The bHLH genes identified here (*bHLH49*, *60*, *63* and *100*) belong to two different clades, 12 and 25 (Figure 2A). While a post-embryonic function has been described for bHLH63, also known as CRYPTOCHROME-INTERACTING bHLH1 (CIB1; Liu et al., 2008) and bHLH100 (Andriankaja et al., 2014; Sivitz et al., 2012; Wang et al., 2007), no embryonic function has been reported for any of these genes. To determine whether these four bHLH genes are involved in S>E transformation, we first studied their embryonic expression domains in detail. Expression of *bHLH49* is found in the basal tier of the embryo and the suspensor, but later (at later heart and torpedo stages) becomes restricted to the proembryo (Figure 2B-D,L). Expression of *bHLH60* is found in the outer cells of the proembryo and a weak transient expression is detected in the suspensor (Figure 2E,F,M). The expression of

bHLH63 is limited to the protoderm at globular stage and from heart stage on is also seen in the subprotodermal cells at the future hypocotyl-cotyledon junction (Figure 2G, 2H). These dynamic, yet largely proembryo-specific expression patterns are consistent with their upregulation in the microarray experiment. Consistent with its downregulation in M0171 >> bdl microarray, *bHLH100* shows suspensor-specific expression until globular stage, after which the domain extends to protodermal cells in the basal tier (Figure 2I-K).

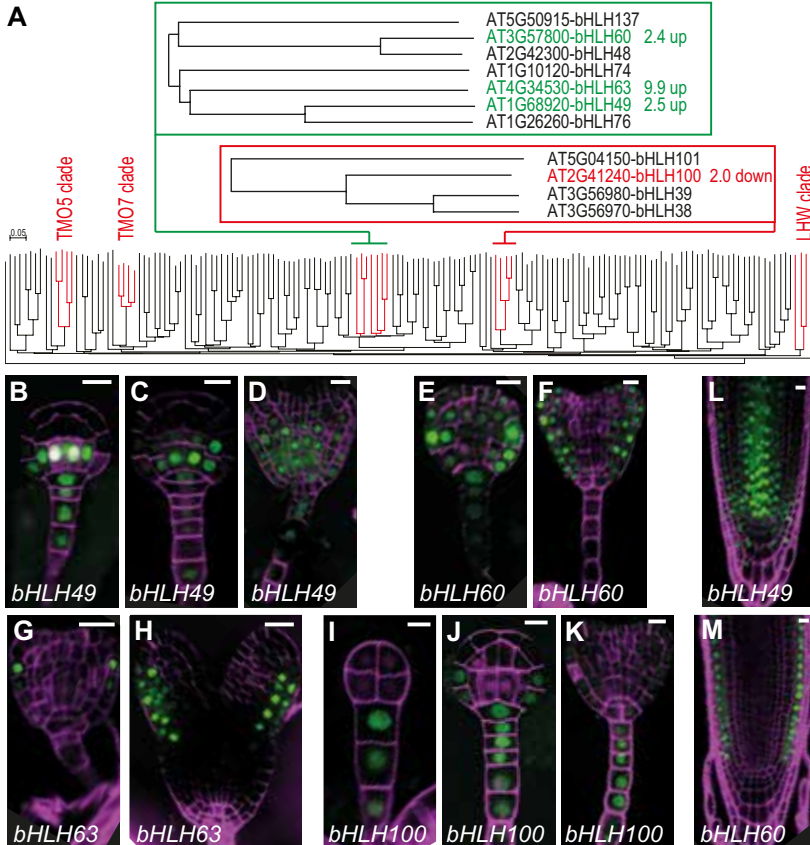


Figure 2: Expression of bHLH genes in the Arabidopsis embryo.

(A) Phylogenetic tree of Arabidopsis bHLH proteins, indicating the divergent positions of the *TMO5*, *TMO7* and *LHW* clades relative to the misregulated bHLH genes. The misregulation of genes in the M0171 >> bdl embryo is indicated in the upper panels. (B-M) Expression of promoter-n3GFP reporters for *bHLH49* (B-D,L), *bHLH60* (E,F,M), *bHLH63* (G,H) and *bHLH100* (I-K) in globular stage (B,C,E,G,J), heart stage (D,F,H,K), octant stage (I) embryos and in roots (L,M). Scale bar indicates 10 μm in all panels.

While some bHLH proteins act cell-autonomously in the cells where the gene is transcribed (e.g. TMO5; Schlereth et al., 2010) others act non cell-autonomously by moving to adjacent cells (e.g. TMO7; Schlereth et al., 2010; UPBEAT1; Tsukagoshi et al., 2010). To determine whether the bHLH genes identified here are mobile, we generated translational fusions of genomic fragments fused to sensitive sYFP protein. Consistent with their function as transcription factors, all four bHLH proteins localized to the nucleus. Protein localization domains exactly matched the promoter expression domains (Figure S1), demonstrating that these proteins likely do not move and that the protein accumulation is transcriptionally controlled.

These bHLH genes were identified based on their misexpression upon auxin response inhibition, and the genes are expressed in the relevant cell types within the embryo. We next examined whether the bHLH genes were indeed rapidly regulated by auxin. We tested the effect of exogenous auxin on transcript levels in seedling roots. *bHLH60* and *bHLH63* were upregulated within 1 hour and *bHLH100* was upregulated within 2 hours of auxin treatment (Figure 3A). In contrast, *bHLH49* was downregulated, suggesting that all four bHLHs are regulated by auxin, although it cannot be concluded whether the regulation is direct or indirect. To examine if bHLH regulation involves ARF transcription factors, we determined transcript levels in a range of *arf* mutants. We showed previously that 6 ARFs (1, 2, 6, 9, 13 and 18) are expressed in the suspensor at globular stage embryo (Rademacher et al., 2011), which qualifies these as potential regulators of bHLH expression. Indeed, although no direct correlation between individual ARFs and bHLHs could be made, the expression of all four genes was destabilized in *arf* mutants (Figure 3B), suggesting complex ARF-dependent regulation. To test if auxin-dependent bHLH regulation is direct, we performed a 2-hour auxin treatment in the presence of cycloheximide (CHX). CHX is an inhibitor of protein biosynthesis and in its presence only direct transcriptional regulation should happen (Franco et al., 1990). Auxin-dependent repression of *bHLH49* expression was still observed in the presence of CHX, which strongly suggests that *bHLH49* is directly regulated by auxin/ARFs. In contrast, activation of the *bHLH60*, *63* and *100* genes was suppressed by CHX, which suggests that auxin regulation is not a direct transcriptional response (Figure 3C).

In summary, our transcriptome studies identified 4 bHLH genes that are expressed in pro-embryo (*bHLH49*, *60*, *63*) or suspensor (*bHLH100*) in an auxin/ARF-dependent fashion, and whose expression changes during S>E transformation.

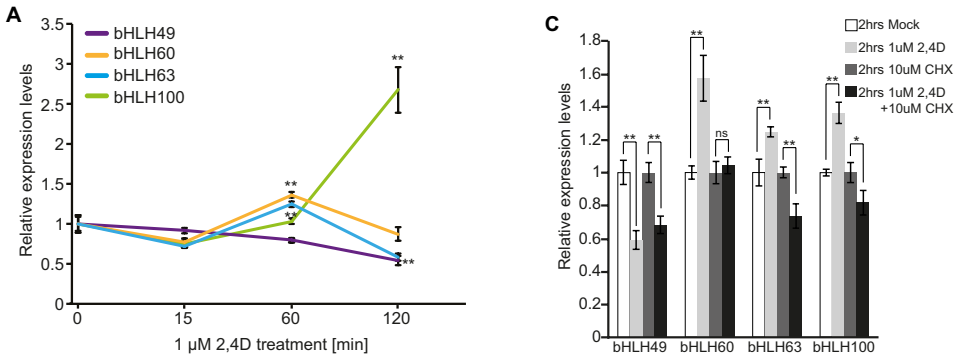


Figure 3: Auxin-dependent expression of bHLH genes.

(A) Relative expression levels (qRT-PCR) of *bHLH49*, *bHLH60*, *bHLH63* and *bHLH100* in roots upon treatment with 1 μM 2,4 dichlorophenoxyacetic acid (2,4D) for 0, 15, 60 and 120 minutes. (B) Relative qRT-PCR expression levels of *bHLH49*, 60, 63 and 100 in roots of *arf1*, *arf2*, *arf6*, *arf9*, *arf13* and *arf18* mutants. (C) Relative expression levels of *bHLH49*, 60, 63 and 100 in roots upon treatment with 1 μM 2,4D, 10 μM cycloheximide (CHX) or both for 2 hours. Expression levels in untreated wild-type (or mock treated samples) were set to 1. Error bars indicate SE; t test: *P < 0,05, **P < 0,001.

bHLH genes mediate suspensor development

To determine the role of *bHLH49*, 60, 63 and 100 during S>E transformation, we identified and characterized insertion mutants with strongly reduced corresponding transcript levels (Figure S2). Strikingly, all lines showed embryo defects with high penetrance (frequencies ranging from 7 to 25%, Figure 4A-E). These phenotypes appear to be the consequence of *bHLH* mutations, rather than unlinked mutations, as the same phenotypes were found in multiple independent alleles for *bHLH49* and *bHLH100*. Regardless of the rather broad expression of the genes, the phenotype of the loss-of-function mutants is specific to the suspensor and/or the hypophysis, suggesting a contribution to normal suspensor and/or hypophysis development. In all mutants, divisions were aberrant at the suspensor-embryo junction. It could not be unequivocally determined if the primary defect originated in the pro-embryo, or in the suspensor (Figure 4A-E). Nonetheless, this result suggests that all 4 *bHLH* genes are required for normal embryo development, in particular for the embryo-suspensor junction. *bHLH100* was identified based on its down-regulation in association with S>E transformation, consistent with its loss-of-function phenotype. The other three

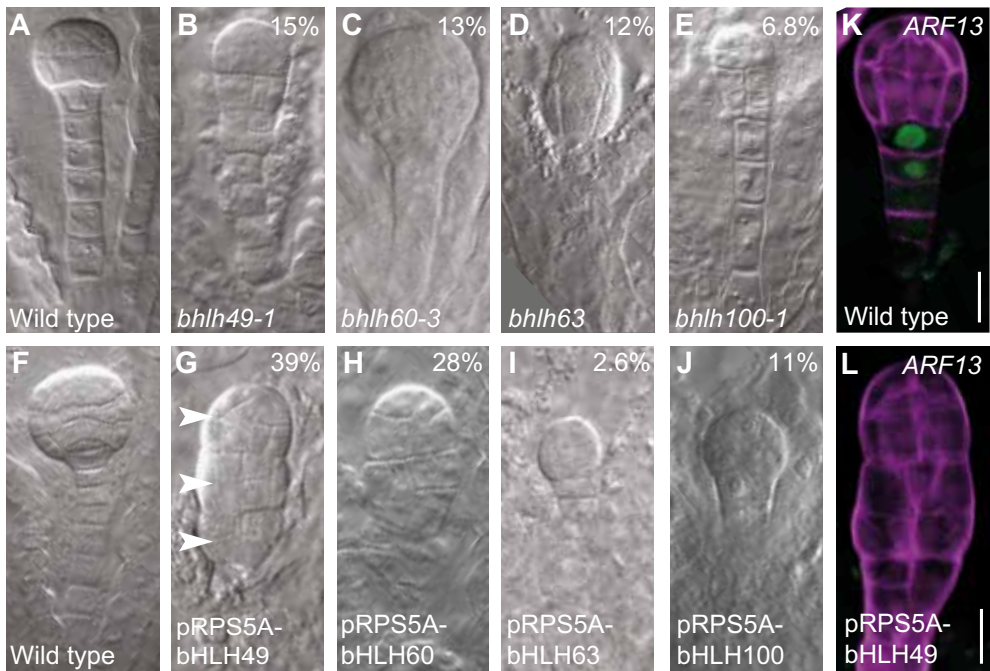


Figure 4: bHLH genes contribute to normal embryo development.

(A-E) Embryo phenotypes of wild-type (A), and mutants in *bHLH49* (B), *bHLH60* (C), *bHLH63* (D) or *bHLH100* (E). (F-J) Embryo phenotypes of wild-type (F), pRPS5A-bHLH49 (G), pRPS5A-bHLH60 (H), pRPS5A-bHLH63 (I) and pRPS5A-bHLH100 (J). (K,L) Expression of pARF13-n3GFP in wild-type (K) and pRPS5A-bHLH49 (L). Scale bar represents 10 μ m in all panels. Arrowheads in (G) indicate the pro-embryos. Percentages indicate penetrance of phenotypes in homozygous mutants and overexpression lines (*bhlh49-1*: 15%, n = 137; *bhlh49-2*: 25%, n = 83; *bhlh60-3*: 13%, n = 265; *bhlh63*: 12%, n = 115; *bhlh100-1*: 6.8%, n = 118; *bhlh100-2*: 22%, n = 101; pRPS5A-bHLH49: 39%, n = 185; pRPS5A-bHLH60: 28%, n = 128; pRPS5A-bHLH63: 2.6%, n = 117; pRPS5A-bHLH100: 11%, n = 175).

genes were identified based on their up-regulation during the same transformation. While the loss-of-function phenotype suggests a requirement for these genes during normal embryo development, it is unclear whether their up-regulation contributes to S>E transformation in M0171 >> bdl embryos. To determine if the up-regulation contributes to abnormal suspensor development, we misexpressed each gene individually, using the *RPS5A* promoter, which drives strong expression in the whole embryo, including the suspensor (Weijers et al., 2001; Weijers et al., 2006). Analysis of four independent lines per construct reveals that bHLH49 and bHLH60 are able to induce severe suspensor phenotypes, while bHLH63 and 100 are less able to do so (Figure 4F-J). Particularly, pRPS5A-bHLH49 embryos showed proliferated suspenders with embryo-like appearance; in some cases 3 pro-embryo-like structures were stacked on top of each other (Figure 4G; arrowheads). Strikingly,

the suspensor-specific expression of *pARF13-n3GFP* (Rademacher et al., 2011) was lost in the proliferating suspensor cells (Figure 4K,L). It is therefore conceivable that bHLH49 is by itself sufficient to induce transformation from extra-embryonic (suspensor) to embryonic identity and thus limiting the expression of *bHLH49* in the suspensor is important for cell identity specification.

bHLH049 mediates auxin-dependent growth

The auxin-dependent *bHLH49* gene is required for normal embryo development and its misexpression is sufficient to induce embryo-like structures in the suspensor. To determine if its function contributes to the developmental output of auxin signaling, we examined auxin sensitivity in *bhlh49* mutants and in the pRPS5A-bHLH49 misexpression lines. As no facile auxin sensitivity assay is available in the embryo, this was performed in post-embryonic roots, in which *bHLH49* is expressed in an auxin-dependent manner. Changes in root growth inhibition by the synthetic auxin 2,4-D can be quantified and used to determine the auxin sensitivity (Lincoln et al., 1990). Our results show that auxin sensitivity was greatly enhanced in pRPS5A-bHLH49 roots, while *bhlh49* mutant seedlings are slightly resistant to low concentrations of auxin (Figure 5A). Hence, *bHLH49* is not only regulated by auxin, but its repression is also required for normal auxin response during root growth.

Finally, because auxin activity is characterized by many feedback loops (Benjamins and Scheres, 2008; Jones et al., 2010; Leyser, 2010), we determined if bHLH49 is a proper auxin output, or if it feeds back into auxin activity. We introduced the *pDR5-GFP* auxin response reporter (Friml et al., 2003; Ulmasov et al., 1997) into pRPS5A-bHLH49 misexpression lines and found DR5-GFP expression to be unchanged (Figure 5B-E), despite the clear developmental defects in this line, suggesting no feedback, but linear activity.

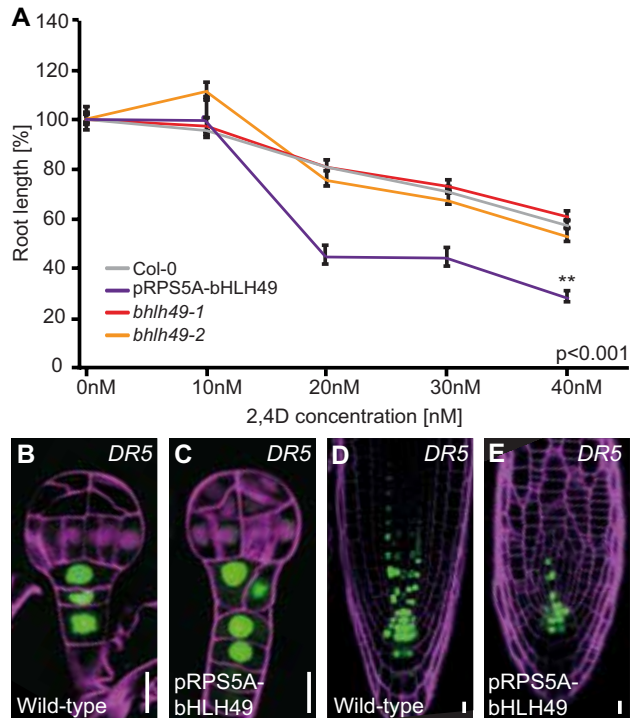


Figure 5: bHLH49 acts downstream of auxin.

(A) Root length of wild type (Col-0), pRPS5A-bHLH49, *bhlh49-1* and *bhlh49-2* seedlings upon treatment with 10 nM, 20 nM, 30 nM and 40 nM 2,4D, compared to untreated control (t test: $**P < 0,001$). Error bars indicate SD. (B-E) Expression of DR5-n3GFP in wild-type (B,D) and pRPS5A-bHLH49 (C, E) globular stage embryos (B,C) and roots (D,E). Scale bar represents 10 μ m in all panels.

bHLH49 homodimers mediate auxin-dependent suspensor regulation

bHLH proteins often form homo and/or heterodimers through the HLH domain in order to bind DNA (Littlewood and Evan, 1998; Toledo-Ortiz, 2003). To determine if bHLH49 acts in complex with other bHLH factors, we performed immunoprecipitation (IP)- mass spectrometry (MS) on lines carrying a pbHLH49-bHLH49-sYFP transgene. We performed two independent experiments on entire seedlings and root tips and in both cases measured triplicate bHLH49-YFP and wild-type samples, followed by quantification and statistical analysis. In both experiments, bHLH49 and YFP were recovered as the most abundant proteins, certifying the quality of the assay (Table S3). However, no unique peptides were found for bHLH proteins other than bHLH49 (Table S4) and since bHLH heterodimers were readily identified in similar experiments with LHW/bHLH156 and TMO5/bHLH032 (De Rybel et al., 2013), this could mean that bHLH49 might act as homodimer rather than as heterodimer.

To test this hypothesis, we used a Förster Resonance Energy Transfer (FRET) - Fluorescence Lifetime Imaging Microscopy (FLIM) assay, where we analyzed and quantified the interaction of cyan fluorescent protein (CFP)- and yellow fluorescent protein (YFP)-tagged bHLH49 in *Arabidopsis* leaf mesophyll protoplasts (Figure 6A-C). In this assay, protein-protein interaction is inferred from energy transfer from the CFP donor to the YFP acceptor molecule, and is measured as a decrease of fluorescence lifetime of the CFP donor (Bucherl et al., 2014). Co-expression of bHLH49-CFP and bHLH49-YFP caused a significant decrease in the lifetime of CFP fluorescence as compared to the co-expression of free YFP, whose effect was less pronounced (Figure 6D). Thus, we conclude that bHLH49 homodimerizes *in vivo* (Figure 6C,D). This is in agreement with the finding that misexpression of bHLH49 alone is able to induce strong developmental defects (Figure 4G), in contrast to e.g. TMO5 or LHW (De Rybel et al., 2013).

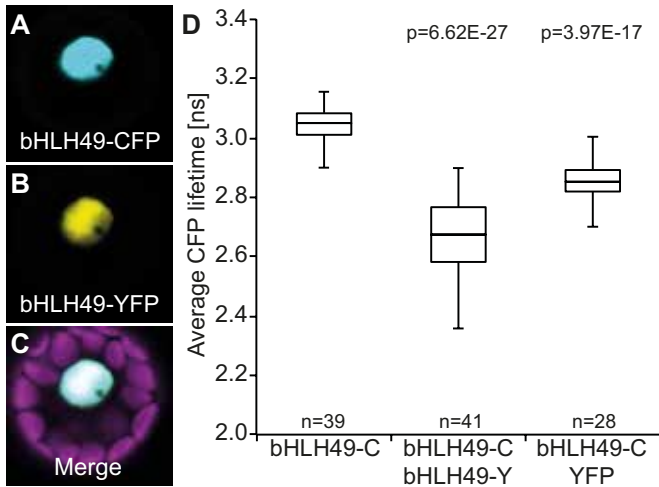


Figure 6: Homodimerization of bHLH49.

(A-C) Expression of bHLH49-CFP (A) and bHLH49-YFP (B) as well as combined signals with chloroplast autofluorescence (C) in protoplasts. (D) FRET-FLIM of bHLH49 in mesophyll protoplasts. The fluorescence lifetime (ns, nanoseconds) of the CFP-tagged bHLH49 protein is presented in boxplots (n is given in each column, p-value for t test is given in top of column). Free-YFP was used as negative control.

To determine if increased *bHLH49* expression, as induced by suppression of auxin response in suspensor cells, contributes to the gene expression program mediating embryo initiation, we subjected root tips of *bhlh49* mutant and RPS5A-bHLH49 overexpression lines to whole-genome transcript profiling. We found 93 genes to be differentially expressed when comparing *bhlh49* versus wild-type (Figure 7A), and 676 gene when comparing pRPS5A-bHLH49 versus wild-type (>1.4-fold

misregulated; FDR < 0.05). Based on general trends, bHLH49 appears to act as an activator, as more genes were upregulated in the overexpression line. We identified 8 genes that were oppositely regulated in overexpression and mutant (Figure 7A).

To validate the bHLH49-microarray data, we selected several genes and determined their expression level in *bhlh49* mutant and RPS5A-bHLH49 misexpression background using qRT-PCR. We selected genes that were downregulated in the *bhlh49* mutant and upregulated in pRPS5A-bHLH49 (*AT4G11650*, *AT2G33790*, and *AT2G46950*) and vice-versa (*AT3G21720* and *AT5G41300*). Our results showed that the expression levels of all the tested genes were in agreement with their misregulation in microarrays (Figure 7D, category I), confirming the validity of the dataset.

We next performed a meta-analysis to determine if there is a significant contribution of *bHLH49* upregulation to auxin-dependent suspensor-to-embryo transformation. This meta-analysis revealed a significant overlap between both sets of differentially expressed genes where 85 and 91 genes were up- and downregulated, respectively in both array experiments (Figure 7B,C). Interestingly, *bHLH60* and *bHLH100* were up- and downregulated respectively, in pRPS5A-bHLH49 roots. This recapitulates their misregulation in M0171 > > *bdl* embryos, which suggests that auxin-dependent regulation of *bHLH60* and *100* is mediated by *bHLH49*. The latter was also confirmed by qRT-PCR (Figure 7D, category II and III).

Finally, to determine the spatio-temporal extent of misregulation, we generated promoter-GFP lines for some of the genes found in the overlap between M0171 > > *bdl* and pRPS5A-bHLH49 datasets and subsequently crossed these with pRPS5A-bHLH49. The severe embryo phenotype of bHLH49 overexpression hampered this analysis, but results from embryonic stage were obtained for p*KMD1*-n3GFP x pRPS5A-bHLH49 (Figure 7E,F). Analysis of p*COBL1*-n3GFP activity in the misexpression line was limited to roots. The expression of *KMD1* was either completely or partially lost from the proliferating embryo-suspensor junction (Figure 7E,F; n = 13), coinciding with its downregulation in the microarray. The same counts for p*COBL1*-n3GFP, whose expression was vanished from the pRPS5A-bHLH49 roots in comparison with its expression in Col-0 (Figure 7G, H). Thus, bHLH49 appears to be a mediator of the auxin-dependent regulation of several genes in the suspensor and root.

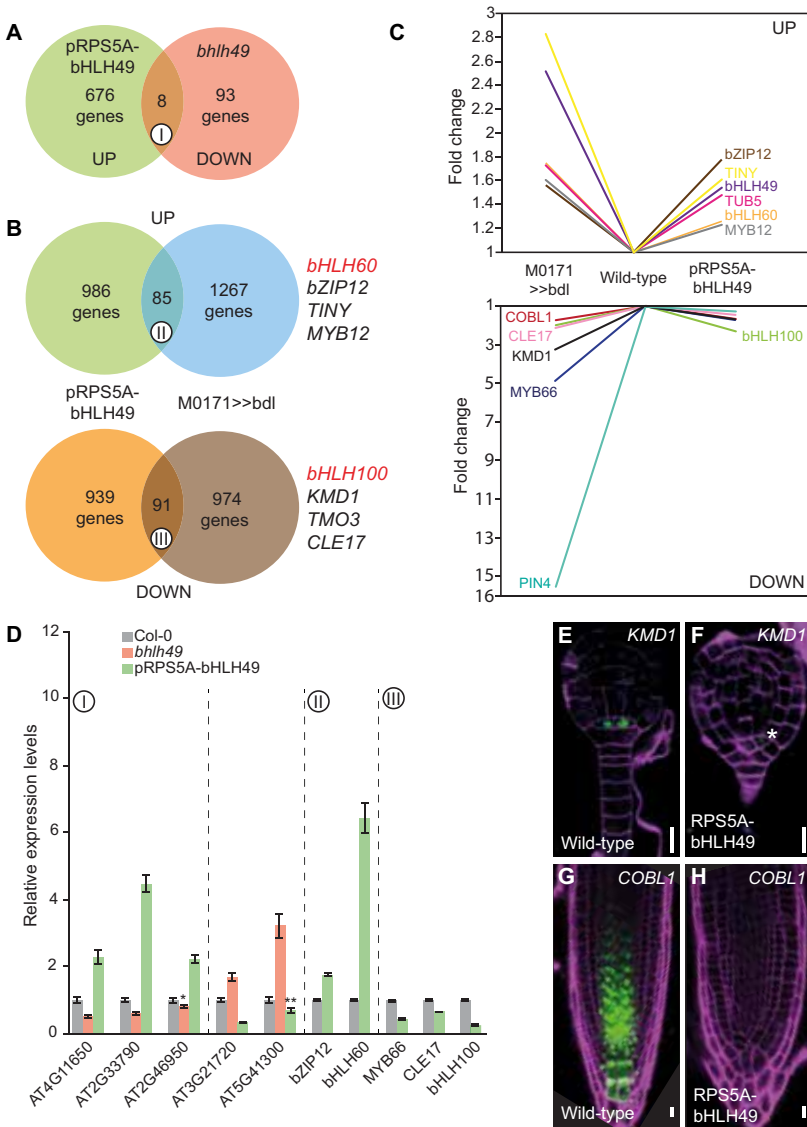


Figure 7: bHLH49 mediates auxin-dependent transcriptional regulation in plant development.

(A) Venn diagram depicting the genes upregulated in pRPS5A-bHLH49 and downregulated in *bhlh49* root tips. (B) Venn diagrams showing the overlap in genes either up-or down-regulated in pRPS5A-bHLH49 seedling roots and M0171 >> bdl embryos. Examples of genes present in the overlap are listed on the right. (C) Expression levels of selected genes in M0171 >> bdl embryos and pRPS5A-bHLH49 seedling roots in micro-array experiments. (D) qRT-PCR validation of expression levels of selected genes in wild-type (Col-0), *bhlh49* and pRPS5A-bHLH49 seedling roots. (E,F) Expression of pKMD1-n3GFP in wild-type (E) and pRPS5A-bHLH49 (F) embryos. (G,H) Expression of pCOBL1-n3GFP in wild-type (G) and pRPS5A-bHLH49 (H) seedling roots. Scale bar represents 10 μ m in all panels.

Discussion

The plant suspensor plays a supportive role during embryogenesis and is mitotically quiescent after an initial 3 rounds of cell division in Arabidopsis (Laux and Jürgens, 1997; Mayer and Jürgens, 1998). This pattern of suspensor development is representative for a large number of plant species, although significant divergence is observed in the plant kingdom (Yeung and Clutter, 1979; Yeung and Meinke, 1993). The quiescence of suspensor cells does not reflect its developmental potential. In several species, suspensor cells can be “reprogrammed” to form a second embryo (Yeung and Meinke, 1993). While mechanisms of conversion and its regulators are largely unknown, our work previously identified transcriptional auxin response as a key pathway suppressing embryo identity in suspensor cells (Rademacher et al., 2012). Here, we further exploited this system as a model for investigating suspensor-to-embryo transformation. Our study identified a number of genes whose misregulation upon auxin response inhibition correlates with embryo identity specification. Importantly, we identified a genetic network involving several bHLH transcription factors that mediates auxin action in controlling embryo identity.

Many genes are misregulated even before suspensor-specific *bdl/iaa12* expression causes morphological defects. This suggests that reprogramming is a complex genetic response, and it may be difficult to isolate the first events. Pro-embryo and suspensor are physically and symplastically connected and indeed many pro-embryo-specific genes are affected by auxin response inhibition in suspensor cells. Thus, we determined expression patterns for a large number of misregulated genes and identified several that are either a) normally expressed in pro-embryo cells and upregulated upon suspensor-specific auxin response inhibition or b) normally expressed in suspensor cells and downregulated. Our detailed investigation of a set of four bHLH genes that follow either pattern revealed that all are indeed regulated by auxin, in an ARF-dependent manner. Based on phenotypic data, as well as on the dependence of auxin-regulation on *de novo* translation, bHLH49 appeared a more direct and biologically more significant regulator. Indeed, when probing the bHLH49-dependent transcriptome, we found that the other bHLH genes are among the targets. Thus, the auxin-repressed bHLH49 gene appears to be an important mediator of auxin-dependent suppression of embryo identity in suspensor cells. Indeed, misexpression of bHLH49 alone could induce excess divisions and even multiple embryo-like structures in suspensors, similar to the effect of auxin response inhibition (Figure 4G).

An important open question is what primary cellular process bHLH49 targets to bring about S > E transformation. Post-embryonic misexpression of bHLH49 induces

2

very strong defects (Figure S3), but these do not include embryo induction. Thus, unlike “embryo inducers” such as LEC1 (Lotan et al., 1998) or BBM (Boutilier, 2002), bHLH49 does not appear to directly promote embryogenesis in a range of contexts, but rather does so in the specific context of the suspensor. Consistent with this notion, bHLH49 is broadly expressed, yet less so in suspensor cells. Given the lower level of expression in wild-type suspensor cells, bHLH49 action is likely dosage-sensitive, only triggering suspensor-derived embryogenesis above a certain threshold. The identity of the bHLH49-dependent genes does not directly reveal a key cellular process that can explain its ability to trigger suspensor-derived embryogenesis. While there is a significant overlap with the genes misregulated in M0171 >> bdl embryos, no gene families or functional categories are over-represented. Interestingly though, the overlap between M0171 >> bdl responsive and bHLH49-dependent genes was found despite the fact that one was carried out in embryos and the other in seedlings. Thus, it is well possible that bHLH49 regulates a rather generic cellular process that is able to trigger embryogenesis in the context of the suspensor. One such process could be cell division. However the post-embryonic misexpression phenotypes are not indicative of excess cell division. Furthermore, several plant species have many more suspensor cells divisions, yet these do not show features of embryo identity (Kawashima and Goldberg, 2010; Yeung and Clutter, 1979; Yeung and Meinke, 1993). Hence, unless Arabidopsis with its minimal number of suspensor cells is an exceptional situation, triggering cell division in suspensor cells may not be sufficient to induce embryogenesis.

This study revealed a novel auxin-dependent bHLH gene, repressed by auxin in an ARF-dependent manner that mediates auxin-dependent development. Previously, the bHLH genes *TMO5* and *TMO7* were found to be activated by auxin in an ARF5-dependent manner, and contribute to auxin-dependent embryonic root and vascular tissue development (De Rybel et al., 2013; Schlereth et al., 2010). Although the bHLH family has about 170 members in Arabidopsis (Bailey et al., 2003; Carretero-Paulet et al., 2010), and many processes are under auxin control, it is striking that multiple auxin responses in embryo development are mediated by phylogenetically unrelated bHLH genes. Based on this finding, it is tempting to speculate that, as auxin responses and the ARF transcription factor family have evolved complexity, ancestral ARF-bHLH regulatory connections have been maintained. As such, these may represent a module common to auxin-controlled development. Transcriptomic analysis of further auxin-dependent developmental processes should help to resolve whether auxin action indeed involves a core regulatory output module.

Materials and Methods

Plant Material

All plants used in this study were Columbia-0 (Col-0) ecotype except M0171-GAL4 enhancer trap line, which was in the C24 background (made available by Dr. Jim Haseloff, Cambridge, UK). T-DNA insertion lines *bhlh49-1* (SALK_135188C), *bhlh49-2* (SALK_087424C), *bhlh60-3* (SAIL_1219_E01), *bhlh63-2* (SAIL_1211_F11), *bhlh100-1* (SALK_150637C), *bhlh100-2* (SALK_074568C), *arf1-5* (SALK_079046; (Ellis et al., 2005)), *arf2-8* (SALK_108995; Ellis et al., 2006), *arf6-1* (CS24606; (Okushima et al., 2005)), *arf9-1* (SAIL_881_H05; Okushima et al., 2005), *arf13-2* (SALK_138188), *arf18-3* (GABI_699B09) and M0171-GAL4 (Rademacher et al., 2012) enhancer trap line were obtained from *Arabidopsis* Stock Centers (NASC-ARBC) and genotyped using the primers listed in Supplemental Dataset 2.

All seeds were sterilized in 25% bleach/75% ethanol solution for 10 minutes and were afterwards washed twice with 70% ethanol and once with 100% ethanol. Dried seeds were plated on half-strength Murashige and Skoog (MS) medium and the appropriate antibiotic (in concentration: 50 mg/l kanamycin or 15 mg/l phosphinothricin) for selection of transgenic seeds. After 24 hours incubation at 4°C, the plants were cultured under long-day conditions at 22°C. Plant transformation was carried out by floral dipping, as described (De Rybel et al., 2011).

Cloning

All cloning was performed using the LIC cloning system and the vectors described (De Rybel et al., 2011). For transcriptional fusions, fragments up to 3 kb upstream of the ATG including 5'-UTR were amplified from genomic DNA using Phusion Flash polymerase (Thermo Scientific). For translational fusions of bHLH genes, the same promoter fragments were amplified along with the genomic coding sequencing excluding the stop codon. In order to generate pRPS5A-driven misexpression, the coding sequences were amplified from complementary DNA (cDNA) clones. All constructs were completely sequenced. The primers used for cloning are listed in Supplemental Dataset 2. At least three independent lines of each construct were analyzed.

Microscopy

Differential interference contrast (DIC) and confocal microscopy were performed as described previously (Llavata-Peris et al., 2013). Cleared embryos were observed

under Leica DMR microscope equipped with DIC optics and confocal imaging was performed on Leica SP5 II system (HyD detector). Cell outlines were generated by counterstaining with SCRI Renaissance Stain 2200 (R2200; Renaissance Chemicals, UK).

Quantitative RT-PCR Analysis

qRT-PCR analysis was performed as described previously (De Rybel et al., 2010). RNA was isolated using TRIzol reagent (Invitrogen) and RNeasy kit (Qiagen). cDNA was prepared from 0.5 µg of total RNA with an iScript cDNA Synthesis Kit (BioRad). qRT-PCR reactions were performed with iQ SYBR Green Supermix (BioRad) and analyzed on a CFX384 Real-Time PCR detection system (BioRad). Reactions were done in triplicate with three biological replicates. Data were analyzed with qBase as described in (Hellemans et al., 2007). Primers were designed with Beacon Designer 8 (Premier Biosoft International). Gene expression levels were normalized relative to *CDKA1;1*, *EEFα4* and *GAPC*. Primers for qRT-PCR are listed in Supplemental Dataset 2.

Auxin Sensitivity Assay

The auxin sensitivity assay was performed according to (Lincoln et al., 1990). Seeds were first germinated on standard half-strength MS medium. Six-day old seedlings were transferred to fresh MS medium supplemented with 10nM, 20nM, 30nM and 40nM 2,4D or lacking 2,4D. After two days, plates were scanned and the length of the newly grown roots was measured using ImageJ software. The percentage of the root growth was then calculated relative to the root growth on MS without auxin.

IP-MS

Immunoprecipitation experiments were carried out as described previously (Zwiewka et al., 2011), where 3g of root tips and/or seedlings of pbHLH49-bHLH49-sYFP transgenic lines in Col-0 background were used. In order to isolate interacting proteins, anti-GFP coupled magnetic beads (Milteny Biotech) were added to the total protein extracts. As described previously (De Rybel et al., 2013; Hubner et al., 2010; Lu et al., 2011), samples were run on nLC-MS/MS and data were analyzed using MaxQuant and Perseus software.

FRET-FLIM

FRET-FLIM analysis in Arabidopsis leaf mesophyll protoplasts was performed as described previously (Bücherl et al., 2013; De Rybel et al., 2013) with minor modifications. Measurements were conducted on a Leica TCS SP5 X system equipped with a 63x/1.20 numeric aperture water-immersion objective lens was used for confocal/FLIM imaging. Confocal and FLIM images were acquired by exciting the respective fluorophores sCFP3A/ sYFP2 (Kremers et al., 2006), using a diode laser (440 nm, pulse frequency 40 MHz) or the 514-nm line of an argon laser was used, respectively. Confocal and FLIM imaging was performed using internal filter-free spectral Hybrid detectors, which have high quantum efficiency and can be used for photon counting detection. The spectral window for sCFP3A and sYFP2 was set in the range of 450 to 500/520 to 560 nm, respectively. Confocal images were acquired with 512 x 512 pixels, while FLIM was acquired using frame size of 128 x 128 pixels. From the fluorescence intensity images, the decay curves were calculated per pixel and fitted with either a mono- or double-exponential decay model using the SPCImage software (Becker & Hickl; version 3.2.3.0). The mono-exponential model function was applied for donor samples with only sCFP3A present. For samples containing two fluorophores, sCFP3A/sYFP2, a double-exponential model function was used without fixing any parameter. All the cloning for this experiment was done in pMON999-sCFP and pMON999-sYFP vectors modified for LIC (De Rybel et al., 2011). The primers used are listed in Supplemental Dataset 2.

Microarray experiments

M0171 > > bdl: After crossing, embryos were isolated in a 5% Sucrose solution, containing 0.1% RNALater as described previously in (Xiang et al., 2011), and the isolated embryos were pooled in a 1.5 mL eppendorf tube on dry ice (300-400 embryos for each biological replicate). Total RNA was extracted according to the protocol of RNAqueous-micro kit and amplified prior labeling following the protocol provided in the MessageAmp aRNA kit with minor modification.

The Arabidopsis 70-mer oligo array slides prepared by University of Arizona were used in all the microarray experiments (version ATV 3.7.2; <http://ag.arizona.edu/microarray>). Antisense RNA was labeled according to the protocol of Wellmer et al, 2004. The aRNA samples representing four biological replicates from experimental and control samples were labeled (two with Cy3 and two with Cy5) and hybridized to the slides following the protocol from <http://eg.arizona.edu/microarray>. Subsequently, the hybridized slides were scanned for Cy3- and Cy5-labeled mRNA

targets with ScanArray 4000 laser scanner (at a resolution of 10 μm . The QuantArray program (GSI Lumonics) was used for the image analysis and signal quantification.

Limma Software (Smyth, 2004) was used for normalization and determination of the modulated genes from the microarray data.

pRPS5A-bHLH49/*bhlh49*: RNA from root tips was isolated as described above (see Quantitative RT-PCR) and total RNA (100ng) was labeled using Ambion WT Expression kit (Life Technologies). The RNA was then hybridized to Arabidopsis gene ST arrays (Affimetrix), which probes the expression of 27,827 unique genes. Sample labeling and hybridization was carried out according to manufacturer's instruction. Microarray analysis was performed as previously described (De Rybel et al., 2014).

Accession Numbers

All microarray data have been deposited in the Gene Expression Omnibus (GEO) at the NCBI (M0171 > > bdl: GSE69854; bHLH49 mutant and overexpression: GSE69700).

Acknowledgements

The authors thank Dr. Jim Haseloff and the Nottingham Arabidopsis Stock Center for distributing seeds, Sjef Boeren for support with Mass Spectrometry and Bert De Rybel and Maritza van Dop for comments on the manuscript. This work was funded by grants from the Netherlands Organization for Scientific Research (NWO; ALW-NSFC grant 846.11.001; ALW Open Competition grant 816.02.014) and the European Union ITN network SIREN; contract number 214788) to D.W.

Author Contributions

T.R., A.S.L., C.I.L.-P. and D.W. designed the research; T.R., A.S.L., C.I.L.-P., J.R.W., D.X., C.-Y.L., and L.V. performed research; all authors contributed to data analysis; T.R. and D.W. wrote the paper with contributions from A.S.L. and C.I.L.-P. and with input from all other authors.

References

- Andriankaja, M.E., Danisman, S., Mignolet-Spruyt, L.F., Claeys, H., Kochanke, I., Vermeersch, M., De Milde, L., De Bodt, S., Storme, V., Skiryicz, A., Maurer, F., Bauer, P., Mühlenbock, P., Van Breusegem, F., Angenent, G.C., Immink, R.G. and Inzé, D. (2014). Transcriptional coordination between leaf cell differentiation and chloroplast development established by TCP20 and the subgroup Ib bHLH transcription factors. *Plant Mol Biol* 85, 233-245.
- Bailey, P.C., Martin, C., Toledo-Ortiz, G., Quail, P.H., Huq, E., Heim, M.A., Jakoby, M., Werber, M. and Weisshaar, B. (2003). Update on the Basic Helix-Loop-Helix Transcription Factor Gene Family in *Arabidopsis thaliana*. *Plant Cell* 15, 2497-2502.
- Benjamins, R. and Scheres, B. (2008). Auxin: the looping star in plant development. *Annu Rev Plant Biol* 59, 443-465.
- Boutillier, K., Offringa, R., Sharma, V.K., Kirft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.M., van Lammeren, A.A., Miki, B.L., Custers, J.B. and van Lookeren Campagne, M.M. (2002). Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. *Plant Cell* 14, 1737-1749.
- Bucherl, C.A., Bader, A., Westphal, A.H., Laptinok, S.P. and Borst, J.W. (2014). FRET-FLIM applications in plant systems. *Protoplasma* 251, 383-394.
- Carretero-Paulet, L., Galstyan, A., Roig-Villanova, I., Martinez-Garcia, J.F., Bilbao-Castro, J.R. and Robertson, D.L. (2010). Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in *Arabidopsis*, poplar, rice, moss, and algae. *Plant Physiol* 153, 1398-1412.
- Chandler, J.W., Cole, M., Flier, A. and Werr, W. (2009). BIM1, a bHLH protein involved in brassinosteroid signalling, controls *Arabidopsis* embryonic patterning via interaction with DORNROSCHEN and DORNROSCHEN-LIKE. *Plant Mol Biol* 69, 57-68.
- Crawford, B.C., Sewell, J., Golembeski, G., Roshan, C., Long, J.A. and Yanofsky, M.F. (2015). Plant development. Genetic control of distal stem cell fate within root and embryonic meristems. *Science* 347, 655-659.
- De Rybel, B., Adibi, M., Breda, A.S., Wendrich, J.R., Smit, M.E., Novak, O., Yamaguchi, N., Yoshida, S., Van Isterdael, G., Palovaara, J., Nijssse, B., Boekschoten, M.V., Hooiveld, G., Beeckman, T., Wagner, D., Ljung, K., Fleck, C. and Weijers, D. (2014). Plant development. Integration of growth and patterning during vascular tissue formation in *Arabidopsis*. *Science* 345, 1255-1261.
- De Rybel, B., Moller, B., Yoshida, S., Grabowicz, I., Barbier de Reuille, P., 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*. *Dev Cell* 24, 426-437.
- De Rybel, B., van den Berg, W., Lokerse, A., Liao, C.Y., van Mourik, H., Moller, B., Peris, C.L. and Weijers, D. (2011). A versatile set of ligation-independent cloning vectors for functional studies in plants. *Plant Physiol* 156, 1292-1299.
- De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W., Audenaert, D., Van Campenhout, J., Overvoorde, P., Jansen, L., Vanneste, S., 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. *Curr Biol* 20, 1697-1706.
- Dharmasiri, N., Dharmasiri, S., Weijers, D., Lechner, E., Yamada, M., Hobbie, L., Ehrismann, J.S., Jurgens, G. and Estelle, M. (2005). Plant development is regulated by a family of auxin receptor F box proteins. *Dev Cell* 9, 109-119.

- Ellis, C.M., Nagpal, P., Young, J.C., Hagen, G., Guilfoyle, T.J. and Reed, J.W. (2005). AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development* 132, 4563-4574.
- Feller, A., Machemer, K., Braun, E.L. and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J* 66, 94-116.
- 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.
- Franco, A.R., Gee, M.A. and Guilfoyle, T.J. (1990). Induction and superinduction of auxin-responsive mRNAs with auxin and protein synthesis inhibitors. *J Biol Chem* 265, 15845-15849.
- Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R. and Jurgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. *Nature* 426, 147-153.
- Gaj, M.D., Zhang, S., Harada, J.J. and Lemaux, P.G. (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of *Arabidopsis*. *Planta* 222, 977-988.
- Haccius, B. (1955). Experimentally Induced Twinning in Plants. *Nature* 176.
- Hamann, T., Benkova, E., Baurle, I., Kientz, M. and Jurgens, G. (2002). The *Arabidopsis* BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes Dev* 16, 1610-1615.
- Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 8, R19.
- Hubner, N.C., Bird, A.W., Cox, J., Spletstoesser, B., Bandilla, P., Poser, I., Hyman, A. and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J Cell Biol* 189, 739-754.
- Jones, B., Gunneras, S.A., Petersson, S.V., Tarkowski, P., Graham, N., May, S., Dolezal, K., Sandberg, G. and Ljung, K. (2010). Cytokinin regulation of auxin synthesis in *Arabidopsis* involves a homeostatic feedback loop regulated via auxin and cytokinin signal transduction. *Plant Cell* 22, 2956-2969.
- Kawashima, T. and Goldberg, R.B. (2010). The suspensor: not just suspending the embryo. *Trends Plant Sci* 15, 23-30.
- Kwong, R.W., Bui, A.Q., Lee, H., Kwong, L.W., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2003). LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. *Plant Cell* 15, 5-18.
- Lakshmanan, K.K. and Ambegaokar, K.B. (1984). Polyembryony. In *Embryology of Angiosperms*, B. Johri, ed. (Springer Berlin Heidelberg), pp. 445-474.
- Laux, T. and Jurgens, G. (1997). Embryogenesis: A New Start in Life. *Plant Cell* 9, 989-1000.
- Leyser, O. (2010). The power of auxin in plants. *Plant Physiol* 154, 501-505.
- Lincoln, C., Britton, J.H. and Estelle, M. (1990). Growth and development of the *axr1* mutants of *Arabidopsis*. *Plant Cell* 2, 1071-1080.
- Littlewood, T.D. and Evan, G.I. (1998). Basic helix-loop-helix transcription factors. Oxford University Press, Oxford.
- Liu, H., Yu, X., Li, K., Klejnot, J., Yang, H., Lisiero, D. and Lin, C. (2008). Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in *Arabidopsis*. *Science* 322, 1535-1539.
- Llavata-Peris, C., Lokerse, A., Moller, B., De Rybel, B. and Weijers, D. (2013). Imaging of phenotypes, gene expression, and protein localization during embryonic root formation in *Arabidopsis*. *Methods Mol Biol* 959, 137-148.
- Lokerse, A.S. and Weijers, D. (2009). Auxin enters the matrix--assembly of response machineries for specific outputs. *Curr Opin Plant Biol* 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.
- Lotan, T., Ohto, M., Yee, K.M., West, M.A., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93, 1195-1205.
- Lu, J., Boeren, S., de Vries, S.C., van Valenberg, H.J., Vervoort, J. and Hettinga, K. (2011). Filter-aided sample preparation with dimethyl labeling to identify and quantify milk fat globule membrane proteins. *J Proteomics* 75, 34-43.
- Mayer, U. and Jürgens, G. (1998). Pattern formation in plant embryogenesis: A reassessment. *Semin Cell Dev Biol* 9, 187-193.
- Murre, C., Bain, G., van Dijk, M.A., Engel, I., Furnari, B.A., Massari, M.E., Matthews, J.R., Quong, M.W., Rivera, R.R. and Stuiver, M.H. (1994). Structure and function of helix-loop-helix proteins. *Biochim Biophys Acta* 1218, 129-135.
- Okushima, Y., Fukaki, H., Onoda, M., Theologis, A. and Tasaka, M. (2007). ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in Arabidopsis. *Plant Cell* 19, 118-130.
- 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. *Plant Cell* 17, 444-463.
- Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Freire Rios, A., Borst, J.W., Lukowitz, W., Jurgens, G. and Weijers, D. (2012). Different auxin response machineries control distinct cell fates in the early plant embryo. *Dev Cell* 22, 211-222.
- Rademacher, E.H., Moller, B., Lokerse, A.S., Llavata-Peris, C.I., van den Berg, W. and Weijers, D. (2011). A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. *Plant J* 68, 597-606.
- Radoeva, T. and Weijers, D. (2014). A roadmap to embryo identity in plants. *Trends Plant Sci* 19, 709-716.
- Raghavan, V. (2006). Life and times of the suspensor - cell signaling between the embryo and suspensor. *Double Fertilization*, 81-100.
- Schlereth, A., Moller, B., Liu, W., Kientz, M., Flipse, J., Rademacher, E.H., Schmid, M., Jurgens, G. and Weijers, D. (2010). MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. *Nature* 464, 913-916.
- Schwartz, B.W., Vernon, D.M. and Meinke, D. (1997). Development of the suspensor: differentiation, communication, and programmed cell death during plant embryogenesis. *Advances in Cellular and Molecular Biology of Plants Volume 4*, 53-72.
- Schwarz, B.W., Yeung, E.C. and Meinke, D.W. (1994). Disruption of morphogenesis and transformation of the suspensor in abnormal suspensor mutants of Arabidopsis. *Development*, 3235-3245.
- Sivitz, A.B., Hermand, V., Curie, C. and Vert, G. (2012). Arabidopsis bHLH100 and bHLH101 control iron homeostasis via a FIT-independent pathway. *PLoS One* 7, e44843.
- Toledo-Ortiz, G. and Huq, E. and Quail, P.H. (2003). The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family. *Plant Cell* 15, 1749-1770.
- Tsakagoshi, H., Busch, W. and Benfey, P.N. (2010). Transcriptional regulation of ROS controls transition from proliferation to differentiation in the root. *Cell* 143, 606-616.
- 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. *Plant Cell* 9, 1963-1971.

- Vernon, D. and Meinke, D. (1994). Embryogenic transformation of the suspensor in twin, a polyembryonic mutant of *Arabidopsis*. *Dev Biol* 165, 566-573.
- Wang, H.Y., Klatte, M., Jakoby, M., Baumlein, H., Weisshaar, B. and Bauer, P. (2007). Iron deficiency-mediated stress regulation of four subgroup Ib BHLH genes in *Arabidopsis thaliana*. *Planta* 226, 897-908.
- Wang, R. and Estelle, M. (2014). Diversity and specificity: auxin perception and signaling through the TIR1/AFB pathway. *Curr Opin Plant Biol* 21, 51-58.
- Weijers, D., Franke-van Dijk, M., Vencken, R.J., Quint, A., Hooykaas, P. and Offringa, R. (2001). An *Arabidopsis* Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. *Development* 128, 4289-4299.
- Weijers, D., Schlereth, A., Ehrismann, J.S., Schwank, G., Kientz, M. and Jurgens, G. (2006). Auxin triggers transient local signaling for cell specification in *Arabidopsis* embryogenesis. *Dev Cell* 10, 265-270.
- Weijers, D., Van Hamburg, J.P., Van Rijn, E., Hooykaas, P.J. and Offringa, R. (2003). Diphtheria toxin-mediated cell ablation reveals interregional communication during *Arabidopsis* seed development. *Plant Physiol* 133, 1882-1892.
- Xiang, D., Venglat, P., Tibiche, C., Yang, H., Risseeuw, E., Cao, Y., Babic, V., Cloutier, M., Keller, W., Wang, E., Selvaraj, G. and Datla, R. (2011). Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in *Arabidopsis*. *Plant Physiol* 156, 346-356.
- Yamaguchi, N., Wu, M.F., Winter, C.M., Berns, M.C., Nole-Wilson, S., Yamaguchi, A., Coupland, G., Krizek, B.A. and Wagner, D. (2013). A molecular framework for auxin-mediated initiation of flower primordia. *Dev Cell* 24, 271-282.
- Yeung, E.C. and Clutter, M.E. (1979). Embryogeny of *Phaseolus coccineus*: the ultrastructure and development of the suspensor. *Can J Bot* 57, 120-136.
- Yeung, E.C. and Meinke, D.W. (1993). Embryogenesis in Angiosperms: Development of the Suspensor. *Plant Cell* 5, 1371-1381.
- Zhao, Z., Andersen, S.U., Ljung, K., Dolezal, K., Miotk, A., Schultheiss, S.J. and Lohmann, J.U. (2010). Hormonal control of the shoot stem-cell niche. *Nature* 465, 1089-1092.
- Zwiewka, M., Feraru, E., Moller, B., Hwang, I., Feraru, M.I., Kleine-Vehn, J., Weijers, D. and Friml, J. (2011). The AP-3 adaptor complex is required for vacuolar function in *Arabidopsis*. *Cell Res* 21, 1711-1722.

Supplementary information

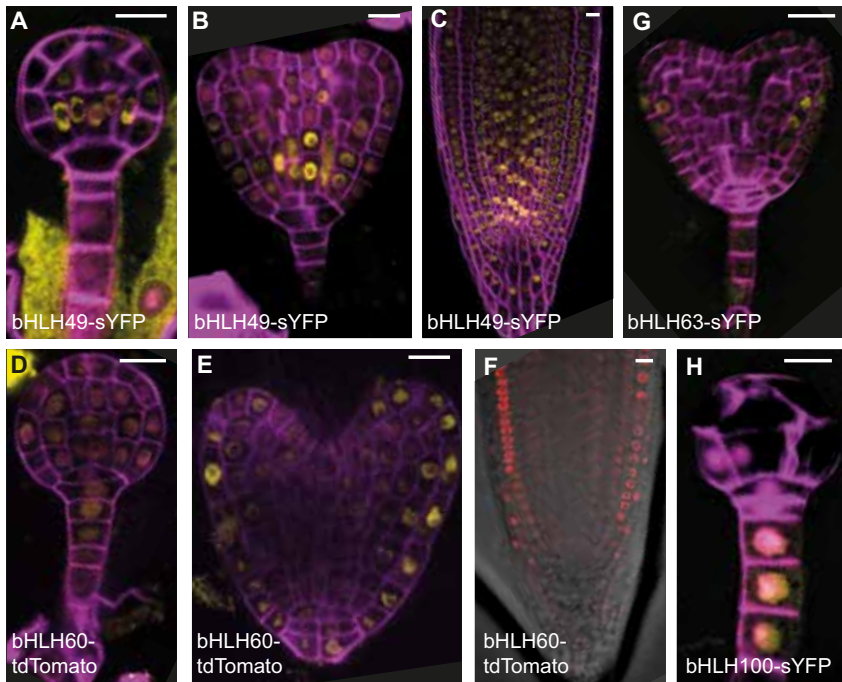


Figure S1. Localization of bHLH proteins.

(A-C) Translational fusion domain of bHLH49, (D-F) bHLH60, (G) bHLH63 and (H) bHLH100 in globular stage (A, D, H), heart stage (B, E, G) embryos and in roots (C, F). Scale bar indicates 10 μm in all panels.

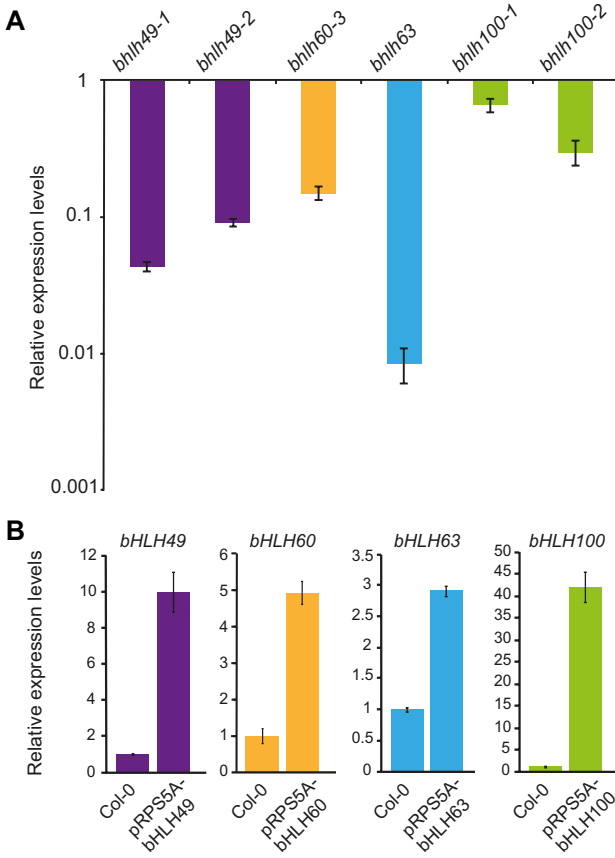


Figure S2: Relative expression of bHLH genes.

(A) Expression of *bHLH49*, *60*, *63* and *100* in *bhlh* mutants and (B) in overexpression lines.

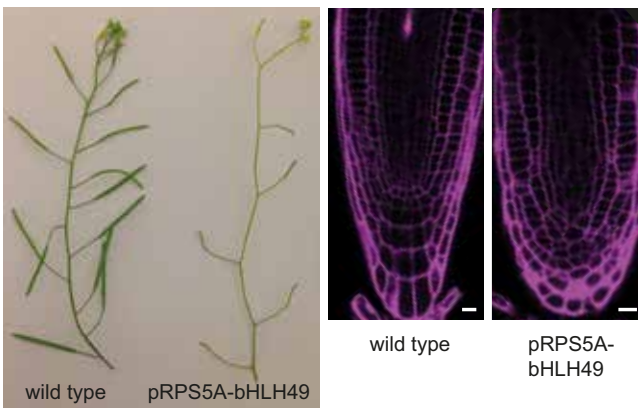


Figure S3: Post-embryonic phenotype of pRPS5A-bHLH49.

Table S1: Primary auxin responsive genes misregulated in the M0171 > > bdl dataset.

AGI = Arabidopsis Genome Initiative identification number. Other name(s) = abbreviated common names for genes. FC = fold change in M0171 > > bdl dataset.

Gene Family	AGI	Other name(s)	FC
AUX/IAA	AT3G15540	IAA19	8.1 down
AUX/IAA	AT3G16500	IAA26	4.8 down
AUX/IAA	AT3G62100	IAA30	3.5 down
AUX/IAA	AT2G46990	IAA20	1.7 down
AUX/IAA	AT1G04550	IAA12	10.6 up
AUX/IAA	AT1G04250	IAA17	3.2 up
GH3	AT1G28130	GH3,17	2.9 down
GH3	AT4G27260	GH3.5, WES1	2.3 down
GH3	AT2G23170	GH3,3	2.0 down
GH3	AT1G59500	GH3,4	1.7 down
SAUR	AT5G27780	-	1.9 down
SAUR	AT1G19840	-	1.7 down
SAUR	AT5G10990	-	1.7 down
SAUR	AT3G20220	-	2.9 up
SAUR	AT1G79130	-	1.7 up

Table S2: Sixty-eight genes selected for validation of M0171 > > bdl dataset.

AGI	Other Name(s)	Short Gene Model Description	Fold Change	Criteria	Expression pattern during embryogenesis
AT1G22590	AGL87	MADS box transcription factor	3.1 up	TF	Suspensor and protoderm
AT1G48910	YUC10	Auxin biosynthesis enzyme	3.5 up	Auxin	no expression detected
AT1G61566	RALFL9	Member of predicted peptide family	2.6 up	Zygote	no expression detected
AT1G68320	MYB62	R2R3-MYB transcription family	3.0 up	TF	suspensor and basal tier of proembryo
AT1G68510	LBD42	LBD transcription factor	2.8 up	Auxin, TF	no expression detected
AT1G68920		basic helix-loop-helix (bHLH) superfamily protein	2.5 up	Auxin, TF	inner cells of proembryo
AT1G76500	SOB3/AHL29	AT hook domain containing protein	2.3 up		no expression detected
AT2G12900		Basic-leucine zipper (bZIP) transcription factor	2.0 up	TF	suspensor and protoderm
AT2G21420		IBR domain containing protein	2.8 up		no expression detected
AT2G26850		F-box family protein	3.4 up	Zygote	no expression detected
AT2G27250	CLV3	One of the three CLAVATA genes controlling SAM size	44.8 up	Top 10%	no expression detected
AT2G35310		Transcriptional factor B3 family protein	2.4 up	TF	apical tier and SAM area of proembryo
AT3G03770		Leucine-rich repeat protein kinase family protein	2.1 up		no expression detected
AT3G25170	RALFL26	Member of predicted peptide family	4.0 up	Zygote	no expression detected
AT3G54320	WR11	Transcription factor of the AP2/ERWEBP class	4.2 up	TF	junction apical and basal tier
AT3G54800		PH and lipid-binding START domain containing protein	24.4 up	Top 10%	no expression detected
AT3G57800	bHLH060	basic helix-loop-helix (bHLH) superfamily protein	2.4 up	Auxin, TF, bHLH	protoderm of proembryo
AT3G61150	HDG1	HD-ZIP IV family homeobox-leucine zipper family protein	2.4 up	Auxin, TF	hypophysis descendents
AT3G61160		Protein kinase superfamily protein	4.1 up	Zygote	no expression detected

AT4G28650	Leucine-rich repeat transmembrane protein kinase family 3.3	up	3.3	suspensor and protoderm
AT4G32540	YUC1 Auxin biosynthesis enzyme	5.9	up	Auxin, Top 10% apical protoderm, SAM area of proembryo
AT4G34530	bHLH063/CIB1 basic helix-loop-helix (bHLH) superfamily protein	9.9	up	Auxin, TF, bHLH, junction apical and basal tier Top 10%
AT5G06250	AP2/B3-like transcriptional factor family protein	3.8	up	TF sam area of proembryo
AT5G11320	YUC4 Auxin biosynthesis enzyme	7.9	up	Auxin suspensor and apical protoderm
AT5G11530	EMF1 Involved in regulating reproductive development	2.2	up	Known function no expression detected
AT5G25830	GATA12 GATA factor family of zinc finger transcription factors	6.3	up	TF, Top 10% suspensor
AT5G26650	AGL36 DNA binding transcription factor	2.7	up	TF no expression detected
AT5G26950	AGL93 DNA binding transcription factor	2.5	up	TF hypophysis descendents and protoderm
AT5G39860	bHLH136/PRE1 basic helix-loop-helix (bHLH) superfamily protein	5.2	up	Auxin, TF, bHLH, no expression detected Top 10%
AT5G47670	LEC1-LIKE Functions as a regulator of embryo development.	2.5	up	TF no expression detected
AT5G52600	MYB82 Member of the R2R3 factor gene family.	2.0	up	TF inner cells of proembryo
AT5G56270	WRKY2 WRKY Transcription factor	2.5	up	Auxin, TF suspensor
AT5G57670	Protein kinase superfamily protein	2.2	up	no expression detected
AT3G05800	bHLH150/AIF1 basic helix-loop-helix (bHLH) superfamily protein	2.9	up	Auxin, TF, bHLH no expression detected
AT3G28857	bHLH164/PRE5 basic helix-loop-helix (bHLH) superfamily protein	2.4	up	Auxin, TF, bHLH no expression detected
AT1G13960	WRKY4 WRKY DNA-binding protein	2.2	down	Auxin, TF no expression detected
AT1G14600	Homeodomain-like superfamily protein	2.5	down	TF apical protoderm and SAM area of proembryo
AT1G15670	Galactose oxidase/kelch repeat superfamily protein	3.1	down	Top 10% suspensor and descendents

AT1G31320	LBD4	LBD transcription factor	2.8 down	Auxin, TF	suspensor and basal tier of proembryo
AT1G62000		unknown protein	2.4 down		no expression detected
AT1G62220		unknown protein	2.2 down		no expression detected
AT1G80440		Galactose oxidase/kelch repeat superfamily protein	3.5 down	Top 10%	apical cells following hypophysis division
AT1G80640		Protein kinase superfamily protein	2.3 down		inner cells of proembryo
AT2G18300	bHLH064	basic helix-loop-helix (bHLH) superfamily protein	2.2 down	Auxin, TF, bHLH	no expression detected
AT2G18890		Protein kinase superfamily protein	3.7 down	Top 10%	inner cells of proembryo
AT2G23050	NPY4	Involved in auxin-mediated organogenesis	3.2 down	Auxin	dynamic, embryo proper and qc
AT2G41170		F-box family protein	2.3 down		no expression detected
AT2G45430	AHL22	AT hook domain containing protein	2.3 down		no expression detected
AT3G13960	GRF5	Growth regulating factor encoding transcription activator	25.4 down	TF, Top 10%	proembryo
AT3G54780		Zinc finger (C3HC4-type RING finger) family protein	4.8 down	Top 10%	no expression detected
AT3G62100	IAA30	Encodes a member of the Aux/IAA family	3.4 down	Auxin, TF	initially suspensor specific, expanding to prevasculature
AT4G04090		BTB/POZ domain-containing protein	2.8 down		suspensor and hypophysis
AT4G10270		Wound-responsive family protein	2.2 down		protoderm
AT4G18740		Rho termination factor	2.1 down		suspensor
AT4G24390		RNI-like superfamily protein	2.1 down		initially suspensor specific, expanding to prevasculature
AT4G31820	ENP/NPY1	member of the NPY family genes	5.9 down	Auxin, Top 10%	lower tier of proembryo
AT4G36240	GATA7	GATA factor family of zinc finger transcription factors	2.0 down	TF	no expression detected

AT5G18270	NAC087	Arabidopsis NAC domain containing protein	3.6 down	TF, Top 10%	protoderm
AT5G48940		Leucine-rich repeat transmembrane protein kinase	3.0 down	Top 10%	basal tier of proembryo
AT5G49490	AGL83	MADS-box transcription factor family protein	2.2 down	TF	not cloned, no analysis performed.
AT5G50915	bHLH137	basic helix-loop-helix (bHLH) superfamily protein	4.6 down	Auxin, TF, bHLH, Top 10%	protoderm of basal tier of proembryo
AT5G66560		Phototropic-responsive NPH3 family protein	2.2 down		same area of proembryo
AT1G29950	bHLH144	basic helix-loop-helix (bHLH) superfamily protein	3.5 down	Auxin, TF, bHLH, Top 10%	no expression detected
AT1G68810	bHLH030	basic helix-loop-helix (bHLH) superfamily protein	2.7 down	Auxin, TF, bHLH	not analysed
AT2G01420	PIN4	Auxin efflux carrier	15.5 down	Auxin	Initially suspensor, later also in lower tier of proembryo
AT1G05710	bHLH153	basic helix-loop-helix (bHLH) superfamily protein	2.2 down	Auxin, TF, bHLH	inner cells of proembryo
AT2G41240	bHLH100	basic helix-loop-helix (bHLH) superfamily protein	2.0 down	Auxin, TF, bHLH	suspensor
AT2G41130	bHLH106	basic helix-loop-helix (bHLH) superfamily protein	2.2 down	Auxin, TF, bHLH	not analysed
AT1G73590	PIN1	Auxin efflux carrier	4.9 down	Auxin	apical tier of proembryo
AT2G42870	bHLH165/PAR1	basic helix-loop-helix (bHLH) superfamily protein	4.2 down	Auxin, TF, bHLH, Top 10%	basal cells following hypophysis division

AGI = Arabidopsis Genome Initiative gene identification number. Other name(s) = abbreviated common names for genes. FC = fold change in M0171 > > bdl dataset. TF, Transcription Factor; Auxin, auxin-regulated; Zygote, zygote-enriched; Top 10%, among top 10% most strongly misexpressed genes; bHLH, member of bHLH family.

Table S3: Immunoprecipitation (IP)- mass spectrometry (MS) on lines carrying a pbHLH49-bHLH49-sYFP transgene.

IP-MS on seedlings			
Protein IDs	Description	RATIO	p-value
Q9CAA9;Q9LK48	Transcription factor bHLH49	389.971	0.001
CON_sp P42212mut3 eGFP	Green Fluorescent Protein	226.136	0.018
F4K4Y5	DEK domain-containing chromatin associated protein	71.957	0.000
P21240;Q9LJE4;C0Z361;Q9C667	Chaperonin 60 subunit beta 1/2/3, chloroplastic;	32.988	0.010
P25857	Glyceraldehyde-3-phosphate dehydrogenase B, chloroplastic	19.116	0.000
Q9FE58	60S ribosomal protein L22-3	18.192	0.423
Q9SN19	unknown protein	15.363	0.002
Q9LTX9	Heat shock 70 kDa protein 7, chloroplastic	14.404	0.006
P19366	ATP synthase subunit beta, chloroplastic	11.917	0.013
IP-MS on root tips			
Protein IDs	Description	RATIO	p-value
Q9CAA9;C0Z2X3;Q6NKN9;Q9C670; Q9LK48;Q9FJL4;Q9SRT2;F4IQH8; Q0JXE7;Q9ZPW3;Q9LV17	Transcription factor bHLH49	3096.873	0.000
F4IGR4;F4IGR5;P51407	60S acidic ribosomal protein P2-1	654.707	0.002
Q38900;F4IL99;Q41934	Peptidyl-prolyl cis-trans isomerase CYP19-1	339.591	0.003
Q05431;F4HU93;C0Z2H6	L-ascorbate peroxidase 1, cytosolic	287.199	0.003
Q08682;B9DG17;F4J4W3;Q8H173	40S ribosomal protein Sa-1	279.894	0.004
Q9FLB6	PRA1 family protein B3	278.785	0.004
P31168;Q0WL48;C0Z2D8	Dehydrin COR47	224.842	0.004
P25696;Q944M3;Q56WK5	Bifunctional enolase 2/transcriptional activator;Enolase	197.330	0.005
B9DHE0;Q9LHB9;P24101; Q9SMU8;Q9LDN9	Peroxidase 32	165.625	0.006
P57752	Acyl-CoA-binding domain-containing protein 6	158.148	0.006
Q9ZU52;Q41992	Probable fructose-bisphosphate aldolase 3, chloroplastic	157.234	0.006

A novel ARF/bHLH module regulates extra-embryonic identity

Q9XF89	Chlorophyll a-b binding protein CP26, chloroplastic	146.294	0.007
P48491;Q43279	Triosephosphate isomerase, cytosolic	142.499	0.007
O80852;Q42132	Glutathione S-transferase F9	140.453	0.007
Q8LBZ8;Q9SXE9;C0Z2B3	Vacuolar calcium-binding protein-related	137.901	0.007
O82089	Copper transport protein CCH	129.198	0.008
O24456	Guanine nucleotide-binding protein subunit beta-like protein A	120.346	0.008
P42761;Q42082;Q8LEQ8	Glutathione S-transferase F10	116.271	0.009
Q9SCX3;Q29PY2;Q84WM9	Elongation factor 1-beta 2	112.646	0.009
P41916;P41917;Q8H156;Q42161	GTP-binding nuclear protein Ran-1;GTP-binding nuclear protein Ran-2;GTP-binding nuclear protein Ran-3	110.288	0.009
B9DGN3;Q9SLF7;F4IGR3;C0Z3G4	60S acidic ribosomal protein P2-2	107.763	0.009
Q9SAJ4;Q8LFV7;C0Z3A6;Q56ZW1	Phosphoglycerate kinase	105.390	0.009
Q9XI93;Q8LEH5	Involved in response to salt stress	95.716	0.010
Q9STW6	Heat shock 70 kDa protein 6, chloroplastic	91.821	0.011
C0SVP9;P31169	Stress-induced protein KIN2	90.813	0.011
P11829;Q0WT41	Acyl carrier protein 1, chloroplastic;Acyl carrier protein	90.100	0.011
Q9ZP06;Q9LKA3;A8MQK3	Malate dehydrogenase 1, mitochondrial;Malate dehydrogenase 2, mitochondrial;Malate dehydrogenase	88.463	0.011
P92549;F4IMB5;G1C2S9;G1C2Z0;A7KNE3	ATP synthase subunit alpha, mitochondrial;ATP synthase subunit alpha	84.586	0.012
P51427	40S ribosomal protein S5-2	84.410	0.012
Q42403;Q42015	Thioredoxin H3	82.690	0.012
P34790;Q0WWG0;Q1H5D7;Q9SKQ0;Q6LAB2	Peptidyl-prolyl cis-trans isomerase CYP18-3;Peptidyl-prolyl cis-trans isomerase;Peptidyl-prolyl cis-trans isomerase CYP19-2	82.164	0.012
Q9LJE4;P21240	Chaperonin 60 subunit beta 2, chloroplastic;Chaperonin 60 subunit beta 1, chloroplastic	81.247	0.012
B9DGD1;Q43127;C0Z2E9;Q56ZK3	Glutamine synthetase;Glutamine synthetase, chloroplastic/mitochondrial	76.872	0.013
P34789	40S ribosomal protein S28-2	76.802	0.013

Q9ZVR3	Putative protein PHLOEM PROTEIN 2-LIKE B4	76.367	0.013
Q8W4E2;P11574;Q680Z0;Q683E8;Q42022	V-type proton ATPase subunit B3;V-type proton ATPase subunit B1	75.004	0.013
Q9SID0;Q9LNE3;Q9LNE4	Probable fructokinase-1	73.080	0.014
Q9SJQ9;B3H6D7	Fructose-bisphosphate aldolase	71.598	0.014

Table S4: Unique peptides of bHLH proteins.

IP-MS on seedlings	
Unique Peptides	Proteins
AVMLDEIINYVQSLQR	bHLH49; bHLH77
DGYIHVR	bHLH49
LATVNPQMDFNLEGLLAK	bHLH49
MDLSAKDEFSAEK	bHLH49
MDLSAKDEFSAEKR	bHLH49
NPDNYDSVNNPSGDWR	bHLH49
RNPDNYDSVNNPSGDWR	bHLH49
SSEQAKPNVPGSGNVSEDTQSSGGNGQK	bHLH49
TITSPLSPMNGGFK	bHLH49
IP-MS on root tips	
Unique Peptides	Proteins
AVMLDEIINYVQSLQR	bHLH49; bHLH77
FLQDLVPGCNK	bHLH49
GQATNSHSLAER	bHLH49; bHLH74; bHLH76
LATVNPQMDFNLEGLLAK	bHLH49
NPDNYDSVNNPSGDWR	bHLH49
QVEFLSMK	bHLH49; bHLH74; bHLH77; bHLH78; bHLH62; bHLH31; bHLH64; bHLH79
RGQATNSHSLAER	bHLH49; bHLH74; bHLH76
RNPDNYDSVNNPSGDWR	bHLH49
SLEMGWNLPNLLPPK	bHLH49
TITSPLSPMNGGFK	bHLH49

A novel ARF/bHLH module regulates extra-embryonic identity

Supplemental dataset 1: M0171 > > bdl datasets.

Downregulated Genes			Continued			Continued		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At3g13960	25.4	0.39	At5g12330	4.1	1.28	At3g14370	3.1	1.47
At1g80100	19.6	0.39	At2g06850	4.0	1.28	At1g51190	3.1	1.28
At2g01420	15.5	0.39	At1g37140	3.9	1.28	At1g15670	3.1	1.45
At1g15760	8.6	0.39	At5g66350	3.9	1.28	At5g01075	3.1	1.28
At2g21050	8.3	0.39	At1g12980	3.9	1.28	At1g17140	3.1	1.45
At3g15540	8.1	0.39	At2g18890	3.8	1.28	At5g48940	3.0	1.45
At2g45190	8.0	0.39	At5g05220	3.8	1.28	At2g07751	3.0	1.45
At1g53860	7.9	0.39	At1g13400	3.8	1.28	At3g21250	3.0	1.90
At1g24590	7.4	0.39	At2g25790	3.8	1.28	At5g53450	3.0	1.45
At5g05250	7.4	0.39	At3g12970	3.7	1.28	At1g56680	3.0	1.45
At4g29030	7.0	1.28	At3g55560	3.7	1.28	At1g28130	2.9	2.32
At3g19380	7.0	1.28	At4g24780	3.6	1.28	At3g18850	2.9	1.45
At3g02000	6.8	1.28	At2g42870	3.6	1.28	At5g15230	2.9	1.45
At1g68870	6.7	1.28	At1g66090	3.6	1.28	At5g22810	2.9	1.45
At2g14247	6.5	1.28	At1g20230	3.6	1.45	At3g56360	2.9	1.45
At1g10060	6.1	1.28	At1g53815	3.6	1.28	At2g33510	2.9	1.28
At1g71691	6.0	1.28	At1g70510	3.6	1.45	At2g39660	2.9	1.90
At5g45720	5.9	1.28	At1g62420	3.6	1.45	At4g05410	2.9	1.45
At1g52070	5.7	1.28	At5g18270	3.5	1.28	At3g51750	2.9	1.45
At1g70830	5.6	1.28	At2g04920	3.5	1.45	At1g52290	2.9	1.45
At2g33530	5.6	1.28	At3g62100	3.5	1.28	At1g49740	2.8	1.45
At4g31760	5.5	1.28	At4g29020	3.5	1.28	At5g13700	2.8	1.90
At4g31820	5.5	1.28	At5g43870	3.5	1.28	At1g13620	2.8	1.47
At1g29270	5.5	1.28	At1g70720	3.5	1.45	At5g19650	2.8	1.90
At3g51060	5.4	1.28	At1g48870	3.4	1.28	At1g48300	2.8	1.47
At2g26180	5.2	1.28	At1g29950	3.4	1.28	At2g37590	2.8	1.47
At3g54780	5.1	1.28	At3g15680	3.4	1.28	At1g14600	2.8	1.90
At1g73590	5.1	1.28	At1g55200	3.4	1.28	At5g26670	2.8	2.13
At5g50915	5.0	1.28	At1g47400	3.3	1.28	At4g09160	2.7	1.45
At5g01240	4.9	1.28	At1g78430	3.3	1.28	At5g54840	2.7	1.45
At5g14750	4.9	1.28	At2g32280	3.3	1.28	At1g68810	2.7	1.90
At2g39060	4.8	1.28	At5g51560	3.3	1.28	At5g50220	2.7	2.32
At3g16500	4.8	1.28	At2g03090	3.3	1.28	At1g55450	2.7	2.13
At4g36630	4.8	1.28	At1g80440	3.2	1.28	At3g06770	2.7	1.45
At5g66940	4.8	1.28	At5g50570	3.2	1.28	At1g31320	2.7	1.47
At1g23020	4.7	1.28	At1g19200	3.2	1.90	At4g09290	2.6	1.47
At3g53450	4.7	1.28	At5g07780	3.2	1.45	At3g24660	2.6	1.45
At1g68480	4.7	1.28	At1g12030	3.2	1.45	At1g74580	2.6	2.73
At4g23750	4.7	1.28	At1g28070	3.2	1.28	At3g56160	2.6	2.32
At5g55340	4.6	1.28	At3g24450	3.2	1.28	At3g42340	2.6	1.90
At3g24225	4.5	1.28	At4g11140	3.2	1.28	At2g07681	2.6	1.47
At1g28110	4.4	1.28	At5g56460	3.2	1.45	At1g78190	2.6	1.90
At1g75500	4.2	1.28	At1g72940	3.2	1.45	At4g15910	2.6	1.90
At3g57010	4.2	1.28	At2g19810	3.1	1.90	At3g07270	2.6	2.13
At1g21090	4.1	1.28	At2g23050	3.1	1.45	At1g60790	2.6	1.90

Continued			Continued			Continued		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At4g04090	2.6	1.47	At1g67340	2.4	2.73	At3g17680	2.2	2.13
At3g26932	2.6	1.90	At2g26710	2.4	2.73	At1g52930	2.2	2.32
At5g56220	2.6	1.47	At3g56760	2.3	3.43	At1g59030	2.2	3.13
At3g20840	2.6	1.90	At1g16220	2.3	2.13	At4g00430	2.2	2.32
At1g78860	2.6	1.90	At5g59790	2.3	2.13	At1g56020	2.2	2.32
At1g02810	2.6	1.47	At1g31040	2.3	1.90	At2g42840	2.2	2.52
At2g07675	2.6	1.47	At2g42110	2.3	2.13	At1g13960	2.2	2.32
At5g28640	2.5	1.90	At5g26692	2.3	2.73	At3g49670	2.2	2.32
At1g77855	2.5	2.52	At5g06940	2.3	2.13	At3g29030	2.2	2.13
At2g18300	2.5	2.52	At1g44860	2.3	3.43	At3g55750	2.2	2.32
At5g18180	2.5	2.32	At1g79580	2.3	2.73	At5g41663	2.2	3.62
At1g62770	2.5	2.32	At4g27260	2.3	2.32	At4g27010	2.2	2.32
At2g05000	2.5	1.90	At3g25900	2.3	2.13	At2g37940	2.2	2.32
At2g26550	2.5	2.73	At3g43800	2.3	2.13	At2g44860	2.2	2.13
At3g48675	2.5	1.90	At3g08570	2.3	2.32	At1g68740	2.2	2.52
At4g00950	2.5	2.32	At1g69700	2.3	2.13	At1g14640	2.2	3.54
At2g06390	2.5	1.47	At5g17970	2.3	2.13	At1g63100	2.2	2.32
At3g57500	2.5	1.90	At2g16850	2.3	1.90	At4g13235	2.2	2.32
At4g16447	2.5	2.13	At1g15080	2.3	2.13	At3g49410	2.2	2.52
At5g13290	2.5	1.90	At3g15720	2.3	2.32	At5g46790	2.2	2.73
At5g05940	2.5	1.47	At1g65710	2.3	2.32	At2g34910	2.2	2.52
At5g41070	2.5	1.90	At2g37880	2.2	2.52	At4g16990	2.2	3.31
At2g15820	2.5	2.73	At3g51920	2.2	3.13	At1g06390	2.1	2.13
At1g80370	2.4	1.90	At4g16370	2.2	2.13	At2g42050	2.1	2.32
At2g26520	2.4	1.90	At4g17470	2.2	2.52	At4g08910	2.1	2.13
At2g45430	2.4	2.32	At2g11910	2.2	2.32	At2g01630	2.1	2.32
At1g70850	2.4	2.13	At1g30135	2.2	2.32	At1g70895	2.1	2.52
At5g43810	2.4	1.90	At4g29360	2.2	2.13	At1g70890	2.1	2.32
At5g52890	2.4	3.43	At2g22680	2.2	2.13	At4g11350	2.1	3.13
At1g29980	2.4	1.90	At4g29190	2.2	2.52	At5g49490	2.1	2.52
At4g38660	2.4	1.90	At3g59190	2.2	3.13	At2g02240	2.1	2.73
At5g67260	2.4	1.90	At4g18740	2.2	2.52	At5g62710	2.1	3.62
At1g30040	2.4	2.13	At5g60200	2.2	2.32	At2g38370	2.1	2.32
At2g26150	2.4	2.13	At2g19090	2.2	2.13	At1g62220	2.1	2.32
At1g71870	2.4	1.90	At1g11400	2.2	2.73	At5g38790	2.1	2.32
At4g26530	2.4	2.32	At2g14590	2.2	3.13	At1g44740	2.1	2.73
At1g10990	2.4	2.32	At5g18310	2.2	2.32	At4g38140	2.1	3.54
At1g67040	2.4	1.90	At5g43730	2.2	3.31	At5g20160	2.1	2.52
At5g22860	2.4	2.13	At1g80640	2.2	2.32	At1g53360	2.1	2.52
At3g17360	2.4	2.52	At5g37260	2.2	2.13	At2g29350	2.1	2.73
At1g70970	2.4	2.13	At2g41130	2.2	3.31	At2g18500	2.1	2.52
At3g47320	2.4	2.13	At4g34160	2.2	2.32	At2g41170	2.1	2.32
At3g04670	2.4	1.90	At5g20330	2.2	2.32	At3g04770	2.1	2.32
At5g61130	2.4	1.90	At5g13830	2.2	2.32	At4g24390	2.1	2.52
At1g62000	2.4	2.13	At1g69295	2.2	2.32	At2g29310	2.1	2.73

A novel ARF/bHLH module regulates extra-embryonic identity

Continued			Continued			Upregulated Genes		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At4g10270	2.1	2.32	At3g25670	2.0	3.54	At2g27250	44.8	0.00
At2g30700	2.1	3.13	At4g25240	2.0	2.52	At2g31980	30.8	0.00
At2g45160	2.1	2.73	At3g42723	2.0	3.43	At4g21020	26.6	0.00
At5g66560	2.1	2.52	At2g41240	2.0	3.13	At4g27160	24.9	0.00
At1g10040	2.1	3.62	At2g34710	2.0	2.73	At3g54800	24.4	3.62
At1g05710	2.1	2.32	At2g23170	2.0	4.42	At3g62230	24.1	0.00
At3g07540	2.1	3.31	At4g36020	2.0	3.13	At4g27150	23.4	0.00
At1g62720	2.1	4.42	At4g25980	2.0	4.68	At3g22640	22.6	0.00
At1g68760	2.1	2.32	At2g38810	2.0	2.52	At3g27660	22.1	0.00
At5g43630	2.1	2.32	At4g25390	2.0	2.52	At1g52690	20.9	0.00
At2g27840	2.1	2.52	At5g61200	2.0	3.13	At4g34135	19.3	0.00
At1g03170	2.1	2.32	At5g48350	2.0	2.73	At2g26400	17.7	0.00
At2g45450	2.1	2.52	At1g69420	2.0	2.52	At4g15390	17.2	0.00
At2g33370	2.1	2.32	At1g72890	2.0	3.54	At3g63040	15.6	0.00
At1g66250	2.1	2.32	At4g22560	2.0	3.13	At2g23190	15.1	0.00
At2g16570	2.1	2.52	At1g15000	2.0	3.31	At4g36700	14.6	0.00
At3g02010	2.1	2.52	At1g07220	2.0	2.73	At3g12430	13.7	0.00
At1g22050	2.1	2.52	At4g16280	2.0	4.68	At4g32510	13.2	3.43
At4g20940	2.0	2.52	At1g16070	2.0	3.43	At2g40600	13.0	3.83
At4g26990	2.0	4.42	At4g38480	2.0	3.94	At1g75830	12.4	0.00
At5g48500	2.0	2.32	At4g04920	2.0	3.13	At3g17520	12.3	0.00
At1g57610	2.0	2.52	At5g19300	2.0	2.52	At3g01570	12.1	0.00
At1g25290	2.0	3.83	At5g05510	2.0	2.73	At5g44630	11.9	0.00
At4g03100	2.0	2.32	At3g21170	2.0	3.62	At4g27140	11.9	0.00
At4g02290	2.0	2.52	At5g37478	2.0	3.54	At4g27170	11.6	0.00
At1g61095	2.0	2.52	At1g28080	2.0	3.31	At4g23680	11.4	0.00
At5g66600	2.0	4.42	At1g25530	2.0	2.52	At5g46070	11.2	4.09
At3g13500	2.0	2.52	At3g15300	2.0	3.54	At1g47540	10.8	0.00
At1g28310	2.0	3.54	At1g18630	2.0	2.73	At5g12910	10.8	3.31
At5g13740	2.0	3.62	At1g62900	2.0	3.43	At1g04550	10.6	0.00
At3g60390	2.0	3.43	At2g31920	2.0	2.73	At3g58410	10.5	0.00
At2g34320	2.0	4.68	At2g34390	2.0	3.62	At1g02450	10.2	0.00
At2g18340	2.0	2.32	At5g13680	2.0	2.73	At3g07850	9.9	0.00
At2g41650	2.0	2.52	At1g02870	2.0	3.13	At3g21180	9.9	4.68
At2g47930	2.0	2.52				At4g34530	9.9	0.00
At1g71692	2.0	3.54				At2g27990	9.8	0.60
At1g70710	2.0	2.52				At5g43430	9.6	2.13
At4g32880	2.0	3.62				At3g49260	9.6	0.00
At2g20490	2.0	2.52				At3g54940	9.5	0.00
At3g21850	2.0	2.73				At5g59230	9.4	2.32
At1g12560	2.0	2.73				At2g40990	9.3	3.94
At3g27580	2.0	3.43				At2g38870	9.0	0.00
At2g17580	2.0	2.52				At1g43780	8.8	0.00
At3g03920	2.0	2.73				At4g36490	8.7	0.00
At1g28580	2.0	3.83				At1g55920	8.7	0.00

Continued			Continued			Continued		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At2g15490	8.5	0.00	At2g42340	5.1	5.49	At3g07820	4.3	0.92
At1g32560	8.4	0.00	At1g61320	5.1	2.52	At2g16730	4.2	0.60
At4g29270	8.0	0.00	At5g04460	5.1	2.52	At3g57690	4.2	0.75
At5g11930	8.0	0.00	At5g24790	5.1	1.18	At1g04920	4.2	0.00
At3g22490	8.0	0.60	At2g34340	5.0	0.00	At1g22100	4.2	0.00
At5g11320	7.9	0.00	At2g43680	5.0	3.83	At3g45310	4.2	0.00
At1g65490	7.9	0.00	At1g69600	5.0	0.00	At3g54320	4.2	0.60
At2g36710	7.2	2.13	At5g50030	5.0	0.00	At1g61563	4.2	0.00
At4g11940	7.2	0.21	At1g12080	5.0	0.00	At4g06724	4.1	4.42
At5g19580	7.2	0.00	At5g13310	5.0	4.09	At1g29640	4.1	2.32
At2g36020	7.0	0.00	At1g28670	4.9	0.00	At3g61160	4.1	0.00
At4g28405	6.9	0.00	At2g32990	4.9	0.00	At5g01300	4.1	0.00
At3g09260	6.8	0.60	At1g48130	4.9	0.00	At5g24530	4.1	0.00
At1g68170	6.7	0.00	At3g23090	4.8	0.00	At1g65090	4.1	0.00
At3g51590	6.7	0.00	At3g42725	4.8	0.39	At4g30670	4.0	0.39
At1g77100	6.5	0.00	At1g09155	4.8	3.31	At3g25170	4.0	0.00
At1g15520	6.4	2.13	At5g09210	4.8	0.00	At1g65550	4.0	1.18
At5g17700	6.3	0.00	At4g25140	4.8	0.00	At1g09750	4.0	0.00
At4g36880	6.3	0.00	At1g75450	4.8	2.32	At2g29090	4.0	1.90
At5g25830	6.3	0.60	At5g16570	4.8	0.00	At5g05190	4.0	3.83
At1g34095	6.2	0.00	At2g47030	4.8	0.00	At1g70410	4.0	0.39
At5g08030	6.1	0.60	At3g29970	4.7	5.13	At4g33600	4.0	0.00
At3g62170	6.0	0.00	At2g21840	4.7	3.13	At5g56140	3.9	2.73
At4g32540	5.9	0.00	At1g02230	4.7	0.60	At4g25433	3.9	0.00
At2g38920	5.9	0.60	At5g44120	4.7	0.60	At5g23320	3.9	0.75
At3g44790	5.9	4.42	At3g01270	4.7	0.00	At5g12420	3.9	0.00
At2g47040	5.9	0.00	At1g52090	4.7	4.09	At1g14420	3.9	0.00
At5g07500	5.7	0.00	At3g26110	4.6	0.00	At1g33280	3.9	1.28
At4g16690	5.7	0.60	At1g33750	4.6	0.00	At3g49540	3.9	0.39
At5g40420	5.7	0.00	At2g27380	4.6	0.00	At2g16660	3.9	0.21
At5g27420	5.7	1.47	At5g28680	4.5	0.00	At4g18596	3.9	0.00
At5g12380	5.6	3.83	At2g47050	4.5	0.00	At1g09790	3.9	0.60
At2g38940	5.6	0.00	At3g03520	4.5	3.62	At3g23870	3.8	0.00
At1g53690	5.5	0.00	At5g01600	4.5	0.00	At5g06250	3.8	0.39
At5g51210	5.5	0.00	At3g20580	4.5	0.00	At5g28000	3.8	0.92
At5g15110	5.4	0.00	At1g32780	4.5	0.21	At3g14040	3.8	1.18
At1g55740	5.4	0.00	At5g62340	4.4	0.00	At3g52970	3.7	1.45
At3g05610	5.3	0.00	At4g25810	4.4	2.13	At1g35400	3.7	5.13
At5g39860	5.2	0.00	At5g48330	4.4	0.75	At1g78500	3.7	0.21
At4g30340	5.2	4.09	At5g04180	4.3	0.00	At5g27200	3.7	0.21
At1g65500	5.2	0.00	At1g48470	4.3	0.00	At3g49040	3.6	2.13
At1g76470	5.2	0.00	At4g35010	4.3	0.39	At1g17620	3.6	5.13
At3g05930	5.2	0.00	At2g31500	4.3	0.00	At5g53870	3.6	0.21
At5g26717	5.2	0.00	At4g02250	4.3	0.00	At1g35560	3.6	3.13
At1g65610	5.1	0.60	At4g15750	4.3	1.47	At4g40020	3.6	3.62

A novel ARF/bHLH module regulates extra-embryonic identity

Continued			Continued			Continued		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At3g46260	3.6	3.62	At5g01520	3.2	3.31	At3g13790	2.9	0.60
At5g27910	3.6	0.75	At1g79160	3.2	0.39	At1g11740	2.9	0.60
At3g28150	3.6	0.39	At2g14290	3.2	0.60	At4g23690	2.9	0.92
At5g38850	3.6	3.54	At2g38530	3.2	0.39	At1g54550	2.8	4.68
At3g28750	3.5	0.39	At1g78320	3.2	0.60	At3g49160	2.8	0.92
At3g61890	3.5	4.42	At4g17500	3.1	2.13	At1g44830	2.8	1.28
At1g48910	3.5	0.60	At3g51420	3.1	0.60	At1g01490	2.8	0.60
At4g27520	3.5	1.28	At2g40610	3.1	1.18	At5g66670	2.8	1.18
At1g78520	3.5	0.39	At3g12580	3.1	0.60	At1g18010	2.8	1.18
At5g51690	3.5	3.54	At5g38160	3.1	0.39	At4g13000	2.8	0.75
At5g44580	3.4	0.60	At2g03980	3.1	0.60	At1g47840	2.8	0.60
At4g37360	3.4	0.39	At5g43935	3.1	4.68	At1g68510	2.8	0.60
At5g52340	3.4	3.62	At4g01890	3.1	5.13	At2g21420	2.8	2.13
At2g26850	3.4	0.39	At1g22590	3.1	0.60	At4g19380	2.8	0.60
At3g06090	3.4	0.39	At2g47115	3.1	5.49	At5g16410	2.8	1.45
At3g28430	3.4	4.09	At2g32270	3.1	0.60	At4g04460	2.8	0.60
At5g01610	3.4	0.60	At1g17060	3.0	0.60	At4g09930	2.8	0.60
At1g08140	3.4	0.92	At2g36550	3.0	0.60	At1g28650	2.8	0.92
At1g78260	3.4	0.60	At3g44910	3.0	5.13	At2g46950	2.8	0.60
At3g13784	3.4	5.13	At4g18020	3.0	2.73	At3g55890	2.8	3.83
At4g34830	3.4	2.32	At3g05260	3.0	4.09	At1g10370	2.8	1.47
At4g28650	3.3	0.75	At1g10770	3.0	3.62	At5g22920	2.8	2.32
At2g28680	3.3	0.21	At1g67810	3.0	0.60	At5g12460	2.8	2.13
At1g02790	3.3	0.75	At3g14205	3.0	3.62	At3g29240	2.8	2.73
At2g25770	3.3	0.39	At1g68320	3.0	0.60	At2g25940	2.8	0.75
At3g28345	3.3	0.39	At4g03965	3.0	0.60	At3g12880	2.8	0.60
At1g56600	3.3	3.13	At1g35750	2.9	2.13	At4g34860	2.7	3.54
At2g30230	3.3	1.28	At1g04560	2.9	1.18	At2g38380	2.7	3.54
At1g28660	3.3	0.39	At1g28640	2.9	0.60	At5g01360	2.7	0.60
At1g62290	3.3	0.39	At1g73190	2.9	1.45	At5g49680	2.7	2.73
At1g66570	3.3	2.73	At5g18900	2.9	2.73	At1g27040	2.7	1.18
At2g23510	3.3	0.60	At2g38060	2.9	5.13	At5g66520	2.7	1.28
At2g24450	3.3	0.60	At3g05800	2.9	3.13	At1g05170	2.7	0.60
At3g22410	3.3	1.18	At5g66020	2.9	0.75	At1g06030	2.7	0.60
At3g05150	3.3	0.60	At4g09760	2.9	3.83	At1g63060	2.7	1.28
At5g42300	3.3	2.73	At2g24762	2.9	4.68	At3g25165	2.7	2.32
At3g58330	3.2	0.60	At3g20220	2.9	0.60	At5g09640	2.7	0.60
At5g39050	3.2	0.60	At1g27170	2.9	0.60	At2g17190	2.7	4.68
At1g25480	3.2	4.42	At3g57020	2.9	2.73	At5g26650	2.7	1.18
At1g06850	3.2	2.32	At3g16180	2.9	0.60	At1g78380	2.7	0.60
At1g04250	3.2	2.73	At5g59340	2.9	0.75	At5g23840	2.7	4.68
At2g41340	3.2	1.90	At4g21900	2.9	0.60	At2g04100	2.7	0.75
At2g14560	3.2	0.60	At2g23790	2.9	3.62	At1g58643	2.7	0.60
At1g58310	3.2	0.60	At1g70560	2.9	0.60	At2g02120	2.7	0.60
At1g44760	3.2	0.39	At1g54860	2.9	1.28	At5g51950	2.7	1.28

Continued			Continued			Continued		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At1g53930	2.7	0.75	At5g07430	2.5	1.18	At3g09340	2.3	1.18
At1g53950	2.7	0.60	At5g26950	2.5	1.47	At2g46450	2.3	1.90
At2g48110	2.7	5.13	At3g44300	2.5	0.92	At2g02990	2.3	0.92
At5g25820	2.7	0.60	At1g66500	2.5	3.13	At2g05420	2.3	2.32
At5g14780	2.7	0.60	At5g47670	2.5	1.28	At5g42220	2.3	4.42
At3g13400	2.7	0.92	At2g24600	2.5	3.43	At3g14990	2.3	1.28
At5g06270	2.6	0.60	At1g17745	2.5	2.73	At2g43520	2.3	1.47
At1g61566	2.6	1.47	At5g50490	2.5	0.92	At2g02780	2.3	5.13
At3g21620	2.6	0.60	At5g19710	2.5	5.49	At1g24520	2.3	2.13
At3g20270	2.6	0.60	At4g37800	2.5	0.92	At5g42030	2.3	4.09
At2g35500	2.6	5.13	At2g40730	2.5	3.62	At4g32870	2.3	0.92
At5g59110	2.6	0.60	At5g26700	2.5	1.18	At2g45830	2.3	3.43
At2g19360	2.6	5.49	At5g63750	2.5	1.18	At2g31570	2.3	1.18
At1g63300	2.6	5.13	At3g28857	2.4	0.92	At5g39180	2.3	1.18
At1g51250	2.6	3.83	At2g45800	2.4	0.92	At3g02410	2.3	1.28
At1g11765	2.6	1.47	At5g19140	2.4	0.75	At2g38750	2.3	1.18
At5g49190	2.6	0.60	At5g52560	2.4	3.83	At4g21380	2.3	1.18
At3g17265	2.6	4.42	At1g70540	2.4	1.47	At1g63360	2.3	1.45
At1g80580	2.6	3.94	At5g20230	2.4	3.54	At5g38950	2.3	1.28
At5g07950	2.6	5.49	At2g39705	2.4	0.92	At5g14860	2.3	1.18
At2g18660	2.6	3.54	At3g28030	2.4	3.31	At1g35290	2.3	1.18
At5g44020	2.6	1.28	At4g20530	2.4	0.92	At5g54140	2.3	4.68
At4g02190	2.6	5.49	At5g48510	2.4	1.45	At3g07775	2.3	1.18
At3g09330	2.6	1.28	At4g21960	2.4	1.18	At1g29600	2.3	1.45
At5g23405	2.6	4.68	At4g23070	2.4	1.18	At1g22620	2.3	2.52
At3g11160	2.6	0.75	At2g35310	2.4	0.92	At3g26860	2.3	1.28
At4g23600	2.6	3.43	At2g25310	2.4	3.83	At4g13050	2.3	1.28
At3g42640	2.6	1.28	At4g39270	2.4	4.42	At5g52230	2.3	1.45
At5g03590	2.6	1.28	At1g54010	2.4	0.92	At5g17330	2.3	1.18
At5g56270	2.5	1.28	At4g22235	2.4	0.92	At1g60970	2.3	1.18
At2g43510	2.5	0.75	At3g61150	2.4	1.18	At4g23090	2.3	1.18
At5g63240	2.5	0.60	At1g10550	2.4	1.18	At2g23970	2.3	1.18
At1g74550	2.5	0.75	At4g24510	2.4	1.18	At1g28420	2.3	3.43
At4g32380	2.5	0.60	At3g57800	2.4	3.31	At4g14740	2.3	1.28
At5g41800	2.5	3.13	At1g68440	2.4	0.92	At1g08630	2.3	1.18
At1g65810	2.5	0.60	At1g24470	2.4	0.92	At1g50650	2.3	1.47
At1g68920	2.5	5.49	At3g13080	2.4	2.32	At1g62360	2.3	1.18
At5g45200	2.5	2.52	At4g26710	2.4	3.83	At1g02820	2.3	2.13
At3g02970	2.5	2.73	At4g00330	2.4	3.83	At3g10200	2.2	1.28
At2g01610	2.5	4.09	At2g14160	2.4	1.18	At5g39530	2.2	4.09
At2g43180	2.5	4.42	At2g37760	2.4	1.18	At2g04230	2.2	2.73
At3g21700	2.5	1.28	At1g23350	2.4	0.92	At3g17810	2.2	1.28
At2g37100	2.5	1.47	At1g35140	2.4	2.32	At4g20830	2.2	1.45
At1g76290	2.5	1.28	At1g76500	2.3	1.18	At2g19000	2.2	3.83
At5g64530	2.5	1.18	At1g60540	2.3	1.47	At1g05320	2.2	1.28
At1g33055	2.5	1.18	At1g22430	2.3	1.18	At2g12940	2.2	1.28

A novel ARF/bHLH module regulates extra-embryonic identity

Continued			Continued			Continued		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At5g06530	2.2	2.73	At1g71450	2.2	4.68	At1g51010	2.0	3.43
At5g07550	2.2	2.32	At2g01031	2.2	1.28	At3g22235	2.0	2.52
At5g11530	2.2	3.43	At5g24480	2.1	1.45	At4g35650	2.0	1.47
At3g15760	2.2	1.47	At2g44240	2.1	3.43	At1g55090	2.0	4.09
At5g53030	2.2	3.43	At5g02020	2.1	3.62	At4g14780	2.0	2.13
At5g57670	2.2	4.42	At4g37250	2.1	1.47	At1g78460	2.0	1.47
At3g08940	2.2	1.28	At2g34970	2.1	2.13	At2g47710	2.0	1.47
At2g44560	2.2	1.18	At4g00990	2.1	1.28	At1g53790	2.0	3.13
At3g05625	2.2	2.52	At1g04700	2.1	5.49	At3g20865	2.0	1.47
At4g22590	2.2	1.45	At5g38930	2.1	1.28	At2g12900	2.0	1.45
At4g15500	2.2	2.32	At4g13840	2.1	1.28	At4g39940	2.0	4.68
At1g06980	2.2	1.28	At3g23730	2.1	1.28	At5g53080	2.0	4.68
At3g02110	2.2	1.18	At3g15820	2.1	1.28	At1g22110	2.0	2.32
At2g25890	2.2	1.45	At1g36970	2.1	4.42	At2g24610	2.0	5.13
At2g27180	2.2	1.28	At1g78370	2.1	1.28	At4g30960	2.0	1.47
At1g25450	2.2	1.28	At5g25470	2.1	3.43	At3g61470	2.0	1.47
At5g41900	2.2	3.54	At2g35300	2.1	2.73	At5g05030	2.0	1.47
At3g19350	2.2	1.28	At1g48930	2.1	1.45	At2g22660	2.0	1.47
At3g54890	2.2	3.31	At5g61190	2.1	1.90	At3g55740	2.0	2.32
At4g24040	2.2	3.54	At5g05460	2.1	1.90	At1g71015	2.0	4.09
At1g68610	2.2	2.13	At3g19000	2.1	1.28	At5g11100	2.0	2.73
At3g01900	2.2	1.28	At3g03770	2.1	5.13	At3g61230	2.0	2.13
At3g61010	2.2	1.90	At5g51550	2.1	1.90	At1g41830	2.0	5.49
At1g76090	2.2	1.28	At3g03270	2.1	1.28	At5g52600	2.0	1.90
At1g26795	2.2	1.28	At2g44550	2.1	1.28	At2g17780	2.0	4.68
At1g12180	2.2	1.18	At5g19980	2.1	2.32	At5g57910	2.0	5.13
At5g59220	2.2	1.45	At3g63200	2.1	1.45	At1g74100	2.0	3.62
At5g46230	2.2	1.28	At5g50480	2.1	3.83	At4g05250	2.0	1.90
At4g34050	2.2	1.18	At2g42000	2.1	5.49	At1g15150	2.0	1.47
At5g02190	2.2	2.13	At2g27080	2.1	2.32	At4g34580	2.0	1.90
At1g64830	2.2	1.28	At4g24220	2.1	2.52	At1g03230	2.0	3.54
At5g39400	2.2	2.73	At1g01130	2.1	1.90	At5g39440	2.0	2.52
At1g17780	2.2	2.32	At2g41190	2.1	1.90	At1g18000	2.0	1.90
At1g66460	2.2	1.28	At3g60130	2.1	1.45	At4g17940	2.0	2.32
At1g22210	2.2	2.73	At1g19320	2.1	2.52	At5g37770	2.0	1.47
At5g46110	2.2	1.28	At5g25450	2.1	2.32	At1g07720	2.0	2.32
At1g68110	2.2	1.90	At1g03630	2.1	1.45	At4g23500	2.0	1.90
At5g47550	2.2	1.28	At5g02160	2.1	1.45	At1g55020	2.0	2.13
At1g67330	2.2	1.47	At5g39030	2.1	2.13	At2g26070	2.0	1.47
At4g15053	2.2	2.13	At1g20620	2.1	1.47	At5g25770	2.0	3.13
At1g69860	2.2	1.47	At5g62620	2.1	1.90	At3g50820	2.0	2.13
At1g03620	2.2	3.31	At3g53990	2.1	1.45	At5g51010	2.0	2.13
At3g01700	2.2	1.90	At1g03210	2.1	2.73	At1g06020	2.0	5.49
At1g62660	2.2	3.31	At5g59700	2.1	3.94	At3g13390	2.0	2.73
At1g13830	2.2	1.28	At5g48000	2.1	4.42	At1g68795	2.0	3.83
At3g19020	2.2	4.09	At3g05320	2.0	5.49	At2g05100	2.0	3.13

Continued			Continued			Continued		
AGI	FC	q-value(%)	AGI	FC	q-value(%)	AGI	FC	q-value(%)
At3g58980	2.0	2.13	At4g19200	2.0	1.47	At4g03510	2.0	3.31
At4g35560	2.0	1.90	At3g22550	2.0	4.09	At1g25400	2.0	1.90
At1g02930	2.0	2.13	At1g27560	2.0	5.13	At1g33700	2.0	2.13
At1g34790	2.0	1.90	At2g35290	2.0	2.32	At5g35460	2.0	1.90
At1g35513	2.0	2.13	At2g29300	2.0	4.42	At1g72130	2.0	2.52
At5g07020	2.0	3.62	At1g67750	2.0	2.32	At3g45600	2.0	1.47
At1g78995	2.0	2.13	At2g44580	2.0	1.47	At1g24540	2.0	1.90
At2g21140	2.0	2.73	At3g16490	2.0	3.31	At2g34420	2.0	1.90
At2g44670	2.0	4.09	At1g01430	2.0	1.47	At4g09650	2.0	1.90
At4g25100	2.0	1.47	At5g54350	2.0	3.83	At5g53240	2.0	2.52

Supplemental dataset 2: Primers used in this study.

AT number	Gene name	Primer orientation	Sequence
Primer pairs used for transcriptional fusions			
AT4G32540	YUC1	Sense	TAGTTGGAATGGGTTCGAATCACTATCAAACCAAGATAAAGT
		Antisense	TTATGGAGTTGGGTTCGAATCTTGATGGATGATGAAAA
AT4G31820	NPY1/ENP/MAB4	Sense	TAGTTGGAATGGGTTCGAAAAATGTAITCAACCAACTCGA
		Antisense	TTATGGAGTTGGGTTCGAAAAATTATGTATATACTTATAT
AT2G23050	NPY4	Sense	TAGTTGGAATGGGTTCGAAAAATTATCGTCTCGCAGGAT
		Antisense	TTATGGAGTTGGGTTCGAAGTCACTTTAAAAAGAAACA
AT1G80640	-	Sense	TAGTTGGAATGGGTTCGAAAAACATCCCTAATTCATAACCATCT
		Antisense	TTATGGAGTTGGGTTCGAAGTGGAGAGAGTGAGGAGGAG
AT3G54320	WR1	Sense	TAGTTGGAATGGGTTCGAAACAAATCACTCTTCGCTAAGATTAC
		Antisense	TTATGGAGTTGGGTTCGAAAAGGAGGAAAGGGCTAATTG
AT1G15670	-	Sense	TAGTTGGAATGGGTTCGAATGTCTTAGAACCAATCTCATG
		Antisense	TTATGGAGTTGGGTTCGAATAAAAAATATCATTTAAATCA
AT1G73590	PIN1	Sense	TAGTTGGAATGGGTTCGAATGTGACTTTTAGTATCAGC
		Antisense	TTATGGAGTTGGGTTCGAAGAGAAGAGACCACATTTTTA
AT1G80440	KMD1	Sense	TAGTTGGAATGGGTTCGAATGAAGAAAATACTTGTTAAG
		Antisense	TTATGGAGTTGGGTTCGAAGGGAAGAGGAGAGTGTGCC
AT2G42870	bHLH165	Sense	TAGTTGGAATGGGTTCGAATCTCCATCTCTCTCTCTT
		Antisense	TTATGGAGTTGGGTTCGAATGAAAGAAAAGAGAGAGATGA
AT3G62100	IAA30	Sense	TAGTTGGAATGGGTTCGAAGACATTGTTTGGATGTTTCG
		Antisense	TTATGGAGTTGGGTTCGAACTAAAACAGGTGATTAATGA
AT3G02210	COBL1	Sense	TAGTTGGAATGGGTTCGAAACTCATGTTTGGTTGTACTA
		Antisense	TTATGGAGTTGGGTTCGAACTGAAGCAAAAAAGAGAGAGAAA
AT1G68920	bHLH49	Sense	TAGTTGGAATGGGTTCGAAGGAAATTGCAAGTCCCAAGG
		Antisense	TTATGGAGTTGGGTTCGAAGTGTGTTCATCAAGTATTAGAAGAAG
AT3G57800	bHLH60	Sense	TAGTTGGAATGGGTTCGAACGATAGAAACGAGCAATGCT
		Antisense	TTATGGAGTTGGGTTCGAAAAACGAAACTCACTGGAAA

A novel ARF/bHLH module regulates extra-embryonic identity

AT4G34530	bHLH63	Sense	TAGTTGGAATGGGTTCGAATGGCTAAAACACCTTATA
		Antisense	TTATGGAGTTGGGTTCGAATAAAGGTGAGGTTAGAAGAG
AT2G41240	bHLH100	Sense	TAGTTGGAATGGGTTCGAAACCTCTCCTACGTGCATTGA
		Antisense	TTATGGAGTTGGGTTCGAATTTGAGTTTATAGATAGTTACTC
Primer pairs used for translational fusions			
AT1G68920	bHLH49	Sense	TAGTTGGAATGGGTTCAAGAAGTTATTAGCATTAACGTG
		Antisense	TTATGGAGTTGGGTTCGAACTGGCTCAACCTTCATATTTG
AT3G57800	bHLH60	Sense	TAGTTGGAATGGGTTCAAGGAAGTGAGAATAGAGATGAG
		Antisense	TTATGGAGTTGGGTTCAACACAGCTCCATTTTGACCTGAT
AT4G34530	bHLH63	Sense	TAGTTGGAATGGGTTCGAAATTAATAATGATAAGTGTCA
		Antisense	TTATGGAGTTGGGTTCAACAACCTCTAAATTGCCATAGA
AT2G41240	bHLH100	Sense	TAGTTGGAATGGGTTCGAAAACCGAAGTGTGTACTGTT
		Antisense	TTATGGAGTTGGGTTCAACTGTAACGAGTGTCCACATT
Primer pairs used for misexpression			
AT1G68920	bHLH49	Sense	TAGTTGGAATAGGTTTCATGGATTTAAGTGCGAAAGA
		Antisense	AGTATGGAGTTGGGTTCTCATGGCTCAACCTTCATAT
AT3G57800	bHLH60	Sense	TAGTTGGAATAGGTTTCATGGATCTGACTGGAGGATT
		Antisense	AGTATGGAGTTGGGTTCTTACAGCTCCATTTTGACCT
AT4G34530	bHLH63	Sense	TAGTTGGAATAGGTTTCATGAATGGAGCTATAGGAGG
		Antisense	AGTATGGAGTTGGGTTCTCAAACCTCTAAATTGCCAT
AT2G41240	bHLH100	Sense	TAGTTGGAATAGGTTTCATGTGTGCACCTGTCCCTCC
		Antisense	AGTATGGAGTTGGGTTCTCATGTAAACGAGTGTCCAC
Primer pairs used for genotyping			
SALK_135188C	<i>bhlh49-1</i>	LP	TTCCGTCGTAACAACGAATC
		RP	CTAGTACGGGTGCAACAAG
SALK_087424C	<i>bhlh49-2</i>	LP	TAACACAGGGCAATGGA AAAAG
		RP	TGCTGAAAACCTCATCTTTCGC
SAIL_1219_E01	<i>bhlh60-3</i>	LP	AAAAGCAGCAAGAGAATGACG
		RP	TCAAGCAACCGATAGCCATAG
SAIL_1211_F11	<i>bhlh63-2</i>	LP	TAAATTTGGGATTAGTCCCGC
		RP	ACAAGATCACAGGCAAAGCAG
SALK_150637C	<i>bhlh100-1</i>	LP	TTGGTCGGTGTAACGAGATC
		RP	TTGTGGTAGAAAAATGTGATTGC
SALK_074568C	<i>bhlh100-2</i>	LP	TCTCAATAGTCCACGTCCACC
		RP	AATGCTTGTGAAACTGTTGCC
Primer pairs used for FRET-FLIM cloning			
AT1G68920	bHLH49	Sense	TAGTTGGAATAGGTTTCATGGATTTAAGTGCGAAAGATG
		Antisense	AGTATGGAGTTGGGTTCTGGCTCAACCTTCATATTTGC
Primer pairs used for qRT-PCR			
AT1G68920	bHLH49	Sense	CCTCTACCTCATGGATTTCATGC
		Antisense	GGATCACATTTTGCAAAATCACC
AT3G57800	bHLH60	Sense	GGACATTGGCTTTTCCTTC
		Antisense	GGATTTGATTCTCAATCGC

AT4G34530	bHLH63	Sense	CCAATCCACAACATAACAATG
		Antisense	CTCTTTGATACATCAACTAATGGG
AT2G41240	bHLH100	Sense	CCGACCAAAAACAGTAAGTCAG
		Antisense	ACACCACTCAAGACATTCCC
	ATR3	Sense	GAAGCTAGTGAGTCTGTTG
		Antisense	GAATCATGACCAGGAGGA
AT4G11650	-	Sense	GAGAATTAATAGACTCGTATC
		Antisense	CGACTAACGTAATTATTATTGC
AT2G33790	-	Sense	AACTCTCCCACCGATCAA
		Antisense	ACACCTCTAACTGCCACTA
AT2G46950	-	Sense	TCTTACTGACCTCTATTTC
		Antisense	TCTTTGGATTCCGATGTT
AT3G21720	-	Sense	ATCTATGTCTCTGGTTGG
		Antisense	TACGGATAATCAGCAAGA
AT5G41300	-	Sense	TGCCITGAATATGCGTTG
		Antisense	GCAATGGTGAAGTTGTGA
AT2G41070	bZIP12	Sense	AATCTCCTGCCACCTCTC
		Antisense	TCTGCCTCGTTAATGACTGA
AT5G14750	MYB66	Sense	CAATCTCCGAATCCTACC
		Antisense	TGAACTCAAGTAGTACITCCTT
AT1G70895	CLE17	Sense	TGTTGGTACGAAGACAGG
		Antisense	CGAGGTGAAGACGAGAATA
Reference gene	CDKA1;1	Sense	ATTGCGTATTGCCACTCTCATAGG
		Antisense	TCCTGACAGGGATACCGAATGC
Reference gene	EEF α 4	Sense	CTGGAGGTTTTGAGGCTGGTAT
		Antisense	CCAAGGTTGAAAGCAAGAAGA
Reference gene	GAPC	Sense	GAAGGGTGGTGCCAAGAAGGTT
		Antisense	AGGGGAGCAAGGCAGTTAGTGG
Other primers			
	SAIL	BP_LB2	GCTTCCTATTATATCTTCCCAAATTACCAATACA
	SALK	LBb1.3	ATTTTGCCGATTCGGAAC
	PGIKN3GFP	Antisense	TTTCTCTTCTTTGAGGCCATGG

Chapter 3

A genome-wide screen for embryo inducers

Tatyana Radoeva, Catherine Albrecht,
Sacco de Vries and Dolf Weijers

Laboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703HA
Wageningen, the Netherlands

Embryogenesis in flowering plants establishes multicellularity from a single reproductive cell, the zygote. After fertilization, the zygote undergoes several rounds of controlled cell divisions to generate a mature embryo. However, embryo formation can also be induced in a variety of other cell types. These non-zygotic embryos go through analogous developmental phases and are morphologically similar to the zygotic embryo, although their initial cell division pattern is often less regular. Despite its fundamental importance and enormous application potential, the mechanisms that alter cell fate from non-embryonic to embryonic are elusive. In the past decades, a variety of different model systems have been used to identify regulators of embryo induction, but it is unclear if these act in a common network. We recently found that inhibition of auxin response in the extra-embryonic suspensor cells cell-autonomously and predictably triggers a switch towards embryo identity. We here use the suspensor as a uniform model system to study the crucial first reprogramming step of embryo initiation. We have tested the ability of fifteen known embryogenesis inducers to promote embryo formation in suspensor cells. Surprisingly, only the RKD1 transcription factor could do so. To identify the entire set of genes capable of inducing embryogenesis from suspensor cells, we developed a systematic local misexpression screen.

Introduction

Embryogenesis in flowering plants begins with fertilization of the egg cell by a sperm cell to form the zygote. However, initiation of embryo development in plants is not restricted to the fertilized egg cell, but has been reported to occur from a variety of different cell types (de Vries et al., 1988; Koltunow et al., 1995; Rademacher et al., 2012; Touraev et al., 1997). In some plant species, these are genetically derived states (e.g. leaf margin embryos in *Kalanchoe* and apomixis in *Taraxacum*; Dijk et al., 2009; Garces and Sinha, 2009), while in other species embryogenesis can be experimentally induced by hormone or stress treatment (Mordhorst et al., 1997; Schmidt et al., 1997). Regardless of the origin of the embryo, the developmental program of the non-zygotic embryos is rather similar to the zygotic ones, and they are thought to go through analogous developmental phases (Dodeman et al., 1997; Harada et al., 2001; Mordhorst et al., 2002). In fact, because the end result is a normal seedling, this widespread ability to induce embryogenesis from somatic cells is often used in clonal propagation of crop plants.

While the initiation of zygotic embryogenesis is marked by fertilization, the processes that trigger other cells to change their fate and enter an embryonic developmental program are unknown. A fundamental question is whether there is a dedicated and conserved genetic pathway to trigger embryogenesis regardless of the cell of origin, and if this pathway is controlled by a specific set of genes. Alternatively, the embryo initiation response may represent a cellular “default” state used by the cell upon particular stresses/signals (Radoeva and Weijers, 2014). As discussed in Chapter 1, several genes have been identified that are able, either in direct or indirect manner, to promote embryogenesis. However, these were identified and tested in diverse experimental systems (e.g. *Brassica* microspores, *Arabidopsis* meristems, embryo cultures or seedlings), and it is therefore impossible to interpret if these are all part of a coherent genetic network (reviewed in Radoeva and Weijers, 2014; Chapter 1).

In *Arabidopsis*, the first zygotic division sets apart an embryonic lineage (apical cell) and an extra-embryonic suspensor lineage (basal cell). The suspensor is a transient structure that is fully completed by the globular stage of embryo development and probably degenerates thereafter via programmed cell death (Bozhkov et al., 2005; Yeung and Meinke, 1993). The general view is that the suspensor serves as a support for positioning the embryo within the seed, and a vital transport route for nutrients and growth regulators to the pro-embryo (Kawashima and Goldberg, 2010). We here use the established developmental plasticity of the extra-embryonic suspensor cells, which extends beyond its normal, rather restricted role (Haccius, 1978). In many plant species, secondary embryos are formed from suspensor cells (Lakshmanan

and Ambegaokar, 1984) and in addition, in several experiments, suspensor-derived embryogenesis has been induced by stress treatments (salt, radiation; Haccius, 1955) or mutations (such as *sus* and *twm2*; Schwartz et al., 1994; Zhang and Somerville, 1997) that impair viability of the original embryo (Haccius, 1955; Liu et al., 2015; Weijers et al., 2003). Thus, suspensor cells can be considered as latent pool of stem cells that can switch to embryo identity upon need. Until recently, the molecular mechanism that controls the switch from suspensor to embryo identity was elusive. It was recently shown that the plant hormone auxin cell-autonomously controls the suspensor fate maintenance. Upon local auxin inhibition, by ectopic expression of stabilized Aux/IAA12 (*bdl*) protein, or by stabilization of the suspensor-specific IAA10 auxin response inhibitor, suspensor cells lose their identity and switch to embryonic cell fate, which can ultimately result in the formation of conjoined twin seedlings (Rademacher et al., 2012). This event represents an alternative, yet very predictable path towards embryogenesis that occurs in a small, defined population of cells. Hence, it can be exploited as an experimental system to unravel the mechanisms underlying plant embryo initiation.

In this chapter, we have first tested the ability of known embryonic regulators to initiate embryogenesis in suspensor cells. Our results show that besides *bdl*, only the plant-specific RWP-RK domain-containing transcription factor RKD1 (Koszegi et al., 2011) could efficiently induce suspensor-derived embryos. This suggests that suspensor to embryo transition likely requires a specific set of genetic regulators. Inspired by this finding, we have developed a genome-wide screen for embryo-inducing genes as a first step towards generating a complete map of genes that can induce embryogenesis in the suspensor.

Results

Testing the activity of embryo inducers in a uniform system

Several genes have previously been reported to trigger embryogenesis when ectopically overexpressed, and have been defined as master embryonic or meristematic regulators (see Chapter 1; reviewed in Ikeuchi et al., 2013). However, their ability to induce embryogenesis has been tested in diverse model systems (Boutilier, 2002; Hecht et al., 2001; Waki et al., 2011; Zuo et al., 2002) and it is therefore difficult to derive if they are part of a coherent network. We therefore decided to use the predictable suspensor-derived embryogenesis as a experimental system and test all known embryo inducer genes, for their ability to convert suspensor cells into embryonic cells.

We selected fifteen genes from literature and divided these in four groups based on the process they regulate (Figure 1A). The first group consists of six genes - *LEAFY COTYLEDON1* (*LEC1*), *LEC2*, *BABY BOOM* (*BBM*), *AGAMOUS-LIKE15* (*AGL15*), *PLETHORA5* (*PLT5* or *EMBRYOMAKER* [*EMK*]) and *RWP-RK DOMAIN-CONTAINING 4* (*RKD4*) – all encoding regulators of re-acquisition of embryonic fate. The second group is composed of regulators of egg cell fate - *RKD1* and *RKD2*. The third group is formed by *WUSCHEL* (*WUS*) and *WUSCHEL-RELATED HOMEODOMAIN-BOX5* (*WOX5*) - regulators in the meristem fate (Ikeuchi et al., 2013; Chapter 1). In the fourth group, we included *SOMATIC EMBRYOGENESIS RECEPTOR KINASE1* (*SERK1*) and its 4 closest homologs *SERK2-5* (Figure 1A). To test if these genes can trigger suspensor to embryo conversion, we misexpressed each in the suspensor using the two-component GAL4/UAS system. In the past, the same approach was used to achieve local inhibition of auxin in discrete domains of the pro-embryo and in the suspensor cells by misexpressing the *bdl/iaa12* protein (Rademacher et al., 2012; Weijers et al., 2003). We fused the coding sequence (CDS) of each of the selected target genes to the GAL4-dependent UAS (upstream activation sequence) promoter, into a binary vector (pPLV132; Wendrich et al., 2015), which was subsequently introduced into the suspensor-specific GAL4/GFP driver line M0171 (Haseloff, 1999). The obtained transgenic lines were afterwards screened for twin seedlings (Figure 1B).

For only two of the selected genes, namely *BBM* and *WUS*, seedlings with altered phenotypes could be retrieved (Figure 1C), mainly affecting the shoot apical meristem (SAM) zone. M0171 >> *BBM* seedlings showed fused wrinkled first leaves, while M0171 >> *WUS* seedlings showed proliferated SAM zone (Figure 1C). This can be explained by the activity of the M0171 driver line in shoot meristem cells during late embryogenesis, as well as post-embryonically (see Figure 1A-D in Chapter 5; Rademacher et al., 2012). At later stages of development, dwarf plants with small leaves were observed for both M0171 >> *BBM* and M0171 >> *WUS*. Importantly, these phenotypes were reminiscent of the ones already described in the literature (Figure 1C and D; Boutilier, 2002; Zuo et al., 2002), and thus validate the biological activity of these genes. Strikingly however, apart from *bdl* (Rademacher et al., 2012), only the *RKD1* gene was able to induce formation of twin seedlings when misexpressed in the suspensor (Figure 2B-F; see next section). This result suggests that the suspensor to embryo transformation is not a general response to local misexpression of a potential embryogenesis inducer. Rather, this fate transition in the suspensor is triggered by a defined set of genetic regulators.

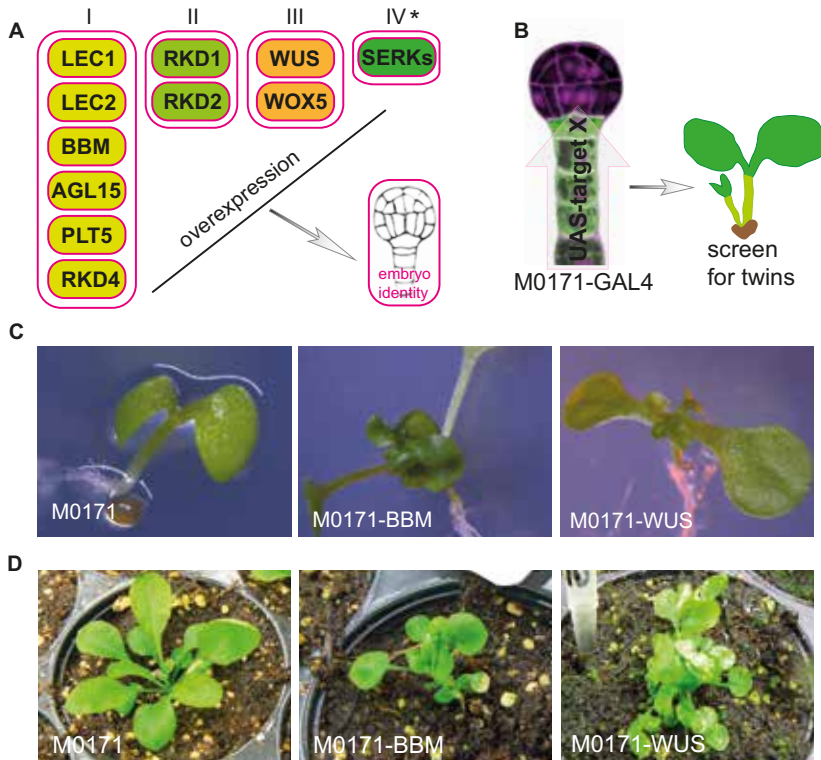


Figure 1: Targeted misexpression of embryo regulators in the suspensor.

(A) Four groups of embryo regulators were selected for local misexpression: I-Embryonic fate regulators (LEC1, LEC2, BBM, AGL15, PLT5 and RKD4), II-Egg cell fate regulators (RKD1 and RKD2), III-Meristem regulators (WUS, WOX5) and IV-SERKs (SERK1, SERK2, BAK1/SERK3, SERK4 and SERK5). All genes in groups I, II and III are able to induce embryo formation upon overexpression, while SERK1 enhances somatic embryogenesis in culture. (B) Scheme of M0171-driven suspensor-specific misexpression. (C) Phenotypes of untransformed M0171-GAL4 seedling, and T1 transgenic seedlings carrying M0171 >>BBM and M0171 >>WUS constructs 7 days after germination. (D) Three-week-old T1 transgenics of M0171 >>BBM and M0171 >>WUS.

RKD1 as a specific inducer of embryogenesis in suspensor cells

The only novel inducer of suspensor to embryo transition, RKD1, is a plant-specific RWP-RK domain-containing transcriptional regulator, which along with RKD2 is involved in regulation of an egg cell-related gene expression program (Koszegi et al., 2011). The *Arabidopsis* genome encodes 14 RWP-RK proteins that can be divided into 2 subfamilies - the NIN-like and RKD subfamily proteins (Schauer et al., 2005). The RKD subfamily is composed of five members: RKD1, RKD2, RKD3, RKD4 and RKD5 (Figure 2L). Based on published activity in promoting aspects of embryo induction (Koszegi et al., 2011; Waki et al., 2011), two other members of the RKD subfamily,

RKD2 and *RKD4*, were also among the selected target genes tested here. Although some dwarf transgenic plants were found (Figure 3A-D), twin seedlings could not be detected upon suspensor-specific misexpression, not even in lines expressing the closest *RKD1* homolog *RKD2* (50 % homology; Figure 2L and M).

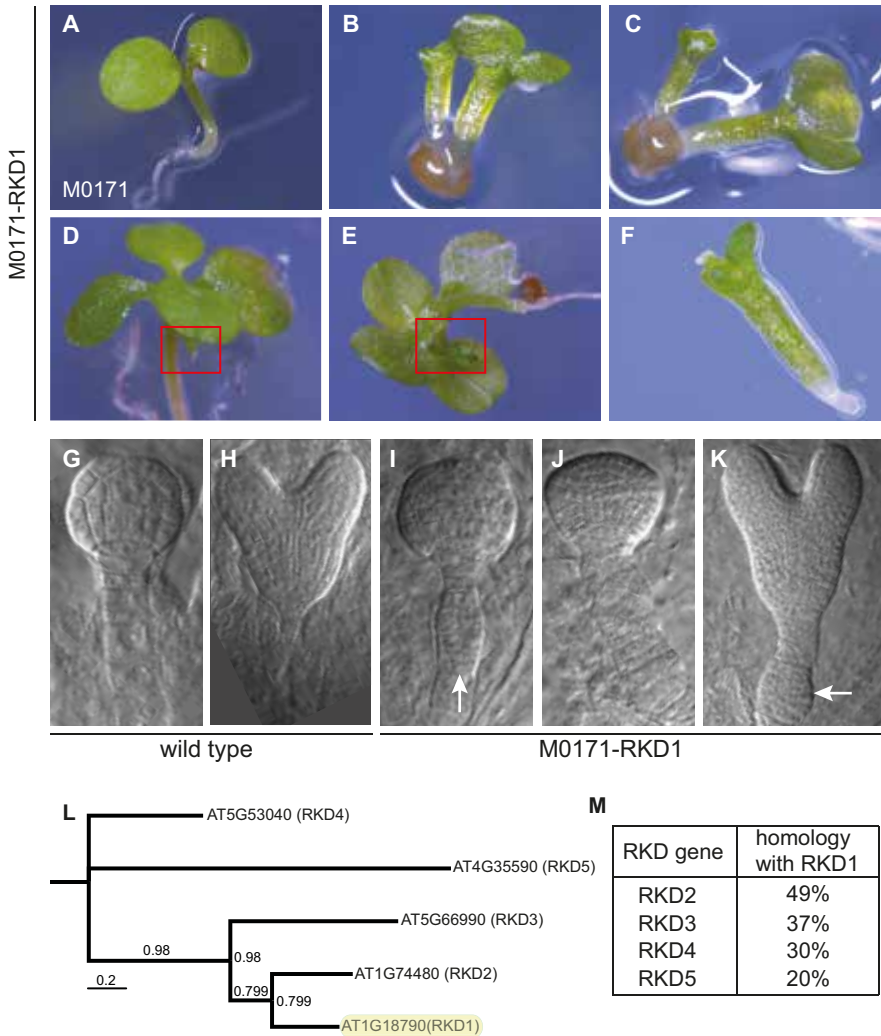


Figure 2: RKD1 induces embryogenesis in suspensor cells.

(A-E) Phenotypes of untransformed M0171-GAL4 seedling (A), M0171 >>RKD1 twin seedlings (B-E) and a separated secondary twin seedling (F). (G-K) Phenotypes of untransformed M0171-GAL4 embryos (G,H) and M0171 >>RKD1 embryos (I-K). In contrast to the single file wild type suspensor, suspensor cells of RKD1-expressing embryos divide aberrantly to form a second embryo. The proliferating suspensor cells are indicated with arrows. (L) Multiple sequence alignment of Arabidopsis RKD proteins. (M) Homology of Arabidopsis RKD genes with RKD1, in percentages (%).

Among the seventeen independent M0171 >>RKD1 transgenic lines that were obtained, five gave twin seedlings (29 % of the primary transformants). Two demonstrated a high penetrance (25%; N > 200) and 3 showed low penetrance (1%; N > 200). While this percentage is low, the twin phenotype was stably inherited in next generations. By comparison, it is important to note that twins are not encountered in wild-type plants. We noticed that in most of the cases in M0171 >>RKD1 lines, one of the twins was smaller than the other (Figure 2B, C and D), indicating a delay in development or later initiation of one of the embryos. Moreover, the seedlings were not always clearly conjoined (Figure 2D and E), or separated after germination

Next, we studied the origin of the second RKD1-induced individual during embryogenesis. Remarkably, there was no precise timing and consistency on when the first aberrant division in the suspensor will arise and which suspensor cell will host the formation of the second embryo (Figure 2I and J). In addition, the growth and development of the two embryos was always asynchronous (Figure 2K), consistent with the observed difference in the twin seedlings' size/developmental stage.

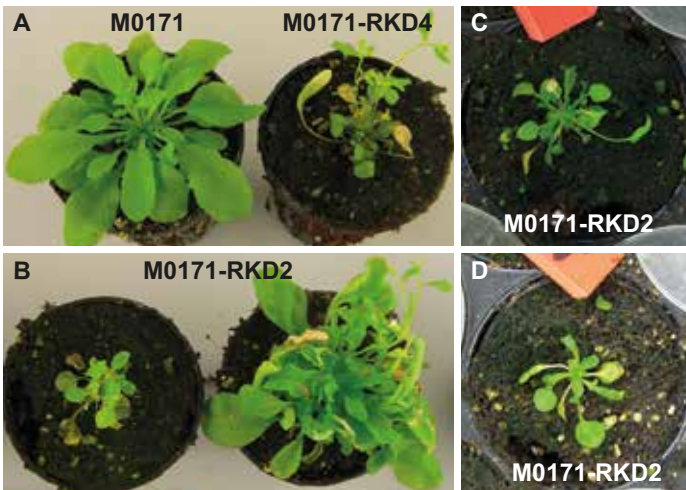


Figure 3: Phenotypes of M0171-driven RKD2 and RKD4 misexpression.

(A-D) Phenotype of untransformed M0171, M0171 >>RKD4 (A) and M0171 >>RKD2 (B-D) six-week-old plants.

In summary, several known embryogenesis regulators were tested for their ability to induce suspensor to embryo transformation. Apart from *bd1*, we found that RKD1 was also capable to induce embryo formation in suspensor cells. The finding that none of the other fifteen genes, including close *RKD1* homologs, could do the same, suggests that the suspensor-derived embryogenesis requires a specific set of genetic regulators.

A local GAL4/UAS activation tagging screen for identification of embryo inducers

Given that the suspensor offers a system in which specific genetic regulators can induce embryogenesis, we aimed to exploit this model to identify novel regulators of embryo induction. These would act in a coherent network with the known *bd1/iaa12* and RKD1 factors. To identify such genes, we here employed a two-component GAL4/UAS activation tagging system. Activation tagging relies on the random insertion of a T-DNA with an outward-oriented promoter in the genome, which can activate nearby genes (Weigel et al., 2000). While a ubiquitous promoter such as 35S can be used (Hayashi et al., 1992; Weigel et al., 2000), local activation tagging can be achieved by using a specific promoter. Such approaches have been successfully used in *Drosophila* (Rørth, 1996) and more recently also in *Arabidopsis* roots (Waki et al., 2013). A schematic overview of the GAL4/UAS activation tagging screen is shown in Figure 4. Instead of using an activation tagging T-DNA with a specific promoter, we used a vector (pBIB-UAS), which harbors five copies of the GAL4-dependent upstream activation sequence (5xUAS) close to the T-DNA left border (Figure 4B; Waki et al., 2013). As host plants, we used the suspensor-specific GAL4/GFP driver line M0171 (Figure 4A). Thus, when the pBIB-UAS T-DNA is inserted next to an endogenous gene, this gene may now be activated in the suspensor by the suspensor specific M0171-GAL4 (Figure 4C). When this gene can induce embryogenesis, one could expect to find twin seedlings (Figure 4C). If the screen is performed at genome-saturating level, we should be able to test the ability of each gene to induce embryogenesis in the suspensor.

We first tested the feasibility of this screen and established conditions for screening. T1 transgenic seedlings were first selected on hygromycin-containing medium, and deviant seedlings were then transferred to medium without hygromycin to restore normal growth. Although it is frequently used as selectable marker in plants (Harrison et al., 2006; Ortiz et al., 1996; Waki et al., 2013; Zheng et al., 1991), hygromycin is known to inhibit protein synthesis and to have an effect on root, hypocotyl and cotyledon growth and development (Harrison et al., 2006; McGuire, 1953), which can “mislead” a mutant selection in such a screen. We therefore analyzed the T2 generation of 17 randomly selected transformants. Of these, 12 lines harbored a single-locus insertion while 2 segregated in an abnormal ratio and 3 lines did not show any resistance to hygromycin (Table 1). Thus this T-DNA mostly inserts as a single copy, and most of the seedlings selected in the screen are true transgenics.

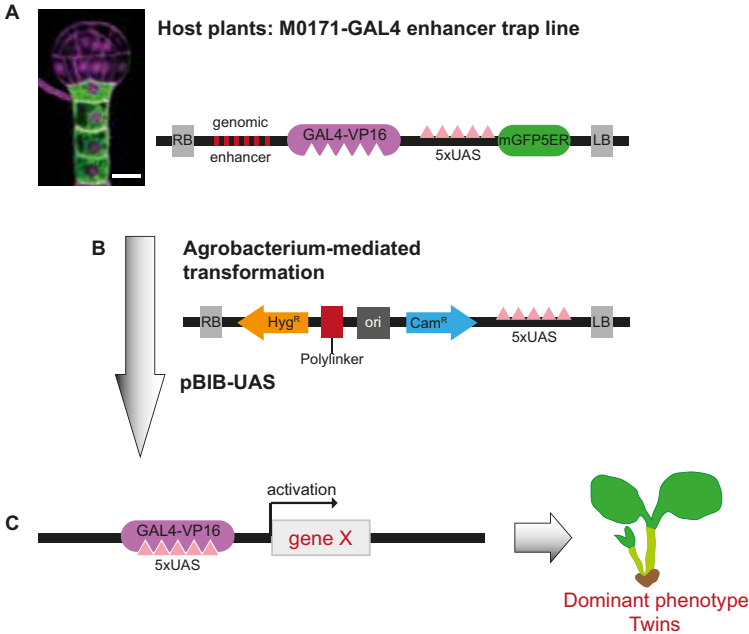


Figure 4: Scheme of GAL4/UAS activation tagging screen.

(A) GFP fluorescence in globular stage embryo of M0171-GAL4 enhancer trap line used in this study. Magenta counterstaining is Renaissance fluorescence. Scale bar represents 10 micrometer. (B) Scheme of the pBIB-UAS tagging vector, which is inserted in host plants via Agrobacterium-mediated transformation. Gene fragments are not drawn to scale. (C) Dominant twin phenotype is expected if the gene next to the insertion can induce embryogenesis. Scheme modified from Waki et al., 2013.

Table 1: Segregation analysis of hygromycin resistance.

HYGROMYCIN SEGREGATION				
line #	hygR	hygS	total	%
96	152	54	206	26
92	150	52	202	26
95	150	50	200	25
11	53	109	162	67
21	128	64	192	33
89	140	60	200	30
90	138	88	226	39
88	170	70	240	29
42	145	80	225	36
87	147	84	231	36
97	150	76	226	34
3	140	73	213	34
14	126	53	179	30
79	107	> 107		nd
91	no R	100% sens		
101	no R	100% sens		
80	no R	100% sens		

We next performed large-scale transformations, and screened a total of 17,321 transgenic seedlings. Among these, we selected 252 mutants (M) showing a range of diverse defects. From these, we selected 13 mutants of which 5 showed a twin-like phenotype (M26, M135, M154, M310, M324 and M340) and 8 that showed a severely distorted shoot meristem phenotype (Table 2; Figure 5B-N) for validation experiments. The twin-like mutants displayed a range of defects.

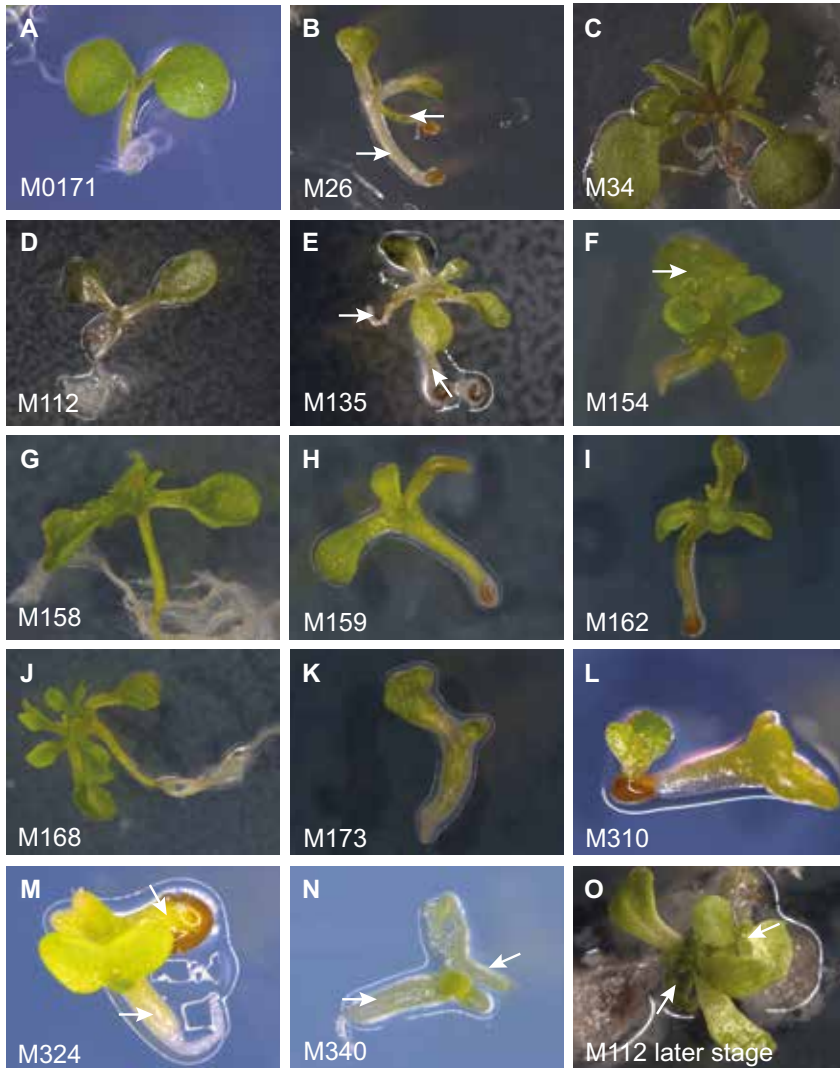


Figure 5: Activation tagging mutants selected for validation.

(A) Untransformed M0171-GAL4 seedling and (B-O) selected T1 mutant seedlings.

Most (M26, M135, M324 and M340) appeared duplicated at some point on the axis (Figure 5B, E, M and N), while one (M154) showed formation of callus-like structures on top of the leaves (Figure 5F) and only 1 mutant (M310) showed a genuine second seedling at the base of the main one (Figure 5L). Among the other 8 mutants, seedlings with proliferative SAM (M34, M158, M159, M162, M168; Figure 5C, G, HI and J) and lacking root (M159, M162 and M173; Figure 5H, I and K) were observed. Interestingly, M112 did not show a very apparent phenotype (Figure 5D), but at later stage there were 2 clearly distinguishable seedlings with cotyledons and leaves, formed on top of the SAM (indicated with arrows, Figure 5O).

Table 2: Overview of the GAL4/UAS activation tagging screen for twins.

Numer of T1 lines screened	17321
Number of T1 lines selected	252
Number of T1 lines with twin-like phenotype	5
Number of lines used for validation	13
Number of lines with reproduced phenotype in T2	0

Identification of tagged genes

To identify the T-DNA insertions potentially responsible for mutant phenotypes, genomic DNA was isolated from the leaves of the primary transformants. Entire mutant seedlings (M154, M173, M310, M159, M324 and M340) were used when no further growth was observed. We next used Thermal Asymmetric Interlaced PCR (TAIL-PCR; Liu et al., 1995) to amplify the T-DNA flanking DNA. In all mutants, a single T-DNA insertion was detected and unequivocal matches in the *Arabidopsis* genome were found. The distance between the insertion point and the start codon of the flanking genes ranged from 180 bp to 7.1 kb. In other activation tagging studies, the activity range of the T-DNA insertion with respect to the distance from the tagged gene was determined to be within 0.8 kb upstream of the first start codon (for 5xUAS promoter; Waki et al., 2013) or between 380 bp to 3.6 kb both upstream and downstream of the tagged gene (for 35S promoter; Weigel et al., 2000). In 5 mutants (M135, M154, M168, M173, M310) the insertion was located within this range, while in 2 mutants (M34 and M324) the distance was larger than 3.6 kb. Interestingly, in 6 of the selected mutants the T-DNA insertion was inserted inside a gene, where in 3 cases (M26, M158, M162) the insertion was found in the opposite coding orientation, and in the other 3 – was in the same orientation (M112, M159, M340). In the following section, we will only discuss the insertion positions of the mutants with the observed twin-like phenotype. A schematic overview of the structures of the different insertion sites is shown in Figure 6 (see also Table S1).

In M26, where ectopic organ formation was observed (Figure 5B), the T-DNA insertion was located in the first intron of AT1G09794 (294 bp downstream of its start codon), coding in the opposite orientation, and 1,670 bp upstream of the start codon of AT1G09790 and in the same coding direction (Figure 6, Table S1). AT1G09794 encodes a Cox19 family protein and AT1G09790 encodes GPI-anchor protein *COBRA-like 6 (COBL6)*. *COBL6* is part of a multigene family in *Arabidopsis* that consist of 11 highly homologous members (Roudier et al., 2002). *COBRA* was recognized as an essential player in the regulation of cell expansion orientation (Schindelman et al., 2001). In this case, the observed phenotype of M26 could be caused by either disrupted function of AT1G09794 by the T-DNA insertion (although being in the opposite direction) or by an activation/misexpression of *COBL6* in the suspensor cells, since the insertion is in the same coding orientation and in a relatively close proximity of its translational start codon.

The T-DNA insert in mutant M135, showing a phenotype similar to M26 (Figure 5E), was in the intergenic region between AT1G28580 and AT1G28590, that both encode GDSL-motif lipases, enzymes with broad substrate specificity. The insertion was located 951 bp upstream of the start codon of AT1G28580, oriented in the same direction, and 178 bp downstream of the stop codon of AT1G28590 (Figure 6, Table S1). An activation-tagged phenotype requires productive contact of enhancer sequences on the T-DNA with neighboring plant sequences (Weigel et al., 2000) and enhancer sequences can function at either 5' or 3' end of a gene (Fang et al., 1989). Thus, in theory, each of the two genes can be tagged, although the insertion is upstream of AT1G28580's ATG and oriented in the same direction.

In mutant M154, which was forming callus-like structures on top of the leaves (Figure 5F, see arrow), the insertion point was found in a region 1,362 bp upstream of the stop codon of AT4G00140 and 735 bp downstream of AT4G00130's stop codon. The T-DNA was in the same coding orientation as AT4G00130 (Figure 6, Table S1), which encodes DNA-binding storekeeper protein-related transcriptional regulator, with unknown function and DUF573 domain. The other gene, AT4G00140, encodes *EMBRYO SAC DEVELOPMENT ARREST 34 (EDA34)*, which as the name suggests was shown to be involved in embryo sac development (Pagnussat et al., 2005). None of these genes is simultaneously downstream of the insertion and oriented in the same direction, which makes it difficult to predict which is more likely to cause the defects in M154.

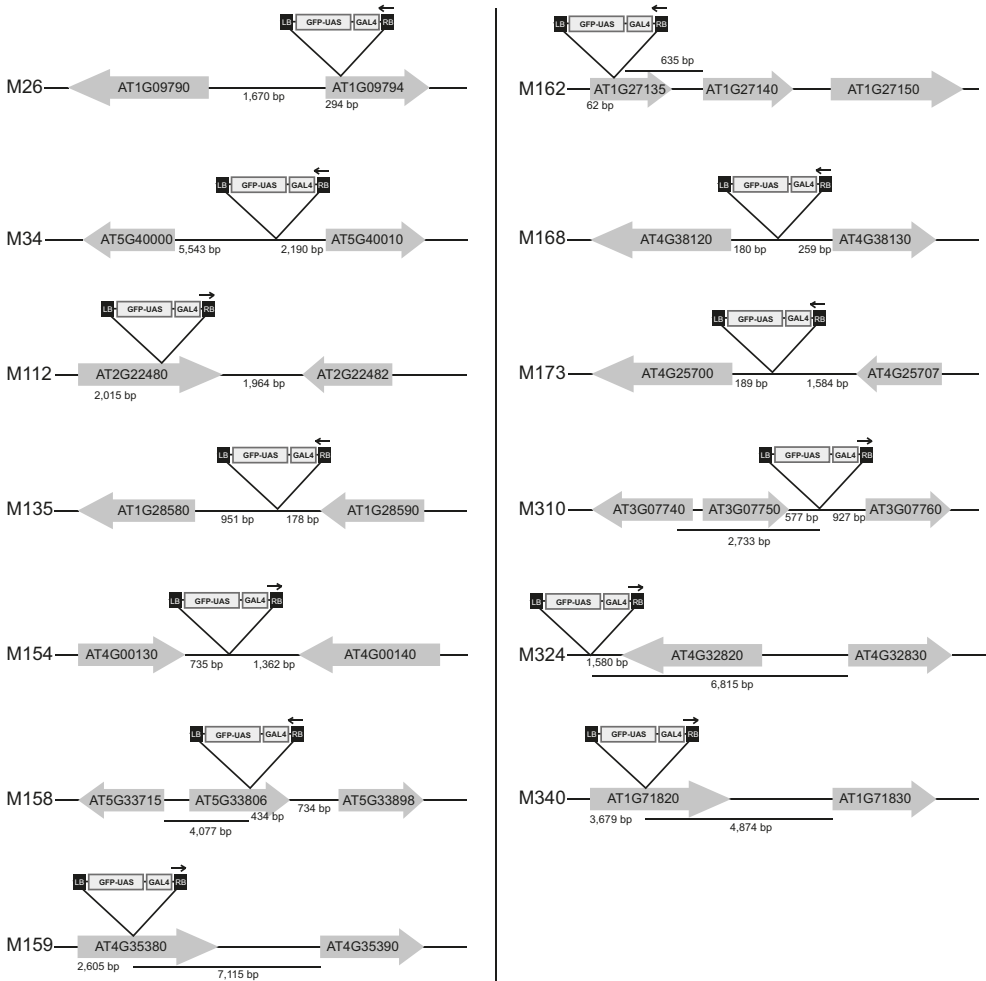


Figure 6: Scheme of the identified T-DNA insertion loci.

Numbers below the T-DNA insertion sites represent the distance (bp) either from the start codon or from the stop codon. Genes are not drawn to scale.

An interesting case and a very good example of a twin phenotype is line M310 (Figure 5L). The T-DNA insertion was positioned 927 bp upstream of the start codon and in the same orientation as AT3G07760, and 577 bp downstream of the stop codon of AT3G07750 in reverse orientation. AT3G07760 and AT3G07750 encode Sterile alfa motif (SAM) domain-containing protein and a 3'-5'-exoribonuclease family protein, respectively. Intriguingly, AT3G07750 shares a relatively small promoter region with AT3G07740, whose start codon is only 2.7 kb away from the insertion site (Figure 6, Table S1). AT3G07740 encodes ADA2A transcriptional adaptor that interacts with

histone acetyltransferase GCN5 homolog and the cold-induced transcription factor CBF1 (Mao et al., 2006). ADA2B, known also as PROPORZ1 (PRZ1) was shown to play a role in the developmental switch from cell proliferation to differentiation in response to variations in hormone (auxin and cytokinin) concentrations (Sieberer et al., 2003). ADA2 proteins contain several domains and most notable among them are Zing-finger and SANT-domains. The latter has a central role in chromatin remodeling by functioning as a unique histone interaction module (Boyer et al., 2004). Furthermore, chromatin-remodeling factors have been shown to play a role in controlling embryonic identity (Ogas et al., 1999; Radoeva and Weijers, 2014; Tsukagoshi et al., 2007).

M324 mutant shows ectopic organ formation (Figure 5M) and the T-DNA insertion was mapped to be 1.5 kb downstream of the stop codon of AT4G32820 and pointing the other direction. AT4G32820 encodes CALCINEURIN BINDING PROTEIN 1 (CABIN1), which is Tetratricopeptide repeat (TPR)-like superfamily protein and it is part of the *Arabidopsis* histone chaperon complex (HIS; Duc et al., 2015). Interestingly, the insertion in M324 was found in the same coding orientation, although 6.8 kb away from AT4G32830 that encodes AURORA1 (AUR1; Figure 6, Table S1). AUR1 is member of a family of Ser/Thr kinases whose activities peak during cell division. It was shown that AUR1 and its homolog AUR2 are involved in lateral root patterning and orienting formative divisions during embryogenesis (Van Damme et al., 2011). Although AUR1 is not within the effective activation tagging distance, discussed above, its relation to embryogenesis makes it a putative candidate.

An ectopic organ formation phenotype was also observed in M340 (Figure 5N). Strikingly, one of the genes adjacent to the T-DNA insertion was *SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1* (*SERK1*; AT1G71830), which encodes a leucine rich-repeat (LRR) transmembrane receptor-like kinase (RLK) shown to increase somatic embryogenesis in culture (Hecht et al., 2001). However, in M340 the insertion was found inside AT1G71820, 3.6 kb downstream of its start codon, which means that it is 4.8 kb upstream of *SERK1* and that the phenotype of M340 could be due to disrupted AT1G71820 function (Figure 6). AT1G71820 encodes SEC6, a member of the exocyst complex that is involved in cytokinesis and cell plate maturation (Fendrych et al., 2010). In summary, unique insertion sites could be identified for each mutant line, and this defines a number of candidate regulators.

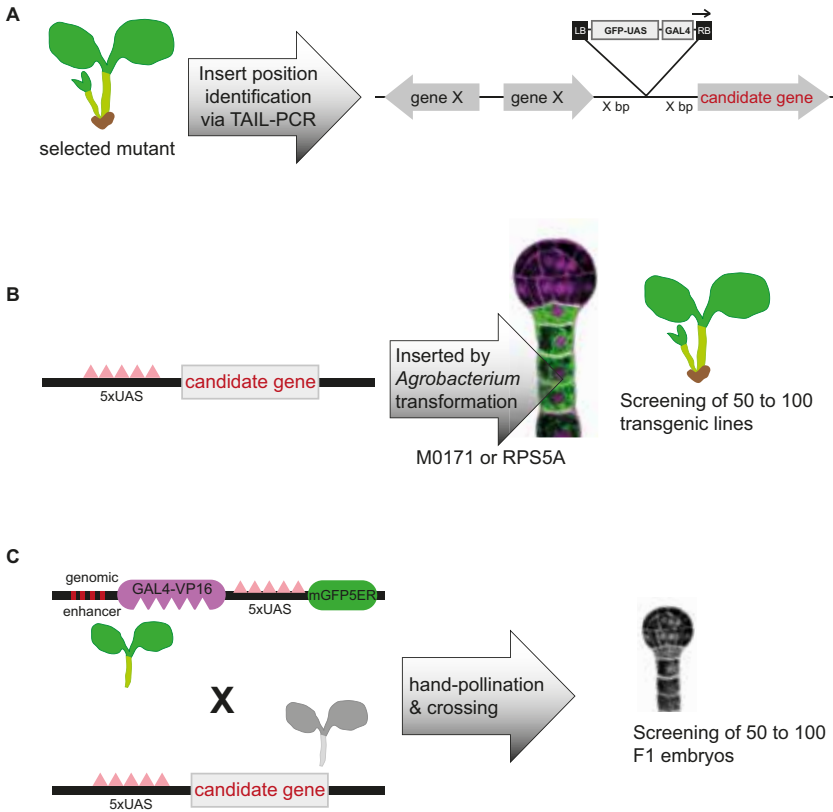


Figure 7: Validation strategy.

(A) Identification of the T-DNA insertion loci. (B) Introduction of UAS-driven coding sequence (CDS) of the candidate gene into M0171-GAL4 and RPS5A-GAL4 lines and subsequent screening of transgenic lines. (C) Crossing of UAS-driven genomic fragment of the candidate gene (introduced in wild type Col-0 background) with M0171-GAL4 and RPS5A-GAL4 lines, and subsequent screening of F1 embryos.

Validation of potential embryogenesis regulators

We designed a validation strategy to confirm the causal relationship between misregulation of the identified tagged genes and the observed phenotype. The hypothesis is that phenotypes are caused by suspensor-specific misexpression of one of the flanking genes. A schematic overview of the strategy is presented in Figure 7.

After identification of the tagged genes (Figure 7A), the coding sequence of each putative candidate was fused to the 5xUAS sequence in a binary vector (pPLV132; Wendrich et al., 2015). These constructs were introduced into the same M0171 enhancer trap line as the one used in the initial screening (Figure 7B). For M26, M34, M135, M168 and M340, transgenic lines were generated for both of the genes flanking the 5xUAS (highlighted in red genes in Table S1). Between 10 and 100

transgenic lines per construct were screened and although some severe phenotypes were occasionally observed (Figure 8A and B; Table 3), none of the newly generated transgenic lines recapitulated the original mutant phenotypes (Figure 5). In addition, we used the ubiquitous expression of other driver line-RPS5A (RIBOSOMAL PROTEIN S5A) to induce stronger overexpression of the candidate gene, but this also did not induce the original phenotypes.

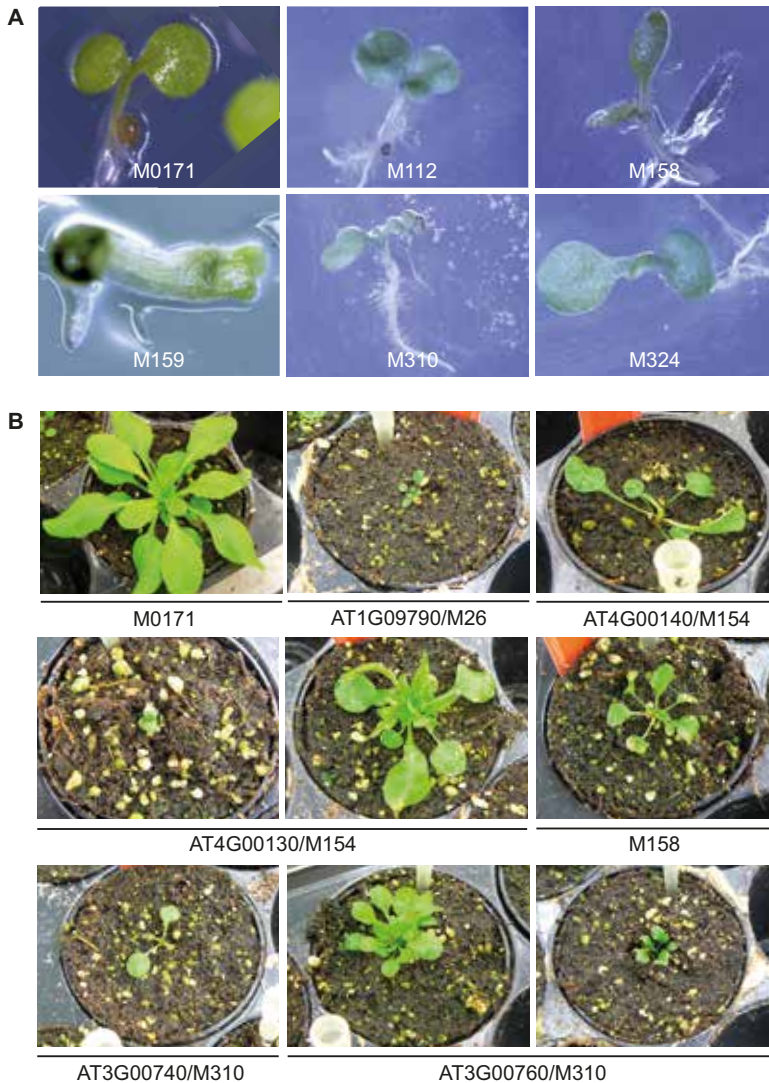


Figure 8: Validation of potential embryogenesis regulators.

(A) Phenotypes of 7-day-old untransformed M0171-GAL4 and T1 seedlings derived from M0171-driven misexpression of candidate genes. (B) Phenotypes of three-week-old untransformed M0171-GAL4 and mutant plants derived from M0171-driven misexpression.

One possibility for the absence of mutant phenotypes could be that to fully mimic the original screen conditions, the full genomic sequence flanking the 5xUAS promoter should be cloned instead of only the downstream open reading frame (ORF). To test this assumption, we fused the full genomic sequence flanking the insertion to the UAS promoter, and introduced these into M0171-GAL4 and RPS5A-GAL4 lines. The newly generated lines were screened for the initially observed phenotypes, but none of the original phenotypic defects could be recapitulated.

Direct transformation of the UAS-driven transgenes in the GAL4-expressing background could induce strong defects that lead to lethality. In that case, the surviving transgenics may not or poorly express the transgene. Therefore, we also generated transgenic lines in Col-0 wild-type background, containing the ORF or the full genomic fragment of the putative tagged genes driven by 5xUAS. Subsequently, three independent T1 lines for each construct were crossed with either M0171 or RPS5A-GAL4 plants (Figure 7C). This strategy allows observation of embryos at early stage of development. Between 50 and 100 F1 embryos were analyzed for each cross, but no phenotypic defects could be observed in suspensor development (data not shown).

Table 3: Number of transgenic lines per construct screened for phenotype.

Line number	Lines		M0171 > > gene X	
	Identity	AGI	# Lines screened	# Phenotypes
M26	COBL6	AT1G09790	16	2
M34	ATPase	AT5G40010	18	0
M112	other RNA	AT2G22482	46	0
	PFK5	AT2G22480	94	1
M135	GDSL-like lipase	AT1G28580	11	0
M154	storekeeper	AT4G00130	44	3
	EDA34	AT4G00140	8	1
M158	DUF3287	AT5G33898	> 105	2
	unknown protein	AT5G33806	39	0
M159	AT-hook protein of GA feedback 1	AT4G35390	49	1
M162	MEG	AT1G27135	50	1
M168	ARM repeat superfamily protein	AT4G38120	not cloned	not cloned
M173	BCH1	AT4G25700	10	1
M310	ADA2A	AT3G07740	90	2
	SAM sterile alfa motif	AT3G07760	89	4
M324	AUR1	AT4G32830	122	2
M340	SERK1	AT1G71830	131	0

In some peculiar cases, for instance M310, two genes can potentially be activated (Figure 6). To further check whether the observed twin phenotype is consequence of the simultaneous activation of both genes, we generated a double transgenic line with RPS5A-ADA2 and RPS5A-SAM constructs, and additionally a line carrying both UAS-ADA2 and UAS-SAM in the M0171 background. The offspring of the crosses was further examined for the observed twin phenotype; alas the latter could not be reproduced (data not shown). In summary, despite a multifaceted validation approach, none of the original mutant phenotypes can be explained by M0171-driven misexpression of a gene flanking the insertion.

Heritability of the mutant phenotypes

In principle, M0171-driven misexpression of a tagged gene should lead to a heritable twinning phenotype. We aimed to test heritability of phenotypes, but due to the severe phenotypes only very few of the selected mutants survived and produced progeny. In fact, from the 13 selected mutants, seeds were obtained only from M26 and M34 (Figure 9B; Suppl. Fig 1B). Five of the remaining eleven mutants (M112, M135, M158, M162 and M168) managed to further grow on soil (Suppl. Figure 1C-E), but were sterile and did not yield seeds. The seedlings of the other 6 mutants were “sacrificed” as material for DNA extraction, which was used for identification of the tagged genes.

In the next (T2) generation of lines M26 and M34, the phenotype and the segregation patterns were assessed. In the case of M26, we found a few embryos (2%; N = 67) with wrongly divided suspensor cells (Figure 9D), reminiscent of the first aberrant suspensor divisions in M0171 >> RKD1. Unfortunately, this phenotype was not stably inherited and did not appear in the next (T3) generation, despite measurable upregulation of the *COBL6* gene in qRT-PCR analysis (Figure 9E). Likewise, the initially reported phenotype of M34 could not be confirmed in the next generation. Thus, for the 2 viable, fertile transgenics isolated in the screen, there is no evidence that the original defect is in fact genetically heritable.

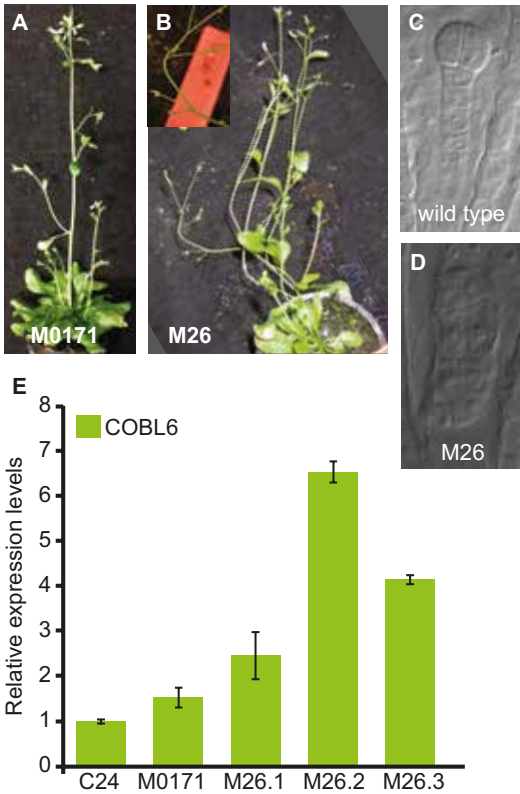


Figure 9: Heritability of the mutant phenotypes.

(A and B) Phenotypes of untransformed M0171-GAL4 (A) and the M26 T1 line (B) six-week-old mature plants. (C and D) Phenotypes of wild type (C) and M26 (D) T2 embryos. (E) Relative expression levels (pRT-PCR) of COBL6 in seedlings. Expression levels in wild type (C24 background) were set to 1. Error bars indicate SE.

Expression pattern analysis of identified tagged genes

Activation tagging can lead to enhanced expression of genes already expressed in the suspensor, or alternatively ectopically activate expression of genes not normally expressed in the suspensor. To study the endogenous expression domains of the genes flanking the mutant insertion sites, we generated transcriptional fusions of each of the genes of interest to nuclear 3xGFP (n3GFP). Promoter-n3GFP transgenics were obtained for 11 genes (Figure 10; Table S2). Multiple independent lines for each gene were analyzed and GFP expression in the root was observed for 7 genes, whose expression was then also studied during embryo development.

Interestingly, during embryogenesis *COBL6/M26* was exclusively and strongly expressed in suspensor cells (Figure 10A-C), matching the expression pattern of M0171 enhancer trap line (see Figure 1A and B in Chapter 5). In the post-embryonic root, the expression of *COBL6/M26* was observed in the vasculature, but only in the elongation zone, and in the distal columella cells (Figure 10D), which are clonal descendants of the suspensor. Another gene detected in the suspensor cells is *ATPase/M34*, but its expression was slightly broader at later stages of embryo

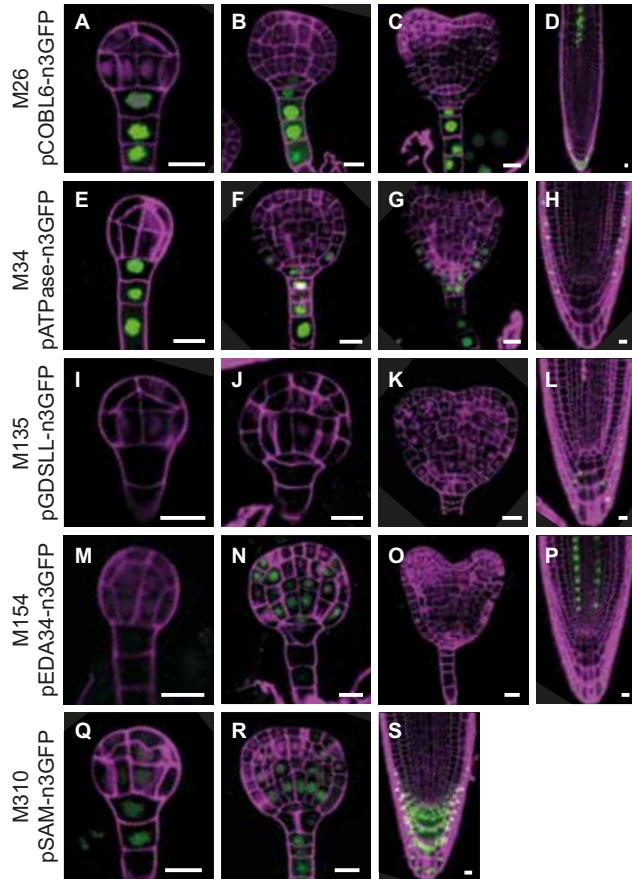


Figure 10: Expression patterns of identified tagged genes.

(A) Expression of promoter-n3GFP reporters for *COBL6* (A-D), *ATPase* (E-H), *GDS-like lipase* (I-L), *EDA34* (M-P) and *SAM* (Q-T) in dermatogen stage (A, E, I, M, Q), globular stage (B, F, J, N, R), heart stage (C, G, K, O, S) embryos and in roots (D, H, L, P, T). Magenta counterstaining is Renaissance fluorescence in (A, B, C, E, F, G, I, J, K, M, N, O, Q, R) and Propidium Iodide in (D, H, L, P, S). All scale bars represent 10 micrometer.

development, including also the basal protoderm cells (Figure 10E-G). In the root, pATPase-n3GFP was detected specifically in the later root cap (LRC) (Figure 10H). GDSL-like lipase/M135 (AT1G28580) could not be detected in the embryo (Figure 10I-K) being preferentially expressed post-embryonically, in columella cells (Figure 10L). *EDA34*/M154 was very specifically expressed in the vasculature during embryogenesis (Figure 10M-O) as well as in the root (Figure 10P). *SAM*/M310 is weakly expressed in the embryo, including the basal tier of the proembryo and the suspensor (Figure 10Q-R). In the root, its expression was observed in the columella, QC, the cells surrounding the QC and LRC (Figure 10S). While two of these genes (M26 and M34) are normally expressed in suspensor cells and could potentially be

enhanced in M0171-driven activation tag lines, the other genes are not normally expressed in the embryo or suspensor at early stages. As such, it is difficult to interpret the identification of these genes in the screen.

Discussion

In plants, there are several pathways leading to embryo formation. These pathways include embryogenesis from somatic cells (in response to exogenous stimulus), apomictic embryogenesis, microspore embryogenesis and extra-embryonic (suspensor) embryogenesis (de Vries et al., 1988; Rademacher et al., 2012; Touraev et al., 1997). The various embryogenic pathways are similar in developmental progression, but differ in the initiation phase where a somatic cell is induced to change its fate towards embryonic cell (Dodeman et al., 1997; Mordhorst et al., 2002). An important fundamental question is what the molecular mechanism involved in this transition is and more notably, whether this path to embryogenesis is controlled by a common gene regulatory network. Alternatively, the embryo initiation response could represent a cellular “default” state used by the cell upon disruption of its homeostasis. In the past two decades, considerable efforts have been made to identify genes that can induce embryogenesis. Most of these genes, however, were tested in diverse experimental systems and it is, thus, impossible to judge if they are all part of the same coherent network. Some regulatory links were shown or suggested between “embryo regulators” (reviewed in Radoeva and Weijers, 2014), but it is unclear to what degree these genes are functionally interconnected.

We here used the predictable suspensor-derived embryogenesis as a uniform model to systematically examine the effect of known embryo inducers on embryo initiation and we demonstrate that, apart from *bdl* (Rademacher et al., 2012), only RKD1 is able to initiate embryo formation in the suspensor. This suggests that suspensor to embryo transformation is not a general response to local misexpression of any potential embryogenesis inducer. Rather, this fate switch requires a defined set of genetic regulators. RKD1 is a member of a small plant-specific transcription factors subfamily that consists of five members (Schäuser et al., 2005). RKD1 and its closest homolog RKD2 (Figure 2L and M) function as regulators of an egg cell-related gene expression program and their ectopic overexpression leads to cell proliferation and callus formation (Koszegi et al., 2011). Interestingly, transient overexpression of another member of the RKD family, RKD4, is sufficient to induce somatic embryogenesis (Waki et al., 2011), but neither RKD2 nor RKD4 possess the newly discovered potential of RKD1 in suspensor cells. RKD1 and RKD2 proteins

share high degree of homology (50%; Figure 2L and M), suggesting that a specific domain of the RKD1 sequence, not present in the other 4 members, is required for the formation of suspensor-derived twin embryos.

In addition, in Chapter 2, we report another transcription factor (bHLH49), whose expression changes during auxin-dependent suspensor to embryo transition and that is also able to induce formation of embryo-like structures in suspensor cells when overexpressed. Hence, so far we have identified three independent factors, *bdl* (Rademacher et al., 2012), bHLH49 (Chapter 2) and RKD1 (this Chapter) that can induce embryo formation in the context of the suspensor. Yet, it is not known if the pathways and the genes involved in the response to each of these factors are conserved. In Chapter 2, a significant overlap between M0171 > > *bdl* responsive and bHLH49-dependent genes was found despite the fact that one was carried out in embryos and the other in roots. Comparison of the transcriptomes of the above-mentioned three factors, generated on suspensor-derived embryos, may reveal whether there is primary set of shared differentially expressed genes that are part of a common embryo program.

Based on our findings, the suspensor appears to be a system in which specific genetic regulators can induce embryogenesis and we next exploited this system to identify novel regulators of embryo induction. We have set up a genome-wide screen in order to find novel embryo inducers and by screening 17,321 transgenic plants, we identified 5 mutants with twin-like phenotype, but only one forming a true secondary seedling on the axis of the main one. Unfortunately none of these lines was both viable and heritable in its phenotype. In addition, none of our attempts to link the observed phenotypes to the identified tagged loci was successful.

One consideration to be taken into account is that hygromycin was used as a selectable marker, which blocks protein synthesis (McGuire, 1953), and even disrupts growth of transgenics, hygromycin-resistant seedlings (Harrison et al., 2006). The long time required to distinguish resistant from sensitive seedlings during selection, coupled to mild toxicity in transgenics may thus contribute to the initially observed phenotypes. This could to some extent explain the absence of phenotypes in our validation experiments, as there would be no genetic basis for the observed defect. Another possibility is that the selected phenotypes display defects that are “freak accidents”, low-frequency stochastic mistakes during early embryogenesis that are caused by abnormal cell divisions in a wild-type background, and are therefore not heritable. Spontaneous twinning in wild type plants is very rare and infrequent phenomenon (Vernon and Meinke, 1994), but as the true twin identified was the only one in more than 17 thousand individuals, we cannot exclude this possibility.

An important consideration in managing expectations from the screen we initiated is the number of lines needed to achieve genomic saturation. Ideally, one would like to identify RKD1, the benchmark gene, and other genes at least twice in the population. To find back RKD1 once with a reasonable chance of approximately 80%, taking into account the size of the *Arabidopsis* genome (120,000 kb), the number of genes (about 30,000) and also the required efficient distance from the promoter (within 1-2 kb from start codon) and orientation of transgene insertion, would be 1 in about 500,000 primary transformants. Thus, from this simple calculation it is evident that much larger transgenic populations would need to be screened to find novel embryogenesis regulators. It is evident that one would need to think of an alternative, more efficient way of selecting transformants than using an antibiotic selection.

In conclusion, we have set up a GAL4-based targeted activation tagging screen motivated to expand the number of genes that are able to induce suspensor derived twin embryos. We have successfully added a new gene to the “collection” by systematically testing known embryonic regulators for their ability to induce embryo formation in the context of the suspensor. Further work is, however, needed to optimize the system towards this goal. In the first place, a better and more efficient way of selecting the transgenic lines will be required in order to avoid “errors” in the initial selection procedure, and to reach saturation.

Materials and Methods

Plant material and growth conditions

The M0171 GAL4/GFP enhancer trap line was generated by Dr. Jim Haseloff in the C24 ecotype (Haseloff, 1999) and was obtained through the Nottingham Arabidopsis Stock Center (NASC). All transcriptional n3GFP fusion lines and the pUAS-gene fusion lines were generated in Columbia (Col-0) ecotype.

Seeds were surface-sterilized in 25% bleach/75% ethanol solution for 10 minutes and were afterwards washed twice with 70% ethanol and once with 100% ethanol. Dried seeds were subsequently plated on half-strength Murashige and Skoog (MS) medium and the appropriate antibiotic (in concentration: 15 mg/l Hygromycin B, 50 mg/l Kanamycin or 15 mg/l Phosphinothricin) for selection of transgenic seeds. After 24 hours incubation at 4°C, the plants were cultured under long-day (16h light, 8h dark) conditions at 22°C.

Cloning

All cloning was performed using the LIC cloning system and the vectors used are previously described (De Rybel et al., 2011). For generating transcriptional fusions, fragments up to 3 kb upstream of the ATG including 5'-UTR were amplified from genomic DNA using Phusion Flash PCR Master Mix (Thermo Scientific) and cloned into vector pPLV(1)04. To generate *pUAS*-fusion lines for M0171-driven misexpression, coding sequences were amplified from genomic DNA or cDNA. All constructs were completely sequenced and transformed into Col-0 wild type Arabidopsis plants by simplified floral dipping (De Rybel et al., 2011). The primers used for cloning are listed in Table S3.

M0171 activation tagging

To perform suspensor-specific activation tagging, the pBIB-UAS construct (Waki et al., 2013) was transformed into the M0171 line by simplified floral dipping (De Rybel et al., 2011). Seeds were plated on half-strength Murashige and Skoog (MS) medium containing 15 mg/l Hygromycin B at a density low enough that seedlings grow individually. After approximately 10 days of growth, resistant seedlings were inspected for twins likely caused during embryogenesis using a stereo-microscope.

Identification of tagged loci

Genomic DNA was isolated from all mutant seedlings using CTAB (Cetyltrimethylammonium bromide) extraction buffer (1% CTAB, 100mM Tris-HCL pH8.0, 20mM EDTA pH8.0, 1.5M NaCl, H₂O) and was afterwards precipitated with isopropanol and washed with 70% ethanol. TAIL-PCR was performed as described previously (Liu et al., 1995) with minor modifications using the specific left and right border primers listed in Table S3, and home-made purified recombinant Taq DNA polymerase. Some PCR products were cloned into pGEM-T vector (Promega) and the PCR products were then sequenced for identification of the T-DNA insertion sites. Database searches were done using the BLAST through the NCBI website (www.ncbi.nlm.nih.gov).

Quantitative RT-PCR analysis

qRT-PCR analysis was performed as previously described (De Rybel et al., 2010). RNA was isolated using TRIzol reagent (Invitrogen) and RNeasy kit (Qiagen). cDNA was prepared from 0.5 µg of total RNA with an iScript cDNA Synthesis Kit (BioRad). qRT-PCR reactions were performed with iQ SYBR Green Supermix (BioRad) and

analyzed on a CFX384 Real-Time PCR detection system (BioRad). Reactions were done in triplicate with three biological replicates. Data were analyzed with qBase software (Hellemans et al., 2007). Primers were designed with Beacon Designer 8 (Premier Biosoft International). Gene expression levels were normalized relative to *CDKA1;1*, *EEFa4* and *GAPC*. Primers for qRT-PCR are listed in Table S3.

Microscopy

Differential interference contrast (DIC) and confocal microscopy were performed as previously described (Llavata-Peris et al., 2013) with minor modifications. For DIC imaging, ovules were isolated in chloral hydrate solution (chloral hydrate, water and glycerol, 8:3:1) and embryos were subsequently observed on a Leica DMR microscope equipped with DIC optics. For confocal imaging of embryos, ovules were isolated in a 4% paraformaldehyde/5% glycerol/1xPBS solution including 1.5% SCRI Renaissance Stain 2200 (R2200; Renaissance Chemicals, UK) for counterstaining of embryos. After squeezing the embryos out of the ovules, R2200 and GFP fluorescence were visualized by excitation at 405 and 488 nm and detection between 430-470 and 500-535 nm, respectively. For confocal imaging of roots, 5 day-old-seedlings were incubated in 10 µg/ml propidium iodide (PI) solution for 1-2 minutes and GFP and PI were visualized by excitation at 488 nm and detection between 500-535 nm and 630-700 nm, respectively. All confocal imaging was performed on a Leica SP5 II system equipped with Hybrid Detectors.

Acknowledgements

We thank Keiji Nakajima for sharing the pBIB-UAS vector used in this study and Jim Haseloff for generating the GAL4 enhancer trap lines. This work was funded by a grant from the Netherlands Organization for Scientific Research (NWO; ALW-NSFC grant 846.11.001) to D.W.

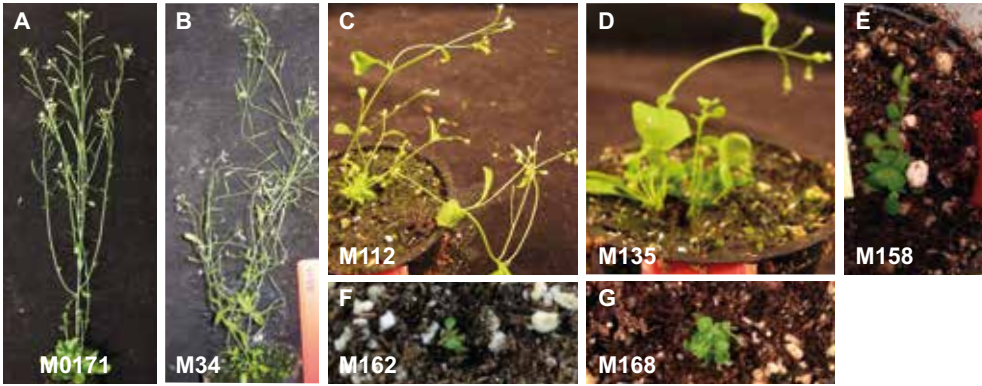
References

- Boutillier, K., Offringa, R., Sharma, V.K., Kirft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.M., van Lammeren, A.A., Miki, B.L., Custers, J.B. and van Lookeren Campagne, M.M. (2002). Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. *Plant Cell* 14, 1737-1749.
- Boyer, L.A., Latek, R.R. and Peterson, C.L. (2004). The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol* 5, 158-163.
- Bozhkov, P.V., Filonova, L.H. and Suarez, M.F. (2005). Programmed Cell Death in Plant Embryogenesis. *Curr Top Dev Biol* 67, 135-179.
- De Rybel, B., van den Berg, W., Lokerse, A., Liao, C.Y., van Mourik, H., Möller, B., Peris, C.L. and Weijers, D. (2011). A versatile set of ligation-independent cloning vectors for functional studies in plants. *Plant Physiol* 156, 1292-1299.
- De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W., Audenaert, D., Van Campenhout, J., Overvoorde, P., Jansen, L., Vanneste, S., 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. *Curr Biol* 20, 1697-1706.
- de Vries, S.C., Booij, H., Meyerink, P., Huisman, G., Wilde, H.D., Thomas, T.L. and van Kammen, A. (1988). Acquisition of embryogenic potential in carrot cell-suspension cultures. *Planta* 176, 196-204.
- Dijk, P., Jong, H., Vijverberg, K. and Biere, A. (2009). An Apomixis-Gene's View on Dandelions. 475-493.
- Dodeman, V., Ducreux G. and Kreis, M. (1997). Zygotic embryogenesis versus somatic embryogenesis. *J Exp Bot* 48, 1493-1509.
- Duc, C., Benoit, M., Le Goff, S., Simon, L., Poulet, A., Cotterell, S., Tatout, C. and Probst, A.V. (2015). The histone chaperone complex HIR maintains nucleosome occupancy and counterbalances impaired histone deposition in CAF-1 complex mutants. *Plant J* 81, 707-722.
- Fang, R.X., Nagy, F., Sivasubramaniam, S. and Chua, N.H. (1989). Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell* 1, 141-150.
- Fendrych, M., Synek, L., Pečenková, T., Toupalová, H., Cole, R., Drdová, E., Nebesářová, J., Šedinová, M., Hála, M., Fowler, J.E. and Žárský, V. (2010). The Arabidopsis Exocyst Complex Is Involved in Cytokinesis and Cell Plate Maturation. *Plant Cell* 22, 3053-3065.
- Garces, H. and Sinha, N. (2009). The 'mother of thousands' (*Kalanchoe daigremontiana*): a plant model for asexual reproduction and CAM studies. *Cold Spring Harb Protoc* 2009, 10: pdb emo133.
- Haccius, B. (1955). Experimentally Induced Twinning in Plants. *Nature* 176.
- Haccius, B. (1978). Question of uncellular origin of non-zygotic embryos in callus cultures. *Phytomorphology* 28, 74-81.
- Harada, J.J., Belmonte, M.F. and Kwong, R.W. (2001). *Plant Embryogenesis (Zygotic and Somatic)*. eLS (John Wiley & Sons, Ltd).
- Harrison, S.J., Mott, E.K., Parsley, K., Aspinall, S., Gray, J.C. and Cottage, A. (2006). A rapid and robust method of identifying transformed *Arabidopsis thaliana* seedlings following floral dip transformation. *Plant Methods* 2, 19.
- Haseloff, J. (1999). GFP variants for multispectral imaging of living cells. *Methods Cell Biol* 58, 139-151.
- Hayashi, H., Czaja, I., Lubenow, H., Schell, J. and Walden, R. (1992). Activation of a plant gene by T-DNA tagging: auxin-independent growth in vitro. *Science* 258, 1350-1353.
- Hecht, V., Vielle-Calzada, J.P., Hartog, M.V., Schmidt, E.D.L., Boutillier, K., Grossniklaus, U. and de Vries, S.C. (2001). The Arabidopsis Somatic Embryogenesis Receptor Kinase 1 Gene Is Expressed in Developing Ovules and Embryos and Enhances Embryogenic Competence in Culture. *Plant Physiol* 127, 803-816.

- Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. and Vandesompele, J. (2007).** qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 8, R19.
- Ikeuchi, M., Sugimoto, K. and Iwase, A. (2013).** Plant Callus: Mechanisms of Induction and Repression. *Plant Cell* 25, 3159-3173.
- Kawashima, T. and Goldberg, R.B. (2010).** The suspensor: not just suspending the embryo. *Trends Plant Sci* 15, 23-30.
- Koltunow, A.M., Bicknell, R.A. and Chaudhury, A.M. (1995).** Apomixis: Molecular Strategies for the Generation of Genetically Identical Seeds without Fertilization. *Plant Physiol* 108, 1345-1352.
- Közegi, D., Johnston, A.J., Rutten, T., Czihal, A., Altschmied, L., Kumlehn, J., Wust, S.E., Kirioukhova, O., Gheyselinck, J., Grossniklaus, U. and Bäumllein, H. (2011).** Members of the RKD transcription factor family induce an egg cell-like gene expression program. *Plant J* 67, 280-291.
- Lakshmanan, K.K. and Ambegaokar, K.B. (1984).** Polyembryony. In *Embryology of Angiosperms*, B. Johri, ed. (Springer Berlin Heidelberg), pp. 445-474.
- Liu, Y., Li, X., Zhao, J., Tang, X., Tian, S., Chen, J., Shi, C., Wang, W., Zhang, L., Feng, X. and Sun, M.X. (2015).** Direct evidence that suspensor cells have embryogenic potential that is suppressed by the embryo proper during normal embryogenesis. *Proc Natl Acad Sci USA* 112, 12432-12437.
- Liu, Y.G., Mitsukawa, N., Oosumi, T. and Whittier, R.F. (1995).** Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* 8, 457-463.
- Llavata-Peris, C., Lokerse, A., Moller, B., De Rybel, B. and Weijers, D. (2013).** Imaging of phenotypes, gene expression, and protein localization during embryonic root formation in *Arabidopsis*. *Methods Mol Biol* 959, 137-148.
- Mao, Y., Pavangadkar, K.A., Thomashow, M.F. and Triezenberg, S.J. (2006).** Physical and functional interactions of *Arabidopsis* ADA2 transcriptional coactivator proteins with the acetyltransferase GCN5 and with the cold-induced transcription factor CBF1. *Biochim Biophys Acta* 1759, 69-79.
- McGuire, P. (1953).** Hygromycin I. Preliminary studies on the production and biological activity of a new antibiotic. *Antibiot Chemother* 3, 1268-1278.
- Mordhorst, A.P., Hartog, M.V., El Tamer, M.K., Laux, T. and de Vries, S.C. (2002).** Somatic embryogenesis from *Arabidopsis* shoot apical meristem mutants. *Planta* 214, 829-836.
- Mordhorst, A.P., Toonen, M.A.J., de Vries S.C. and David, D.M. (1997).** *Plant Embryogenesis*. *Crit Rev Plant Sci* 16, 535-576.
- Ogas, J., Kaufmann, S., Henderson, J. and Somerville, C. (1999).** PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in *Arabidopsis*. *Proc Natl Acad Sci USA* 96, 13839-13844.
- Ortiz, J.P., Reggiardo, M.I., Ravizzini, R.A., Altabe, S.G., Cervigni, G.D., Spitteler, M.A., Morata, M.M., Elias, F.E. and Vallejos, R.H. (1996).** Hygromycin resistance as an efficient selectable marker for wheat stable transformation. *Plant Cell Rep* 15, 877-881.
- Pagnussat, G.C., Yu, H.J., Ngo, Q.A., Rajani, S., Mayalagu, S., Johnson, C.S., Capron, A., Xie, L.F., Ye, D. and Sundaresan, V. (2005).** Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development* 132, 603-614.
- Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Freire Rios, A., Borst, J.W., Lukowitz, W., Jurgens, G. and Weijers, D. (2012).** Different auxin response machineries control distinct cell fates in the early plant embryo. *Dev Cell* 22, 211-222.
- Radoeva, T. and Weijers, D. (2014).** A roadmap to embryo identity in plants. *Trends Plant Sci* 19, 709-716.
- Rørth, P. (1996).** A modular misexpression screen in *Drosophila* detecting tissue-specific phenotypes. *Proc Natl Acad Sci USA* 93, 12418-12422.

- Roudier, F., Schindelman, G., DeSalle, R. and Benfey, P.N. (2002). The COBRA Family of Putative GPI-Anchored Proteins in Arabidopsis. A New Fellowship in Expansion. *Plant Physiol* 130, 538-548.
- Schauser, L., Wieloch, W. and Stougaard, J. (2005). Evolution of NIN-Like Proteins in Arabidopsis, Rice, and *Lotus japonicus*. *J Mol Evol* 60, 229-237.
- Schindelman, G., Morikami, A., Jung, J., Baskin, T.I., Carpita, N.C., Derbyshire, P., McCann, M.C. and Benfey, P.N. (2001). COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in Arabidopsis. *Genes Dev* 15, 1115-1127.
- Schmidt, E.D., Guzzo, F., Toonen, M.A. and de Vries, S.C. (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* 124, 2049-2062.
- Schwartz, B.W., Yeung E.C. and Meinke D.W. (1994). Disruption of morphogenesis and transformation of the suspensor in abnormal suspensor mutants of Arabidopsis. *Development* 120, 3235.
- Sieberer, T., Hauser, M.T., Seifert, G.J. and Luschnig, C. (2003). PROPORZ1, a Putative Arabidopsis Transcriptional Adaptor Protein, Mediates Auxin and Cytokinin Signals in the Control of Cell Proliferation. *Curr Biol* 13, 837-842.
- Touraev, A., Vicente, O. and Heberle-Bors, E. (1997). Initiation of microspore embryogenesis by stress. *Trends Plant Sci* 2, 297-302.
- Tsukagoshi, H., Morikami, A. and Nakamura, K. (2007). Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in Arabidopsis seedlings. *Proc Natl Acad Sci USA* 104, 2543-2547.
- Van Damme, D., De Rybel, B., Gudesblat, G., Demidov, D., Grunewald, W., De Smet, I., Houben, A., Beeckman, T. and Russinova, E. (2011). Arabidopsis α Aurora Kinases Function in Formative Cell Division Plane Orientation. *Plant Cell* 23, 4013-4024.
- Vernon, D. and Meinke, D. (1994). Embryogenic transformation of the suspensor in twin, a polyembryonic mutant of Arabidopsis. *Dev Biol* 165, 566-573.
- Waki, T., Hiki, T., Watanabe, R., Hashimoto, T. and Nakajima, K. (2011). The Arabidopsis RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. *Curr Biol* 21, 1277-1281.
- Waki, T., Miyashima, S., Nakanishi, M., Ikeda, Y., Hashimoto, T. and Nakajima, K. (2013). A GAL4-based targeted activation tagging system in Arabidopsis thaliana. *Plant J* 73, 357-367.
- Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E.J., Neff, M.M., Nguyen, J.T., Sato, S., Wang, Z.Y., Xia, Y., Dixon, R.A., Harrison, M.J., Lamb, C.J., Yanofsky, M.F. and Chory, J. (2000). Activation tagging in Arabidopsis. *Plant Physiol* 122, 1003-1013.
- Weijers, D., Van Hamburg, J.P., Van Rijn, E., Hooykaas, P.J. and Offringa, R. (2003). Diphtheria toxin-mediated cell ablation reveals interregional communication during Arabidopsis seed development. *Plant Physiol* 133, 1882-1892.
- Wendrich, J.R., Liao, C.Y., van den Berg, W.A., De Rybel, B. and Weijers, D. (2015). Ligation-independent cloning for plant research. *Methods Mol Biol* 1284, 421-431.
- Yeung, E.C. and Meinke, D.W. (1993). Embryogenesis in Angiosperms: Development of the Suspensor. *Plant Cell* 5, 1371-1381.
- Zhang, J.Z. and Somerville, C.R. (1997). Suspensor-derived polyembryony caused by altered expression of valyl-tRNA synthetase in the *twn2* mutant of Arabidopsis. *Proc Natl Acad Sci* 94, 7349-7355.
- Zheng, Z., Hayashimoto, A., Li, Z. and Murai, N. (1991). Hygromycin Resistance Gene Cassettes for Vector Construction and Selection of Transformed Rice Protoplasts. *Plant Physiol* 97, 832-835.
- Zuo, J., Niu, Q.W., Frugis, G. and Chua, N.H. (2002). The WUSCHEL gene promotes vegetative-to-embryonic transition in Arabidopsis. *Plant J* 30, 349-359.

Supplementary information



Supplemental Figure 1: Phenotypes of selected primary transformants.

(A-G) Phenotypes of untransformed M0171-GAL4 (A) and T1 transgenic (B-G) 6-week-old plants.

Table S1: Overview of the identified insertion sites and flanking genes.

Line number	Insertion position	Chromosomal position	Strand	Flanking gene(s)	Other name/Short gene description
M26	- 1670 bp upstream of ATG + 294 bp downstream of ATG	3,172,489	Complement	AT1G09790	COBL6
M34	- 5,543 kb upstream of ATG 2190 bp downstream of TAA	16,018,027	Complement	AT5G40000	Cox19 family protein P-loop containing nucleoside triphosphate hydrolases superfamily protein
M112	- 1964 bp upstream of TGA + 2015 bp downstream of ATG	9,547,685	Forward	AT2G22482	Encodes Mitochondrial ATPase involved in seed and silique development other RNA
M135	- 951 bp upstream of ATG - 178 bp upstream of TGA	10,047,330	Complement	AT1G28580	PFK5; Phosphofructokinase 5 GDSL-like lipase
M154	+ 735 bp downstream of TGA of AT4G00130	49,804	Forward	AT4G00130	GDSL-like lipase DNA-binding storekeeper protein-related transcriptional regulator
M158	- 1362 bp upstream of TGA of AT4G00140 - 734 bp upstream of ATG + 2026 bp downstream of ATG	12,786,899	Complement	AT4G00140	EDA34-embryo sac development arrest 34
M159	+ 2605 bp downstream of ATG - 7,115 bp upstream of ATG	16,822,488	Forward	AT5G33898	DUF3287
M162	- 632 bp upstream of ATG + 62 bp downstream of ATG	9,427,225	Forward	AT5G33806	unknown protein
M168	- 180 bp upstream of ATG 252 bp downstream of TAA	17,896,240	Complement	AT5G33715	unknown protein
M173	- 189 bp downstream 1584 bp downstream of TAA	13,096,055	Forward	AT4G35380	SEC7-like
M310	- 2733 bp upstream - 927 bp upstream	2,475,796	Forward	AT4G35390 AT1G27140 (150) AT1G27135	AT-hook protein of GA feedback 1 GST13
			Complement	AT4G38120	ARM repeat superfamily protein
			Complement	AT4G38130	HD1- encode histone deacetylase that enhances ATERF7 transcriptional repression
			Complement	AT4G25700	BCH1
			Complement	AT4G25707	small protein
			Complement	AT3G07740	ADA2A
			Forward	AT3G07760	SAM sterile alfa motif

M324	- 6815 bp upstream of ATG	Forward	AT4G32830	AURI	
	- 1,580 bp upstream of TCT	Complement	AT4G32820	Tetratricopeptide repeat (TPR)-like superfamily protein	
M340	- 4,874 bp upstream of ATG	Forward	AT1G71830	SERK1	
	+ 3,679bp downstream of ATG (inside)	Forward	AT1G71820	SEC6	

Table S2: Expression patterns overview.

AGI- Arabidopsis Genome Initiative identification number

Promoter - n3GFP lines						
Line number	AGI	Other name	Embryo expression	Root Expression	Internal line #	
M26	AT1G09790	COBL6	suspensor	vascular in the elongation zone, columella	P18	
M34	AT5G40000	ATPase	suspensor; basal protoderm	Lateral root cap (LRC)	P19	
M112	AT2G22480	phosphofructokinase	no expression	no expression	P1	
M135	AT1G28580	GDS like lipase	no expression	columella	P2	
M154	AT4G00140	EDA34	weak in vascular cells	vasculature	P4	
M158	AT5G33898	DUF 3287	-	-	P6	
	AT5G33806	unknown		root	P5	
M159	AT4G35380	SEC7	n.a.	no expression	P7	
M162	AT1G27140	glutamate s transferase	no expression	no expression	P10	
	AT1G27135	MEG	-	-	P9	
M168	AT4G38120	ARM repeat		root	P11	
M173	AT4G25700	Beta catotene hydroxylase	-	-	P12	
	AT3G07740	ADA2	-	-	P13	
M310	AT3G07760	SAM	suspensor; later basal tier proembryo	calumella, LRC	P14	
M324	AT4G32820	Tetratricopeptide repeat	no expression	no expression	P16	

n.a. not analysed

Table S3: Primers used in this study.

AT number	Gene name	Line number	Primer orientation	Sequence
Primer pairs used for transcriptional fusions				
AT1G09790	COBL6	M26	Sense	TAGTTGGAATGGGTTTCGAAATTCCTCTCGCTCTCGCTTC
			Antisense	TTATGGAGTTGGGTTTCGAAATTTGTCTCACAGACGATGA
AT5G40010	ASD	M34	Sense	TAGTTGGAATGGGTTTCGAAATACGATACAACTATCTAT
			Antisense	TTAIGGAGTTGGGTTTCGAAATTTTGGATGCTCIGTTCG
AT2G22480	PFK5	M112	Sense	TAGTTGGAATGGGTTTCGAAATACATAACTGGATGACTGAT
			Antisense	TTATGGAGTTGGGTTTCGAAAGCGGAAATAGAAGATCCGA
AT1G28580	GDSL	M135	Sense	TAGTTGGAATGGGTTTCGAAAGAGACAAAGAGAAAGATGGA
			Antisense	TTATGGAGTTGGGTTTCGAAATGCCAGAGAGCTGAAGGTTG
AT4G00140	EDA34	M154	Sense	TAGTTGGAATGGGTTTCGAAATTTCCACCGGAGTTTCAAT
			Antisense	TTATGGAGTTGGGTTTCGAAATGTTAGCTAATTTAGGAATTTG
AT5G33806	unknown protein	M158	Sense	TAGTTGGAATGGGTTTCGAAATTTGGGATGCATCGGGCGTGT
			Antisense	TTAIGGAGTTGGGTTTCGAAATTTGGTTCAGAGATGATGGG
AT4G35380	SEC7	M159	Sense	TAGTTGGAATGGGTTTCGAAACAGCCGACCATGAACCACAT
			Antisense	TTATGGAGTTGGGTTTCGAAACGCGGAAATCAAAGTTT
AT1G27140	Glutamate S transferase	M162	Sense	TAGTTGGAATGGGTTTCGAAATTCCTTTTATTAGAGAATA
			Antisense	TTATGGAGTTGGGTTTCGAAAGTTTCTCTTGGTTTGTG
AT4G38120	ARM repeat superfamily	M168	Sense	TAGTTGGAATGGGTTTCGAAAGACTCGGAGCTTCTAATTTCTTG
			Antisense	TTATGGAGTTGGGTTTCGAAATTTCACTCCCGGAAAA
AT4G25700	β -carotene hydroxylase	M173	Sense	TAGTTGGAATGGGTTTCGAAAGGTTTCAATGGAGAAGAT
			Antisense	TTAIGGAGTTGGGTTTCGAAACCTAATGGAAGGAGGAGCTA
AT3G07760	SAM	M310	Sense	TAGTTGGAATGGGTTTCGAACTTTGAGTCTCAGAGAATC
			Antisense	TTATGGAGTTGGGTTTCGAACTTTCCCTTCGCACAAAAT
AT4G32820	Tetrapeptide repeat	M324	Sense	TAGTTGGAATGGGTTTCGAAAGGAAGAAGAGATCAACCCAA
			Antisense	TTATGGAGTTGGGTTTCGAACTTTGATACTCGAAGCTGGTAATG

Primer pairs used for CDS amplification

AT1G09790	COBL6	M26	Sense	TAGTTGAATAGGTTCAATGGGTGCAATGCTAAATCTT
			Antisense	AGTATGGAGTTGGGTTCCCTCAAACCTAGTAGAAGAAGAT
AT5G40010	ASD	M34	Sense	TAGTTGAATAGGTTCAATGGCGTATCCCGGTTCCGCT
			Antisense	AGTATGGAGTTGGGTTCCCTATGTGTGAGATGACCCAC
AT2G22480	PFK5	M112	Sense	TAGTTGAATAGGTTCAATGGATGCTCTTTCTCAGGCG
			Antisense	AGTATGGAGTTGGGTTCCCTTAGATGAAATCGGGTTGGC
AT2G22482	Other_rna	M112	Sense	TAGTTGAATAGGTTCAATGAACTTTCTGAAAGCAA
			Antisense	AGTATGGAGTTGGGTTCCCTAATTTAGCCAGGTGGTTA
AT1G28580	GDSL	M135	Sense	TAGTTGAATAGGTTCAATGGCGTATCCCGGTTCCGCT
			Antisense	AGTATGGAGTTGGGTTCCCTATGTGTGAGATGACCCAC
AT4G00130	DUF573	M154	Sense	TAGTTGAATAGGTTCAATGATGAGGTTAAGAAGATTGC
			Antisense	AGTATGGAGTTGGGTTCCCTCAAGATTCAGCTTTGAACA
AT4G00140	EDA34	M154	Sense	TAGTTGAATAGGTTCAATGGCGGAGGGGAGAATCGGAG
			Antisense	AGTATGGAGTTGGGTTCCCTCACCGTGGACCTTCCCGTT
AT5G33806	unknown protein	M158	Sense	TAGTTGAATAGGTTCTTCTAAAGCAAAAGTTTGGAAAG
			Antisense	AGTATGGAGTTGGGTTCCCTCAGAGAGGGAGCGAAGCTC
AT5G33898	DUF3287	M158	Sense	TAGTTGAATAGGTTCAATGGAAAACAAGGAGCCTCC
			Antisense	AGTATGGAGTTGGGTTCCCTTAGACTCCTTGGCTAGCGA
AT4G35380	SEC7	M159	Sense	TAGTTGAATAGGTTCAATGTCAAAGCTCACAAAACCTCCG
			Antisense	AGTATGGAGTTGGGTTCCCTAAGCCAAAATAGCAACCAATGCAC
AT4G35390	AHL25	M159	Sense	TAGTTGAATAGGTTCAATGTCTAGTTATATGCAACCC
			Antisense	AGTATGGAGTTGGGTTCCCTAAAAGCGGGTCTAGTCCG
AT1G27135	MEG	M162	Sense	TAGTTGAATAGGTTCAATGAAATTCGGCAAAATAGCT
			Antisense	AGTATGGAGTTGGGTTCCCTTAAATCTCTTTGAAATATCC
AT1G27140	GST13	M162	Sense	TAGTTGAATAGGTTCACTGAATTAGTAAAGCAAAG
			Antisense	AGTATGGAGTTGGGTTCCACCTCTCGACGGTTGGCAT
AT4G38120	ARM repeat superfamily	M168	Sense	TAGTTGAATAGGTTCAATGTTACAGGGGGCAGCTTCG
			Antisense	AGTATGGAGTTGGGTTCCCTCAATTCACATTCATCTAAC
AT4G25700	β -carotene hydroxylase	M173	Sense	TAGTTGAATAGGTTCAATGGGGCAGGACTCTCAACC
			Antisense	AGTATGGAGTTGGGTTCCCTCAAGAACTCGAAGCTCGACC
AT3G07740	ADA2A	M310	Sense	TAGTTGAATAGGTTCAATGGGTGCTTCGAAACTAGCT
			Antisense	AGTATGGAGTTGGGTTCCCTCATGTTGAGTCACCTATTCC

AT3G07760	SAM	M310	Sense	TAGTTGAATAGGTTTCATGACCAGAGAACAATCCTCCC
			Antisense	AGTATGGAGTTGGGTTCTCAAGATTCAAGCTTGAGAG
AT4G32820	Tetrapeptide repeat	M324	Sense	TAGTTGAATAGGTTCAATGTTTCGATTCGACGGATC
			Antisense	AGTATGGAGTTGGGTTCTTATAAATGGTGTGCAGGTG
AT4G32830	AUR1	M324	Sense	TAGTTGAATAGGTTCAATGGGATCCCTACGGAGACA
			Antisense	AGTATGGAGTTGGGTTCTTAAACTCTGTAGATTCCAG
Primer pairs used for amplification of the genomic fragment from the insertion point				
AT1G09790	COBL6	M26	Sense	TAGTTGAATAGGTTCTCATAAAGTCATAAAGAAAAACA
			Antisense	AGTATGGAGTTGGGTTCTCAAACCTAGTAGAAGAAGAT
AT1G28580	GDSL	M135	Sense	TAGTTGAATAGGTTCTGTATCCTTAAATTTGTTGGACC
			Antisense	AGTATGGAGTTGGGTTCCCTATGTGTCAGATGACCCAC
AT4G00130	DUF573	M154	Sense	TAGTTGAATAGGTTCTACTTCGGGAGAACTCTACGAAGC
			Antisense	AGTATGGAGTTGGGTTCCCTAAGAAGAGTCTGATACAT
AT4G00140	EDA34	M154	Sense	TAGTTGAATAGGTTCCCCCAATTTCTTGATTTAAG
			Antisense	AGTATGGAGTTGGGTTCTCAGGGTTGGGAGTTGTGGT
AT5G33806	unknown protein	M158	Sense	TAGTTGAATAGGTTCTTCTAAAGCAAAAGTTGGAAG
AT5G33898	DUF3287		Antisense	AGTATGGAGTTGGGTTCTTACTAGACTCCTTGGGCTAGCGA
AT4G35380	SEC7	M159	Sense	TAGTTGAATAGGTTCAATGTCAAAGTCACAAAACCCCTCG
AT4G35390	AHL25		Antisense	AGTATGGAGTTGGGTTCCCTAAAACGGGGTCTAGTCCG
AT1G27135	MEG	M162	Sense	TAGTTGAATAGGTTTCATGAAATTTATCGCAAATAGCT
AT1G27140	GST13		Antisense	AGTATGGAGTTGGGTTCCACCTCTTCGACGGTTGGCAT
AT4G38120	ARM repeat superfamily	M168	Sense	TAGTTGAATAGGTTCTGAACATCATCAACTAGTGAAG
			Antisense	AGTATGGAGTTGGGTTCTCAAATTCACATTTGCTATCTAAC
AT4G25700	β -carotene hydroxylase	M173	Sense	TAGTTGAATAGGTTCCGGGTGTTCAATGGAGAAGATG
			Antisense	AGTATGGAGTTGGGTTCTCAAAGAACTCGAACTCGACC
AT4G32830	AUR1	M324	Sense	TAGTTGAATAGGTTCCGCAAACTCATTAGCAGGGCAA
			Antisense	AGTATGGAGTTGGGTTCTTAAACTCTGTAGATTCCAG
AT1G71830	SERK1	M340	Sense	TAGTTGAATAGGTTCCGTAGTCGGAGTTATTGAC
			Antisense	AGTATGGAGTTGGGTTCTTACCTTGGACCAGATAACTC

Primer pairs used for qRT-PCR			
AT1G09790	COBL6	M26	Sense ATCCAGATGAGGAAGTAT Antisense CACCACTAAAGTTCCAATT
Reference gene	CDK1;1		Sense ATTGGGTATTGCCACTCTCATAGG Antisense TCCTGACAGGGATACCGAATGC
Reference gene	EEF α 4		Sense CTGGAGGTTTTGAGGCTGGTAT Antisense CCAAGGCTAAAGCAAGAAGA
Reference gene	GAPC		Sense GAAGGTTGGTGCCCAAGAAGGTT Antisense AGGGGAGCAAGGCAGTTAGTGG
Primers used for TAIL-PCR			
UAS-TAIL1		Specific I	GAATGGCAGAAAATTCAGCTTGGC
UAS-TAIL2		Specific II	CCGGCTCGTATGTTGTGGAA
UAS-TAIL3		Specific III	CATGCCTGCAGGTCGACAGATCTC
AD1		Arbitrary	NGTCGASWGANA WGAA
AD2		Arbitrary	WGTGNAGWANCANAGA

Chapter 4

Optimization of suspensor-specific activation tagging screen

Tatyana Radoeva, Catherine Albrecht,
Sacco de Vries and Dolf Weijers

Laboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703HA
Wageningen, the Netherlands

The life cycle of plants can be initiated sexually by the formation of an embryo from a single fertilized cell, or asexually when embryo formation is initiated from other cell types. A key, however unanswered, fundamental question is what molecular mechanisms underlie plant embryo initiation, and how these mechanisms are leveraged to initiate alternative embryogenesis. Using auxin-dependent suspensor embryogenesis as a uniform model, we showed that this process is complex and likely requires defined genetic networks. To identify a complete set of genes that can induce embryogenesis, we established a systematic targeted misexpression screen using the two-component GAL4/UAS system. In an initial screen using selection based on hygromycin, the workflow was optimized to isolate twin seedlings following suspensor-specific activation tagging, but no heritable twin mutants were identified. Here, we redesigned the activation tagging vector to use red seed fluorescence for transgenic selection, and next developed a fluorescent seed sorting procedure. This procedure now allows efficient large scale transformation, purification of transgenic seeds and phenotypic screening on media lacking additives. In a pilot screen of 63,000 primary transformants, we isolated 7 true twin seedlings. For one, we confirmed that twinning was genetically heritable, dominant, and caused by excessive division of suspensor cells. Thus, this new generation screen will now allow to saturate the genome with insertions and systematically identify all genes whose activation can trigger embryogenesis in suspensor cells.

Introduction

Seeds are important for plant dispersal and reproductive success. The seed contains an embryo, a new miniature plant that is formed during embryogenesis after fertilization of the egg cell. Zygotic embryogenesis is the most prevalent type of plant embryogenesis, but several other forms of embryogenesis are known, including somatic (Gaj, 2004; Quiroz-Figueroa et al., 2006), apomictic (Koltunow et al., 1995), microspore-derived (Segui-Simarro and Nuez, 2008; Touraev et al., 1997) and extra-embryonic (suspensor) (Rademacher et al., 2012; Yeung and Meinke, 1993). Even though there are strong morphological analogies in the progression of embryo development among the diverse forms of embryogenesis, little is known about the cellular processes that make a non-embryonic cell competent to undergo embryogenesis.

In previous Chapters, we have used the predictable and uniform extra-embryonic (suspensor) embryogenesis to shed light on the molecular mechanisms that control the initial phase of embryo initiation. The suspensor is derived from the basal cell of the two-cell embryo, formed by the asymmetric division of the zygote (Laux and Jurgens, 1997; Radoeva and Weijers, 2014). Traditionally, the suspensor is considered as a support structure for spatial positioning of the developing embryo in the ovule and as vital transport route for nutrients and growth regulators (Kawashima and Goldberg, 2010; Raghavan, 2006). However, the suspensor is not just suspending the embryo, but it can be considered as a “fountain of youth” in plants due to its developmental potential to allow secondary embryo formation (Lakshmanan and Ambegaokar, 1984). It was recently shown that this potential of the suspensor is cell-autonomously controlled by auxin (Rademacher et al., 2012) and in Chapter 3 we show that the suspensor to embryo transition can be achieved by misexpression of a single gene. To further expand the number of these regulators, we adapted a cell-type-specific two-component GAL4/UAS activation tagging screen (Waki et al., 2013). However, our initial screen using the activation tagging vector described previously (Waki et al., 2013) proved to be less efficient in reaching the required number of transformants for isolating genes conferring suspensor-derived twin formation. A second drawback was the requirement of selective agents like hygromycin, known to introduce developmental defects at low frequency.

Here, we have modified this system to establish an alternative and more efficient high-throughput method for selection and screening of transgenic seeds. The main modification is that we introduced a red fluorescence gene that allows fluorescent seed sorting. In addition, we have optimized the sorting procedure to obtain a large number of transgenic seeds that can be stored for later germination without loss of

viability. With this new system, we have performed a pilot screen and identified a fertile and genetically heritable twin mutant, indicating the reliability of this new generation screen.

Results

Construction of a new tagging vector

To develop an optimized activation tagging vector, we designed a plasmid based on the following principles: (1) The plasmid should be as compact as possible, as smaller plasmid size may enhance replication and stability in *E. coli* and *A. tumefaciens* hosts and smaller T-DNA size may increase transformation efficiency; (2) the T-DNA should contain the 5xUAS part from the original pBIB-UAS plasmid (Waki et al., 2013), and contains primer sequences for TAIL-PCR (see Chapter 3); (3) for facile transgenic identification and antibiotic-free selection, the T-DNA should contain a marker for seed fluorescence. Based on these principles, a minimal vector, pGIIF-UAS-AT was constructed (Figure 1A) based on the plasmid backbone of the pPLV104 (or pPLV4_v2) plasmid. This is a plasmid based on pGreenII, from which a cryptic *Lac* promoter has been removed to increase stability in *E. coli* (De Rybel et al., 2011; Wendrich et al., 2015). pGreenII is a small binary vector, and a second plasmid, pSoup is included in the *Agrobacterium* strain to allow plant transformation (Hellens et al., 2000). The pPLV104 backbone included T-DNA Right Border and Left Border elements. The T-DNA of pGIIF-UAS-AT was composed of two parts. One part was taken directly from pBIB-UAS and contained the 5xUAS element as well as some sequence upstream of this that includes the TAIL PCR primer sequences. The other part was taken from pGoldenGate SE7 (Emami et al., 2013), and contained a 35S promoter-driven *PIP2A-mCherry*. This gene encodes a red fluorescent membrane-localized protein that, when expressed from the 35S promoter, causes red fluorescence in T1 seeds (Emami et al., 2013). The entire pGIIF-UAS-TA vector is only 5.7 kb, and the T-DNA is only 3.1 kb in length. Upon construction and complete sequencing of the entire vector, it was transformed to wild-type plants. Among the seeds from transformed plants were fluorescent individuals that varied in fluorescence intensity (Figure 1B). Importantly, T2 progeny seeds of 10 randomly selected fluorescent T1 seeds segregated red fluorescence (Figure 1C), showing that the construct is stable and provides heritable red seed fluorescence. Since the *PIP2-mCherry* is driven by the 35S promoter, fluorescence could also be observed in seedlings, but intensity varied strongly between lines (not shown).

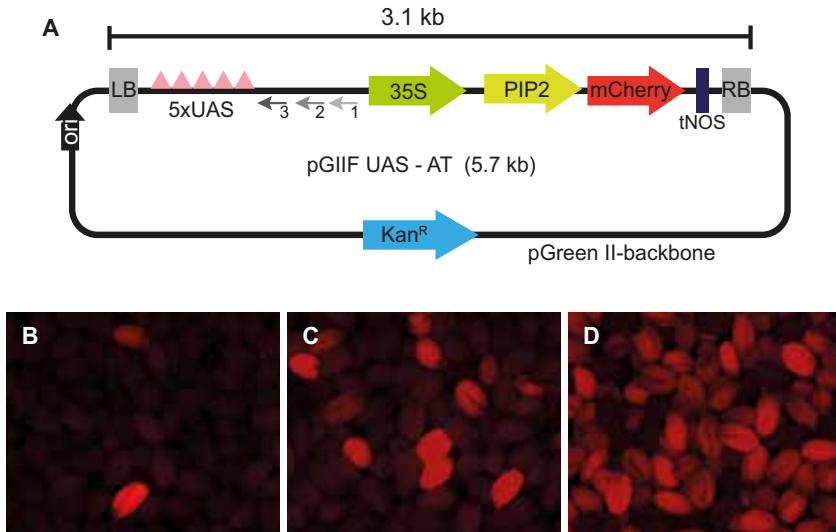


Figure 1: Design of a new activation tagging vector.

(A) Layout of the pGIIF UAS-AT vector. This vector of 5.7 kb total length carries a T-DNA of 3.1 kb with 5xUAS elements pointing towards the T-DNA Left Border (LB). Upstream is a region to which the 3 TAIL-PCR primers can anneal (1,2,3), and a 35S-driven *PIP2-mCherry* gene followed by the NOS terminator near the Right Border (RB). The backbone carries a Kanamycin resistance gene (*Kan^R*; for *E. coli* and *A. tumefaciens* hosts) and an origin of replication (*Ori*). (B-D) Red seed fluorescence in a population of seeds harvested from transformed plants (B; T1), segregating T2 seeds (C) and homozygous T3 seeds (D).

Optimization of seed-sorting conditions

The transformation efficiency of floral dip transformation can range between 0.1 and 2.28%, depending on conditions (Ghedira et al., 2013). To efficiently separate the large excess of non-transgenic seeds from the fluorescent transgenic seeds, we used a BioSorter Large-Particle Flow Cytometry (LPC) platform (Union Biometrica). This platform works on the principal of flow cytometry and allows analyzing, sorting and dispensing objects based on their size, specific light scattering and fluorescent characteristics (Herzenberg et al., 2002). A scheme of the seed-sorting process is presented in Figure 2A. A mixture of *Arabidopsis* seeds is introduced into the flow cell, where they are surrounded by a sheath solution. The sheath solution hydrodynamically guides the seeds through the laser beam, where the size, optical density and fluorescence is measured for each seed. Sorting and dispensing decisions are made based on user-selected parameters. The BioSorter is controlled by software with real-time data acquisition (Union Biometrica), which allows graphical and numerical recording of the optical and fluorescent intensity changes of each seed (Figure 2B, Bi and Bii). Via the software, real-time analysis of the measured

parameters is used to define sort criteria or region of interest (Ri; Figure 2B). If the seed does not meet the sort criteria (falls outside the Ri, Figure 2Bii), it is sent to a waste reservoir by an air dispensing mechanism. The fluorescent seeds that fall in the Ri can be collected into tubes, multi-well plates or petri dishes (Figure 2A, 2Bi). Samples are introduced into flow cell via a sample cup (in our case 50 ml conical tube) with suspended stirrer. According to the nature of the sample, different introduction and collection solutions can be used, for instance water, ethanol or physiological salt solutions. The concentration of seeds determines the sorting time and efficiency.

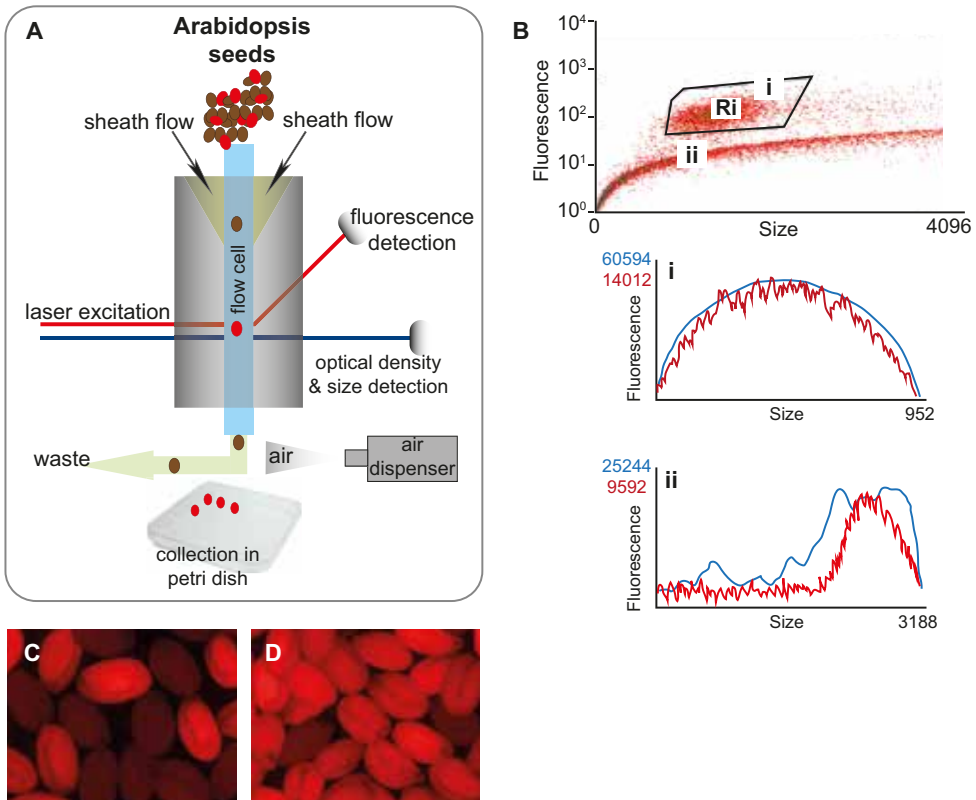


Figure 2: High-throughput seed sorting.

(A) Schematic overview of *Arabidopsis* seed sorting process. (B) Fluorescence over seed size dot-plot. Note that not every dot represents a seed, but could also be auto fluorescent flower remnant. Hence, the plot is overestimation rather than precise quantification of the number of fluorescent seeds. Ri, region of interest; (i) Single-sorted seed signal (ii) Single-unsorted seed signal; red line depicts fluorescence, blue line depicts extinction (optical density or absorbance). (C) Mixture of transgenic and non-transgenic *Arabidopsis* seeds prior to sorting. (D) Sorted *Arabidopsis* transgenic seeds.

As no fluorescent seed sorting protocol had yet been reported for *Arabidopsis*, we first tested a number of conditions and their feasibility to obtain dry and viable seeds with high germination rate. The various conditions and solutions that were tested are summarized in Table 1. *Arabidopsis* seeds were diluted in a conical tube with 40 mL of 60% ethanol to a final concentration of about 300 seeds per ml. A region of interest was defined on fluorescence (FLU) versus seed size dot-plot (Figure 2B). The seeds that were fluorescent and fell in the defined region (Ri) were directly collected on filter paper, in a petri dish. The sorted seeds were immediately dried and could be stored for a later germination without significant loss of viability. Visual inspection confirmed that almost all sorted seeds from Ri were fluorescent (Figure 2D). As a result, with this method efficient and rapid selection of transgenic seeds can be achieved within minutes and further screening of the transgenic seedlings can be directly done by germination on sterile growth medium lacking antibiotics or herbicides.

Table 1: Overview of the tested sorting conditions.

Time	Introduction solution	Time	Collection solution	Germination rate (%)	Comment
5 min	water	10 min		70	seeds uptake water very quickly; cannot be stilized
		30 min	96% ethanol	50	
		60 min		10	
5 min	50 % glycerol	10 min	96% ethanol	90	seeds are sticky and stay on top of the solution; cannot be stilized
10 min		30 min		80	
5 min	70 % ethanol	10 min	96% ethanol	100	seeds sink immediately and cannot enter the flow cell; cannot be stilized
10 min		30 min		90	
5 min	50 % ethnol	10 min	96% ethanol	90	seeds cannot be stilized and stored
10 min		30 min		80	
5 min	60 % ethanol	-	filter paper	100	dried seeds <u>can</u> be sterilized and stored
10 min					

A new generation GAL4/UAS activation tagging screen for identification of embryo inducers

To identify novel genes that possess the ability to trigger embryo formation in suspensor cells, we here used the standardized system of suspensor-derived embryogenesis in combination with the GAL4/UAS activation tagging system, as

described in the previous Chapter (Chapter 3; Figure 4A-C). After transformation, transgenic seeds were separated from the majority of non-transgenic seeds (Figure 2C and D) by fluorescence-based seed sorting (Figure 2A). The transgenic seeds were then plated on growth medium without antibiotics.

To first examine the reliability of the fluorescence selection and to confirm presence of the transgene, we isolated DNA from 12 randomly selected transformants and performed PCR genotyping. As shown in Figure 3A, 10 out of 12 seedlings resulting from the fluorescent seeds tested positive in this PCR, in contrast to the wild type control sample. Thus, most of the seedlings coming from the red fluorescent seeds are true transgenics.

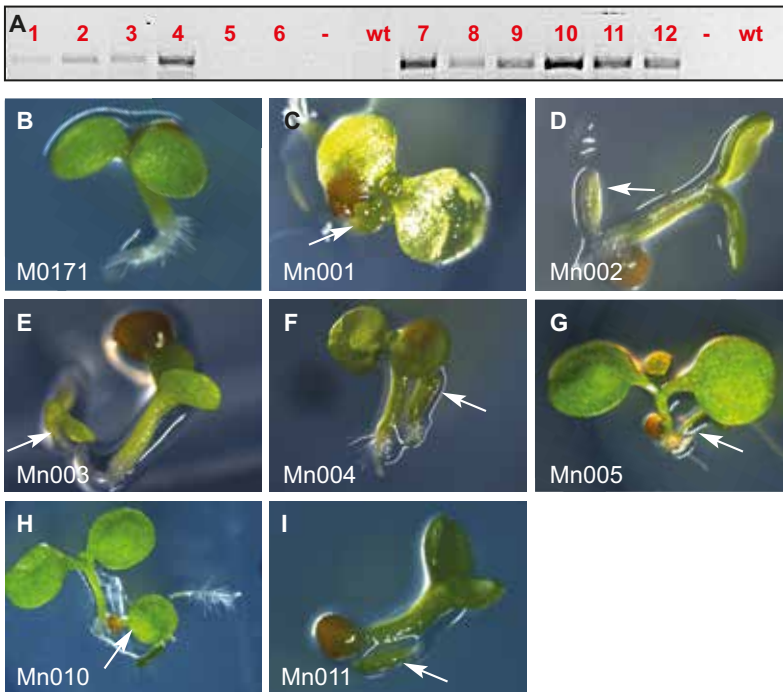


Figure 3: Mutants selected after activation tagging in M0171 background.

(A) PCR confirmation of transgene presence in seedlings germinated from sorted seeds. (B) Untransformed M0171-GAL4 seedling and (C-I) Selected T1 twin mutant seedlings. Arrows indicate the position of the secondary seedling.

With this confirmation in hand, we next performed large-scale transformations, followed by seed sorting, and screened transgenic T1 seedlings for twin seedling phenotypes. After screening a total of 63,000 red fluorescent seed derived seedlings, we identified 7 clearly distinguishable twins (Figure 3B-I). In contrast to the twin-

like phenotypes identified upon more lenient phenotypic criteria in our first screen (Chapter 3), all of the newly selected mutants (Mn class; Figure 3B-I) showed a genuine second seedling sharing a seed with the primary seedling. We consider this first set of 63,000 seedlings a pilot that provides proof of concept. Based on the transformation frequency obtained in this pilot, we estimate that with our currently available transformed seed pools, between 500,000 and 1,000,000 primary transformants can be isolated and screened. Given the size of the genome and constraints for insertion in the vicinity of transcriptional start sites (see discussion), this should suffice to attain an approximately 80% chance at identifying any single gene that is able to confer the twin seedling phenotype upon misexpression in the suspensor cells.

Heritability of the mutant phenotypes

As shown in Chapter 3, a heritable twinning phenotype can be observed upon suspensor-specific RKD1 misexpression. However, none of the twin-like individuals isolated in the first generation activation tagging screen (Chapter 3) passed this phenotype on to their progeny. Therefore, as a first step in isolating new embryogenesis regulators, we transferred twin seedlings to soil for analysis of heritability. From 7 selected twins, only 4 (Mn003, Mn005, Mn010 and Mn011; Figure 3E, G, H and I) grew vigorously enough to survive transfer to soil. At the time of completing this thesis, only Mn003 could be analyzed. Importantly, Mn003 mature plants are comparable to untransformed M0171-GAL4 plants (Figure 4A), which suggests that twinning is not accompanied by general defects or growth retardation. Ovules were dissected from Mn003 siliques and prepared for examination of T2 embryos. Strikingly, 85% (N=56) of the T2 embryos showed proliferation or formation of secondary embryos in suspensor cells (Figure 4B-F). Moreover, twin mature embryos with well-established cotyledons and hypocotyl were found (Figure 4F). Thus, Mn003 is fertile and the initially observed twin phenotype is genetically heritable. Importantly, the high frequency of twin embryos in this line (85%) is consistent with the expected dominant nature of the activation tagging. Further analysis will reveal the identity of the tagged gene. In any event, the Mn003 line provides the first proof of concept and demonstrates the potential for genome-wide screening for novel embryo regulators using this experimental strategy.

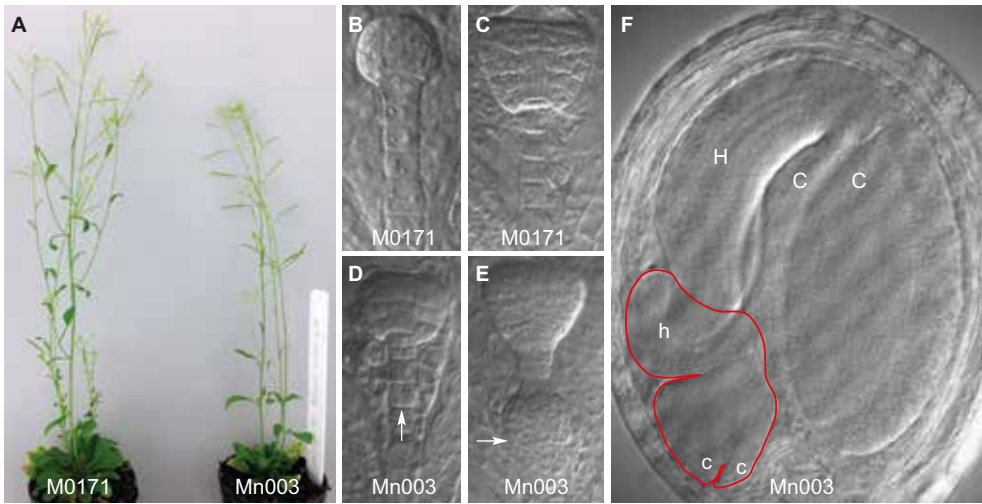


Figure 4: Heritability of the Mn003 twin mutant phenotype.

(A) Phenotypes of flowering untransformed M0171-GAL4 and mutant Mn003 plants. (B-F) Phenotypes of untransformed M0171-GAL4 embryos (B and C) and Mn003 globular stage (D), heart stage (E) and mature (F) embryos. Suspensor cells of Mn003 embryos are dividing (D) to form a second embryo (E), in comparison to the normal wild type suspensor cells. Arrows indicate the dividing suspensor cells. Red line indicates the second mature embryo. C, cotyledon and H, hypocotyl.

Discussion

To mechanistically dissect the initiation phase of embryo formation process, we combined the concept of suspensor-derived embryogenesis with a targeted GAL4/UAS activation tagging approach towards the identification of more “embryo inducers”. The aim of this study was to optimize the activation tagging system used previously (Waki et al., 2013; Chapter 3) by modifying the construct and the initial selection procedure to enable much larger numbers of seedlings to be screened for twins. We constructed a new antibiotic marker-free tagging vector by including a red fluorescence cassette as visual selection marker to allow fluorescent seed sorting as means to identify *Arabidopsis* transgenes. Furthermore, we have established a simple protocol for high-throughput seed sorting using Large-Particle Flow Cytometer (Union Biometrica).

Identification and selection of transgenic seedlings is often a time-consuming process, which first requires germination of the transformed seeds and growth of the resulted seedlings under the stress of selective agents (antibiotics, herbicides), followed by an additional period of time to distinguish between resistant and sensitive seedlings. In addition, the selective agents can have negative effect on seedling growth and viability,

which can mislead the visual selection process in a mutant screen. Using fluorescence-tagged transformed seeds allows rapid identification and selection of transformed *Arabidopsis* seeds without germination (Stuitje et al., 2003). In combination with high-throughput seed sorting this results in significant saving in time, materials and effort as compared to the traditional antibiotic transgenic plant-selection protocol. The obtained transgenic seeds can be simply grown on normal medium and the transgenic seedlings can be safely phenotypically characterized. Here we made use of a 35S promoter-driven *PIP2-mCherry* (Emami et al., 2013). This was shown to cause discernable red seed fluorescence, and we confirmed that this fluorescence is indeed a good selectable marker to identify transgenics. The fluorescence can be followed over generations, and can be used to quickly scan seeds for presence of transgene. It should be mentioned that the use of the 35S promoter may be a disadvantage for some applications, for example when the ubiquitous red fluorescence in post-embryonic tissue interferes with a fluorescence marker-based screen. To make the screening approach more generic, a further improvement could be the use of the Oleosin-RFP maker that is used in the pFAST-R vector (Shimada et al., 2010). This transgene confers much stronger red fluorescence (our own unpublished observations), but because a seed storage protein is tagged, fluorescence quickly fades after germination.

We applied the above-discussed strategy to set up a new generation genome wide screen for identification of embryo inducers. The new T-DNA is nearly 3 times shorter than the one in pBIB-UAS, and we indeed obtained transgenics with relatively high frequency in large-scale transformations. Based on an initial set of sorting and screening experiments, we estimate that screening between 500,000 and 1,000,000 transgenic seedlings, on a weekly basis of 5 to 10,000 per researcher, should be within practical reach using this setup. Given that efficient activation tagging is expected when the insertion is within 1 kb of the transcriptional start site, and should ideally be in the sense orientation, one would like to generate a population with a T-DNA insertion every 1 kb, and in both orientations. In a genome of 120 MB (or 120,000 kb; the Arabidopsis Genome Initiative, 2000), this would equal 240,000 lines, assuming random insertion. Thus, with 500,000 to 1,000,000 seedlings, we can realistically expect every gene to be hit once at an 80% chance. So far, we have screened 63,000 red seed derived seedlings, which is already about 3.5 times more than the total amount of transformants screened in the first screen. We have selected 7 genuine twin seedlings, 4 of which were viable. Up to this point, we can confirm that at least one of the selected mutants is fertile and its twin phenotype is dominant and genetically heritable. Further analysis is required to identify the tagged gene, which will be followed by validation experiments to reproduce the original twin phenotype.

In summary, we have successfully established a new generation activation tagging screen, to identify more genes able to trigger embryo formation in suspensor cells. We have now an efficient and rapid way of identification and selection of mature transgenic seeds in a large background of non-transgenic seeds by fluorescence-based seed sorting.

Materials and Methods

Plant material and growth conditions

The M0171 GAL4/GFP enhancer trap line was generated by Dr. Jim Haseloff in the C24 ecotype (Haseloff, 1999) and was obtained through the Nottingham Arabidopsis Stock Center (NASC).

After sorting, seeds were sterilized in 70% ethanol solution for 10 minutes and were afterwards washed once with 100% ethanol. Dried seeds were subsequently plated on half-strength Murashige and Skoog (MS) medium. After overnight incubation at 4°C, the seedlings were grown for 5–6 d under long-day (16h light, 8h dark) conditions at 22°C.

Construction of an optimized activation tagging vector

The tagging vector pGIIF UAS-AT was constructed from three fragments. A vector backbone fragment, including T-DNA Right Border (RB) and Left Border (LB) elements, was amplified from pPLV104 (also named pPLV04_v2; De Rybel et al., 2011; Wendrich et al., 2015) using primers listed in Table S1. Restriction sites for *EcoRI* and *XhoI* enzymes were included in these primers such that an *EcoRI* site was present at the RB and an *XhoI* site was present at the LB. A fragment encompassing the 5xUAS elements and an upstream sequence including the TAIL-PCR primer binding sites was amplified from the pBIB-UAS plasmid (Waki et al., 2013) using primers listed in Table S1. *Acc65I* and *XhoI* restriction sites were included in these primers such that an *XhoI* site was present near the 5xUAS and an *Acc65I* site was present at the other end. Finally, a fragment containing 35S-PIP2-mCherry-*tNOS* was amplified from the pGoldenGate SE7 plasmid (Emami et al., 2013) using primers listed in Table S1. These primers included *Acc65I* and *EcoRI* restriction sites, such that the *Acc65I* site was at the 35S end, and the *EcoRI* site was at the *tNOS* end. All fragments were amplified using Phusion Flash DNA polymerase (Phusion Flash PCR Master Mix, Thermo Scientific), and digested with the appropriate restriction enzymes. Following gel extraction, a three-fragment ligation was used (using T4 DNA ligase; Thermo Scientific) to obtain a plasmid with all three fragments. The plasmid was sequenced completely.

Seed sorting

Arabidopsis seeds were sorted in a BioSorter Large-Particle Flow Cytometer (LPC; Union Biometrica). mCherry fluorescence was excited by a 488 nm solid-state laser and fluorescence was detected between 586-640 nm. Seeds were introduced in 60% ethanol for sorting and were dispersed directly on filter paper. Sorted seeds were subsequently dried and stored before ethanol-only based sterilization and subsequent further screening.

Screening M0171 activation tagged seedlings

To perform suspensor-specific activation tagging, the pGIIF UAS-AT construct was introduced in *Agrobacterium tumefaciens* strain GV3101 (pMP90) carrying pSoup plasmid (Hellens et al., 2000), and this strain was used to transform the M0171 line by simplified floral dipping (De Rybel et al., 2011). For screening, seeds were plated on half-strength Murashige and Skoog (MS) medium at a density low enough that seedlings grow individually. After approximately 5 days of growth, transgenic seedlings were visually inspected for twins likely caused during embryogenesis using a stereo-microscope.

PCR confirmation of transgene presence

Genomic DNA was isolated from randomly selected transgenic plants and wild type M0171-GAL4 line using CTAB (Cetyltrimethylammonium bromide) extraction buffer (1% CTAB, 100mM Tris-HCL pH8.0, 20mM EDTA pH8.0, 1.5M NaCl) and was afterwards precipitated with isopropanol and washed with 70% ethanol. The isolated DNA was used for PCR amplification using home-made purified recombinant Taq DNA Polymerase. Program: 35 cycles (95°C, 30s; 55°C, 30s; 72°C, 30s). Primers used for amplification are listed in Table S1.

Acknowledgements

The seed sorting was performed in Utrecht University Large-Particle Flow Cytometry Facility (UU-LPC) operated by Dr. Juliane Teapal. This work was funded by a grant from the Netherlands Organization for Scientific Research (NWO; ALW-NSFC grant 846.11.001) to D.W.

References

- De Rybel, B., van den Berg, W., Lokerse, A., Liao, C.Y., van Mourik, H., Moller, B., Peris, C.L. and Weijers, D. (2011). A versatile set of ligation-independent cloning vectors for functional studies in plants. *Plant Physiol* 156, 1292-1299.
- Emami, S., Yee, M.C. and Dinneny, J.R. (2013). A robust family of Golden Gate Agrobacterium vectors for plant synthetic biology. *Front Plant Sci* 4, 339.
- Gaj, M.D. (2004). Factors Influencing Somatic Embryogenesis Induction and Plant Regeneration with Particular Reference to *Arabidopsis thaliana* (L.) Heynh. *Plant Growth Regulation* 43, 27-47.
- Ghedira, R., De Buck, S., Nolf, J. and Depicker, A. (2013). The Efficiency of *Arabidopsis thaliana* Floral Dip Transformation Is Determined Not Only by the Agrobacterium Strain Used but Also by the Physiology and the Ecotype of the Dipped Plant. *Mol Plant Microbe Interact* 26, 823-832.
- Haseloff, J. (1999). GFP variants for multispectral imaging of living cells. *Methods Cell Biol* 58, 139-151.
- Hellens, R.P., Edwards E.A. Leyland, N.R., Bean, S. and Mullineaux, P.M. (2000). pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Mol Biol* 42, 819-832.
- Herzenberg, L.A., Parks, D., Sahaf, B., Perez, O., Roederer, M. and Herzenberg, L.A. (2002). The History and Future of the Fluorescence Activated Cell Sorter and Flow Cytometry: A View from Stanford. *Clin Chem* 48, 1819-1827.
- Kawashima, T. and Goldberg, R.B. (2010). The suspensor: not just suspending the embryo. *Trends Plant Sci* 15, 23-30.
- Koltunow, A.M., Bicknell, R.A. and Chaudhury, A.M. (1995). Apomixis: Molecular Strategies for the Generation of Genetically Identical Seeds without Fertilization. *Plant Physiol* 108, 1345-1352.
- Lakshmanan, K.K. and Ambegaokar, K.B. (1984). Polyembryony. In *Embryology of Angiosperms*, B. Johri, ed. (Springer Berlin Heidelberg), pp. 445-474.
- Laux, T. and Jurgens, G. (1997). Embryogenesis: A New Start in Life. *Plant Cell* 9, 989-1000.
- Quiroz-Figueroa, F.R., Rojas-Herrera, R., Galaz-Avalos, R.M. and Loyola-Vargas, V.M. (2006). Embryo production through somatic embryogenesis can be used to study cell differentiation in plants. *Plant Cell Tiss Organ Cult* 86, 285-301.
- Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Freire Rios, A., Borst, J.W., Lukowitz, W., Jurgens, G. and Weijers, D. (2012). Different auxin response machineries control distinct cell fates in the early plant embryo. *Dev Cell* 22, 211-222.
- Radoeva, T. and Weijers, D. (2014). A roadmap to embryo identity in plants. *Trends Plant Sci* 19, 709-716.
- Raghavan, V. (2006). Life and times of the suspensor - cell signaling between the embryo and suspensor. *Double Fertilization*, 81-100.
- Segui-Simarro, J.M. and Nuez, F. (2008). How microspores transform into haploid embryos: changes associated with embryogenesis induction and microspore-derived embryogenesis. *Physiol Plant* 134, 1-12.
- Shimada, T.L., Shimada, T. and Hara-Nishimura, I. (2010). A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J* 61(3): 519-28.
- Stuitje, A.R., Verbree E.C., van der Linden, K.H., Mietkiewska, E.M., Nap, J.P. and Kneppers, T.J. (2003). Seed-expressed fluorescent proteins as versatile tools for easy (co)transformation and high-throughput functional genomics in *Arabidopsis*. *Plant Biotechnol J* 1, 301-309.
- the *Arabidopsis* Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815.
- Touraev, A., Vicente, O. and Heberle-Bors, E. (1997). Initiation of microspore embryogenesis by stress. *Trends Plant Sci* 2, 297-302.

- Waki, T., Miyashima, S., Nakanishi, M., Ikeda, Y., Hashimoto, T. and Nakajima, K. (2013).** A GAL4-based targeted activation tagging system in *Arabidopsis thaliana*. *Plant J* 73, 357-367.
- Wendrich, J.R., Liao, C.Y., van den Berg, W.A., De Rybel, B. and Weijers, D. (2015).** Ligation-independent cloning for plant research. *Methods Mol Biol* 1284, 421-431.
- Yeung, E.C. and Meinke, D.W. (1993).** Embryogenesis in Angiosperms: Development of the Suspensor. *Plant Cell* 5, 1371-1381.

Supplementary information

Table S1: Primers used in this study.

Fragment	Primer orientation	Sequence
Backbone	Sense	GGATATGAATTCACGGTTATCCACAGAATCAG
	Antisense	GGATATCTCGAGATTCCTAAAACCAAAATCCA
UAS	Sense	GCATATGGTACCGAATGGCAGAAATTCAGCTTGGCC
	Antisense	GGATATCTCGAGAAATTGACGCTTAAGTAGAG
mCherry	Sense	AGATATGGTACCGCCTACTCCAAAAATGTCAAAGAT
	Antisense	GGATATGAATTCCTCCGATCTAGTAACATAGA
Genotyping	Sense	CCACTGACGTAAGGGATGAC
	Antisense	GCACCTTGAAGCGCATGAACTC

Chapter 5

Molecular characterization of Arabidopsis GAL4/UAS enhancer trap lines identifies novel cell-type-specific promoters

Tatyana Radoeva, Colette A. ten Hove,
Shunsuke Saiga and Dolf Weijers

Plant Physiology (2016) 171, pp. 1-13

Laboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703HA
Wageningen, the Netherlands

Cell-type-specific gene expression is essential to distinguish between the numerous cell types of multicellular organism. Therefore, cell-type-specific gene expression is tightly regulated and for most genes RNA transcription is the central point of control. Thus, transcriptional reporters are broadly used markers for cell identity. In *Arabidopsis thaliana*, a recognized standard for cell identities is a collection of GAL4/UAS enhancer trap lines. Yet, while greatly used, very few of them have been molecularly characterized. Here we have selected a set of 21 frequently used GAL4/UAS enhancer trap lines for detailed characterization of expression pattern and genomic insertion position. We studied their embryonic and postembryonic expression domains and grouped them into three groups (early embryo development, late embryo development, and embryonic root apical meristem lines) based on their dominant expression. We show that some of the analyzed lines are expressed in a domain often broader than the one that is reported. Additionally, we present an overview of the location of the T-DNA inserts of all lines, with one exception. Finally, we demonstrate how the obtained information can be used for generating novel cell-type-specific marker lines and for genotyping enhancer trap lines. The knowledge could therefore support the extensive use of these valuable lines.

Introduction

Differences among the numerous cell types of multicellular organism are instructed by unique cell-type-specific gene expression. Understanding how cell identities are genetically controlled is therefore a major challenge in developmental biology. While epigenetic, transcriptional and post-transcriptional mechanisms all contribute to cell-type-specific gene expression, gene activity primarily depends on being actively transcribed. Hence, transcriptional reporters, such as gene promoter-reporter fusions or enhancer trap lines are widely used markers for cell identity in model organisms (De Rybel et al., 2013; Fendrych et al., 2014; Levesque et al., 2006). Marker genes are thus essential tools for determining identity in for example mutant cells (Mylona et al., 2002; Sabatini et al., 2003; Wolters et al., 2010). The need for robust identity markers is especially urgent in plants, where cell identity is very flexible, and can easily be reprogrammed during regeneration (Sena et al., 2009; Sugimoto et al., 2010), somatic embryo induction (de Vries et al., 1988), or even under influence of environmental signals (e.g hydropatterning or stress; Ikeda-Iwai et al., 2003). In addition, cell-type-specific markers are often used to drive ectopic gene expression as part of studying gene function (Waki et al., 2013; Weijers et al., 2006).

Many markers have been generated in the model plant Arabidopsis (*Arabidopsis thaliana*), but in plant developmental biology, a collection of GAL4/UAS enhancer trap lines has become an accepted standard for cell identities. These so-called “Haseloff” lines (Haseloff, 1999), as well as later derivations of the same principle (Ckurshumova et al., 2009; Gardner et al., 2009) thus provide an essential resource in Arabidopsis biology. The GAL4/UAS system is a two-component gene expression system widely used for targeted gene misexpression. It was first developed for use in *Drosophila melanogaster* (Brand and Perrimon, 1993) and later also successfully optimized for Arabidopsis (Haseloff, 1999) and other model organisms (Hartley et al., 2002; Kawakami et al., 2004; Ornitz et al., 1991; Scheer and Campos-Ortega, 1999). The two-component system requires two lines: one that contains GAL4-VP16 - the DNA-binding domain of the yeast GAL4 transcriptional activator fused to the potent Herpes simplex VP16 transcriptional activation domain (Sadowski et al., 1988), which can be driven by a (characterized) cell-type-specific promoter. Another line carries a selected target gene placed under control of a GAL4-dependent promoter (Upstream Activation Sequence [UAS]) that is silent in the absence of GAL4. Genetic crossing between these two lines will specifically activate the target gene in particular tissue or cell types.

The “Haseloff” GAL4/UAS collection (Haseloff, 1999) consists of around 250 plant lines. In these lines, a T-DNA carrying GAL4-VP16 and modified UAS-driven GFP-

gene (mGFP5ER), is randomly inserted into the *Arabidopsis* genome. A minimal CaMV 35S promoter is placed upstream of GAL4-VP16, and expression depends on insertion near an endogenous enhancer element. Thus, GFP expression in these enhancer trap lines reports the activity of genomic enhancer sequences in the proximity of minimal promoter elements (Haseloff, 1999). The set of 250 lines was selected based on GFP expression in the root. From this set, a number of lines, expressed in defined domains of the root, have risen to prominence as very widely used tools in root biology (De Rybel et al., 2013; Levesque et al., 2006; Sabatini et al., 1999; Xu et al., 2013). In addition, as the lines cell-specifically express GAL4-VP16, introduction of a second UAS-driven gene will target expression of that gene in the GFP-marked domain. This strategy has also been widely used for local misexpression (Kang et al., 2013; Sabatini et al., 2003; Weijers et al., 2006).

Yet, while the “Haseloff” lines are extensively used in the *Arabidopsis* community, very few of them have been molecularly characterized (Cary et al., 2002; Gardner et al., 2009; Laplaze et al., 2005). Thus, it is not known where the insertions are located and what gene expression these lines actually report, or if perhaps the insertion disrupts the gene it is inserted in. A specific caveat is that the lines were generated in an ecotype (C24) that is no longer commonly used, and intercrossing with other ecotypes such as Columbia causes strong phenotypic variation, including extreme delay of flowering in the F1 generation. Knowledge of insertion sites would allow migrating the useful expression driver into other ecotypes. Finally, knowing insertion sites in these lines would facilitate introgression in mutant backgrounds as plants could be easily tested for homozygosity or heterozygosity of the marker using PCR.

Here, we selected a set of the 21 most commonly used GAL4/UAS-GFP enhancer trap lines for a detailed characterization of expression and transgene insertion. Given that most of the cell identities in the root are specified during embryogenesis, we extended the description of expression to the embryo. In addition, we explored other seedling organs for GFP expression. We report the genomic insertion sites of these enhancer trap lines and show that this information can be used to infer expression of neighboring genes, as well as to generate novel cell-type-specific markers, as well as to genotype enhancer trap lines. We expect that this resource will aid the extensive use of these important identity markers in the community.

Results and Discussion

Twenty-one GAL4/UAS-GFP enhancer trap lines from the “Haseloff” collection (Haseloff, 1999) were selected based on their widespread use in the Arabidopsis community (De Rybel et al., 2013; Moller et al., 2009; Mylona et al., 2002; Rademacher et al., 2012; Tsugeki et al., 2010; Waki et al., 2013; Webb et al., 2002; Weijers et al., 2006; Wenzel et al., 2012), and on their documented expression pattern (Haseloff, 1999; <http://data.plantsci.cam.ac.uk/Haseloff/>). For each line, we first confirmed and extended the expression analysis, and determined the genomic coordinates of the T-DNA insertion. As a proof of concept, we show how this information can be applied to generate cell-type-specific reporter lines and for genotyping of enhancer trap lines.

Expression patterns of GAL4/UAS enhancer trap lines

We preselected GAL4/UAS lines for root and embryo expression that are among the most used/popular ones from the “Haseloff” collection (De Rybel et al., 2013; Moller et al., 2009; Mylona et al., 2002; Tsugeki et al., 2010; Waki et al., 2013; Webb et al., 2002; Weijers et al., 2006; Wenzel et al., 2012). Although the expression patterns of some of the selected enhancer trap lines were previously described with respect to embryo development and/or root development (e.g. J0571, Q0990, J0121; Laplaze et al., 2005; Wenzel et al., 2012), the expression of all 21 lines was systematically characterized in detail at different stages of development and in different parts of the plant (including embryo, root, leaf, shoot apical meristem [SAM] and cotyledons; Table S1).

Based on confocal microscopy observations, 15 out of 21 GAL4/UAS enhancer trap lines showed GFP expression during embryo development (Table S1) while three (J0121, Q0171 and RM1000) were expressed only later during post-embryonic development (Table S1; Suppl. Figure 1F-J). In the remaining three lines (J2501, J2661 and M0136), no GFP signal could be detected. In previous studies, the expression pattern of J2661 and J2501 was described to be restricted to the pericycle (Levesque et al., 2006; Petersson et al., 2009). A plausible explanation for the lack of GFP expression in our analysis might be a gene silencing event or differences in growth conditions between labs.

In the following, we describe in detail the expression patterns observed for the remaining 18 lines, categorized by their dominant expression domain. All expression patterns are summarized in Table S1.

Early embryo development

A first group, consisting of three lines (M0171, Q0990 and Q2500), showed cell type specific expression during early embryo development (Fig. 1). The GFP expression of M0171 was observed as early as the octant cell stage in all suspensor cells and only later on, when the suspensor is no longer present, the expression was switched on in the cotyledon junction (Fig. 1A-D; Suppl. Fig. 1A). Subsequently, during post-embryonic development, M0171 expression remained cotyledon junction-specific (Fig. 1D). Expression of the Q0990 line commenced at dermatogen stage in the inner basal cells (Fig. 1E), which are the precursors of ground and vascular tissues (Peris et al., 2010; ten Hove et al., 2015). The expression of this line remains vascular tissue-specific later during embryogenesis as well as post-embryonically (Fig. 1F-H). This line has been extensively used as vascular marker (De Rybel et al., 2013; Donner et al., 2009; Levesque et al., 2006) and in misexpression studies (Weijers et al., 2006; Wenzel et al., 2012).

In Q2500 embryos, very weak expression was first detected during late-globular stage in the ground tissue precursors (Fig. 1I). Additional expression foci were found with the progression of embryogenesis (at heart stage) in the mesophyll precursors and protoderm of future cotyledons. Expression extended through the ground tissue of the hypocotyl and root, and also included the pericycle and QC (Fig. 1J). Post-embryonically, the GFP expression of Q2500 is found in the root ground tissue, pericycle and QC (Fig. 1K) and cotyledon junction as predicted by its expression during embryogenesis; but also in the leaf marking the guard mother cells (Fig. 1L; Suppl. Fig. 1C). In contrast to the other two lines in this group, Q2500 has broader expression domain found in multiple tissue types.

Late embryo development

The second group consists of six lines (J0571, J2731, M0148, M0164, M0167 and M0223) whose expression marks specific embryonic tissues during late embryo development. The first line in this group is J0571, which is widely used as ground tissue specific marker (Mylona et al., 2002; Tsugeki et al., 2010; Xu et al., 2013; Zhang et al., 2015) and for targeted gene expression studies (Haseloff, 1999; Waki et al., 2013; Wenzel et al., 2012). Indeed, starting from heart stage, GFP expression in J0571 was observed in the ground tissue, but weak expression could be also seen in the basal protoderm cells (Fig. 2A, B). In addition to the ground tissue expression in the post-embryonic root, GFP could be detected in some, but not all QC cells, and in addition this line showed strong GFP expression in the shoot meristem margins, young leaf primordia and in the leaf epidermis (guard and pavement cells) (Fig. 2C, D and Suppl. Fig. 1D). The expression pattern in line J2731 is very remarkable.

In developing embryos, its expression was first observed in vascular tissue at early heart stage, but at a distance of 1-2 cells from the QC (Fig. 2E, F). This pattern is exactly opposite to that of recently reported “stem cell” zone markers, which are expressed in the cells closest to the lens-shaped cell, but not in the cells above it (Wendrich et al., 2015). Strikingly, J2731 roots reveal a GFP signal in the cells immediately above the QC (Fig. 2G), which suggests that expression shifts towards more juvenile vascular cells as development progresses. Very strong expression was also detected in the guard mother cells and in the pavement cells, similar to line Q2500 (Suppl. Fig. 1C and E).

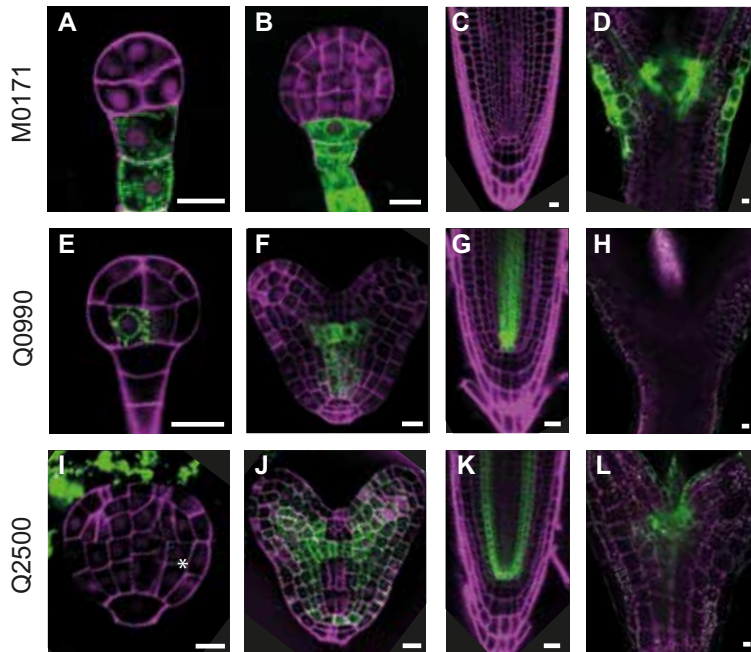


Figure 1: GFP expression in Early embryogenesis lines.

GFP fluorescence in preglobular or globular stage embryos (A, E, I), late-globular or heart stage embryos (B, F, J), root tips (C, G, K) and shoot apex (D, H, L) of M0171 (A-D), Q0990 (E-H) and Q2500 (I-L) lines. Magenta counterstaining in (A, B, E, F, I, J) is Renaissance fluorescence, Propidium Iodide in (C, G, K) and chlorophyll autofluorescence in (D, H, L). All scale bars represent 10 micrometer.

During embryogenesis, expression of M0148 was found in two narrow files of cells surrounding the SAM, while post-embryonically, the marker was expressed in pericycle cells in the differentiation zone of the root (Fig. 2I-K), and in a broad domain surrounding the SAM (Fig. 2L). M0164 marks the vascular strands in the embryonic cotyledon primordia (Fig. 2N). Like M0148, GFP signal of M0164 was detected in the root pericycle (Fig. 2O). Furthermore, strong expression was found

in shoot vascular tissues (Fig. 2P). Finally, lines M0167 and M0223 have similar expression in the intersection of the two cotyledon primordia (Fig. 2Q-X).

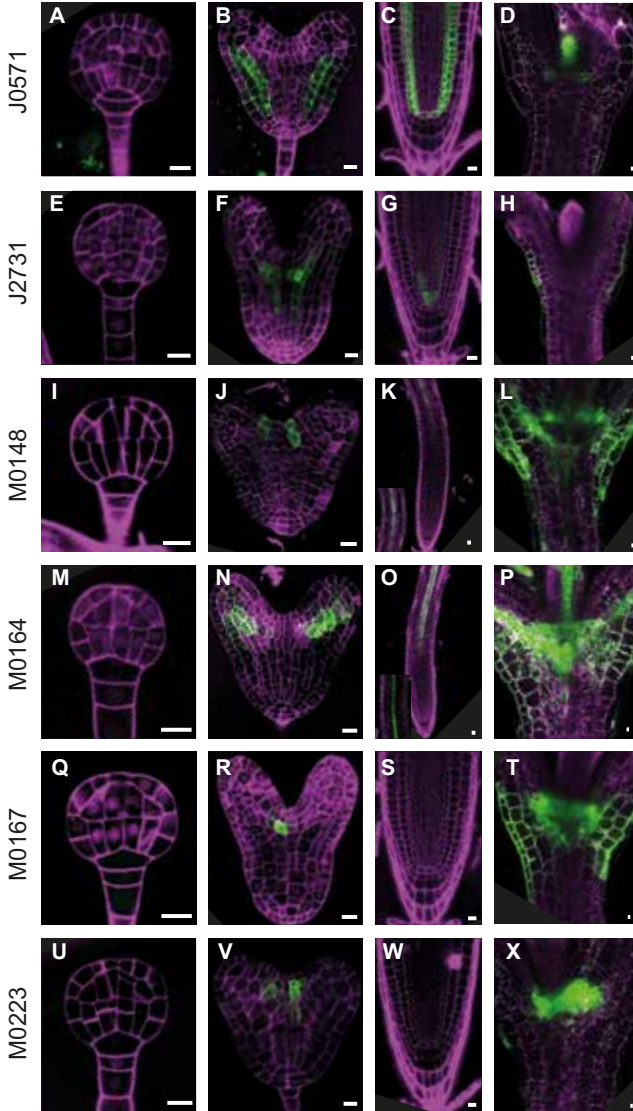


Figure 2: GFP expression in Late embryogenesis lines.

GFP fluorescence in globular stage embryos (A, E, I, M, Q, U), heart stage embryos (B, F, J, N, R, V), root tips (C, G, K, O, S, W) and shoot apex (D, H, L, P, T, X) of J0571 (A-D), J2731 (E-H), M0148 (I-L), M0164 (M-P), M0167 (Q-T) and M0223 (U-X) lines. Magenta counterstaining in (A, B, E, F, I, J, M, N, Q, R, U, V) is Renaissance fluorescence, Propidium Iodide in (C, G, K, O, S, W), and chlorophyll autofluorescence in (D, H, L, P, T, X). All scale bars represent 10 micrometer.

Embryonic RAM lines

Six lines (J1092, J3281, J3411, M0028, Q0680 and Q1630) are expressed mainly or specifically in the embryonic root meristem. All lines in this group show similar expression patterns observed in the future RAM cells and in the case of J1092 and J3411 extended to the lateral root cap (LRC) precursors (Fig. 3C and I). The marked cell types originate from the uppermost suspensor cell, specified as hypophysis at early globular stage of embryonic development. An asymmetric division of the hypophysis then generates a lens-shaped cell from which the quiescent center (QC) will form, and a basal cell, which is the precursor of the root cap (Scheres et al., 1994; ten Hove et al., 2015). The expression of these lines during post-embryonic root development fully recapitulates their embryonic expression domain (Fig. 3) and could not be found in other parts of the plant body, except J3281 whose expression in the post-embryonic root extends to include the young vascular tissue (Fig. 3F). Hence, the lines in this group are excellent candidates for RAM/columnella markers.

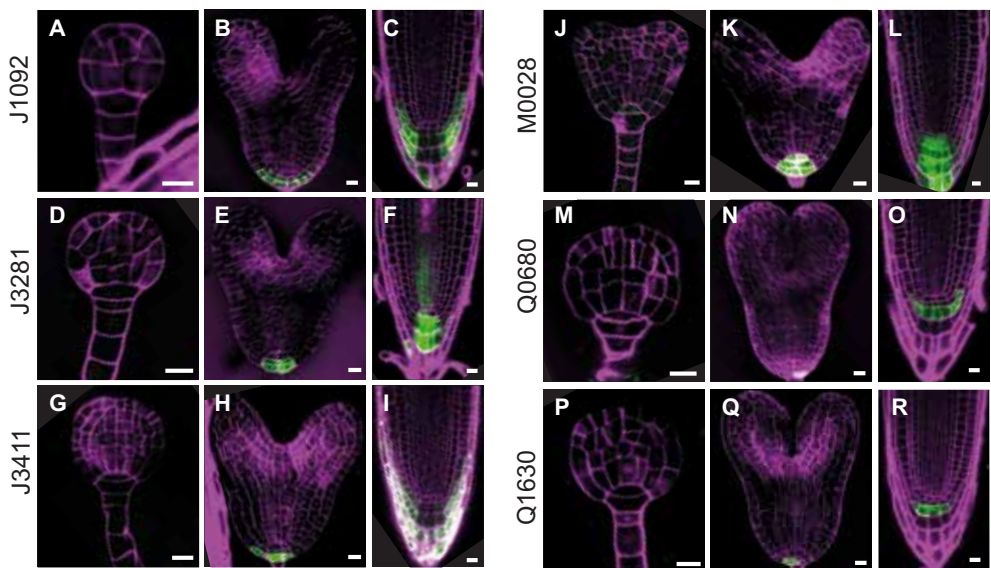


Figure 3: GFP expression in Embryonic root meristem lines.

GFP fluorescence in globular stage embryos (A, D, G, J, M, P), heart stage embryos (B, E, H, K, N, Q) and root tips (C, F, I, L, O, R) of J1092 (A-C), J3281 (D-F), J3411 (G-I), M0028 (J-L), Q0680 (M-O) and Q1630 (P-R) lines. Magenta counterstaining in (A, B, D, E, G, H, J, K, M, N, P, Q) is Renaissance fluorescence, and Propidium Iodide in (C, F, I, L, O, R). All scale bars represent 10 micrometer.

Even though lines J0121, Q0171 and RM1000 did not fall in any of the above-described groups, they possess interesting expression patterns. J0121 is one of the most frequently used xylem pole pericycle marker lines (Laplaze et al., 2005; Parizot et al., 2008; Sugimoto et al., 2010), but we could also detect a very distinct

GFP signal in the seed coat and young leaf primordia (Suppl. Fig. 1H and F). The remaining two lines Q0171 and RM1000 show very specific expression in the post-embryonic root cap and in the leaf margins, respectively (Suppl. Fig. 1I and J).

Identification of insert location

To molecularly characterize the GAL4/UAS enhancer trap lines, as well as to identify novel cell-type-specific promoters during embryo development, we determined the genomic coordinates of the T-DNA insertion. To amplify the genomic DNA flanking the T-DNA inserts, we performed Thermal Asymmetric Interlaced PCR (TAIL-PCR; Liu et al., 1995) with degenerate primers and a set of nested primers in the T-DNA left border. The TAIL-PCR products were subsequently sequenced.

The insert location in lines J0571, J0121, J2731, M0167 and M0223 was previously reported (Cary et al., 2002; Laplaze et al., 2005; Miyashima et al., 2011; Moller et al., 2009) yet, as the description of some was not very explicit, we included these lines as well. In the following we will focus on notable cases only. However, a detailed overview of all insert locations and the flanking genes is presented in Figure 4 and Table S1.

We were able to map the T-DNA insert positions in all lines (Fig. 4; Table S1), with one notable exception. Despite repeated attempts, we did not manage to amplify a T-DNA flank from the M0171 line. Since the line clearly expresses GFP, has been used to trans-activate other genes (Rademacher et al., 2012), and is kanamycin resistant, the T-DNA must be intact. Therefore, the inability to amplify flanking regions may be related to the genomic location (e.g. high GC content or repeats). We also included the three lines (J2501, J2661 and M0136), which did not show any GFP expression, but identified insertion sites for each, suggesting that the lack of GFP expression is not due to absence of an insertion.

Identification of insert location in the early embryo development lines

Intriguingly, in the vascular-specific Q0990 line, the insertion is mapped to a large intergenic region 15.4 kb upstream of the ATG of AT5G43810 and 3 kb upstream of the ATG of AT5G43790 (Fig. 5A). The T-DNA is in the same coding direction as AT5G43810, which encodes ARGONAUTE10 (AGO10, known also as ZWILLE (ZLL); Moussian et al., 1998), a member of the ARGONAUTE (AGO)/ELONGATION INITIATION FACTOR 2C (EIF2C) class of proteins. The other gene, AT5G43790, is in the opposite coding direction to the T-DNA insertion and encodes putative Pentatricopeptide Repeat (PPR) superfamily protein. In addition, a hypothetical

transposable element (AT5G43800) is located between the Q0990 insertion and the ZLL gene. In theory, the insertion should report on the activity of local enhancer elements that normally drive expression of one or both of the adjacent genes. According to microarray data (Belmonte et al., 2013; Brady et al., 2007), expression of both genes is found in the vasculature in the root, but only ZLL is expressed during embryogenesis (Moussian et al., 1998; Tucker et al., 2008). We generated a construct containing the promoter region of ZLL fused to GFP and tested its expression pattern in wild-type plants. As shown in Figure 5, the pattern of expression of *pZLL-GFP* is identical to the one of Q0990 (Fig. 1E-H), which suggests that (a) regulatory element(s) in the *ZLL* promoter (downstream of the insertion) is responsible for the GFP expression pattern in Q0990.

Q2500 enhancer trap line is inserted on chromosome IV in the third exon of UNFERTILIZED EMBRYO SAC12 (UNE12; AT4G02590), 1.1 kb downstream of its ATG and coding in the same direction (Fig. 4; Table S1). UNE12 is a bHLH transcription factor involved in the double fertilization event forming the zygote and the endosperm (Pagnussat et al., 2005). The other gene flanking the T-DNA, in the opposite direction, is AT4G02600. The latter encodes MLO1, a member of a large plant-specific family of seven-transmembrane domain proteins. As shown by microarray experiments (Brady et al., 2007; Moussian et al., 1998), both UNE12 and MLO1 are expressed in the ground tissue, including pericycle, but only UNE12 is expressed during embryo development with a peak at heart stage and might, thus, be responsible for the specific expression pattern in Q2500. Interestingly, the T-DNA insertion in this line (third exon of UNE12) does not induce the developmental defects that were reported for loss of function alleles (Pagnussat et al., 2005).

Identification of insert location in the late embryo development lines

The T-DNA insert in line J0751 is, as reported previously by (Miyashima et al., 2011), in the intergenic region between AT4G39900 and AT4G39910 that encode an unknown protein and a nuclear ubiquitin-specific protease (UBP3), respectively. The insert is located, in the promoter region of AT4G39900, 732 bp upstream of the start codon and oriented in the same direction (Fig. 5B), which makes it more likely to have ground tissue-specific expression. To test if this is true, we made reporter lines using 3 kb and 1.7 kb fragments upstream of the start codon of AT4G39900 and fused both to sensitive nuclear-localized 3xGFP (n3GFP). Both short and long promoter fusions showed slightly broader expression patterns as compared to J0571 (Fig. 5G-L). During embryogenesis, weak expression of *pAT4G39900-n3GFP* was first detected in globular-stage embryos in the protoderm and in the ground tissue

precursor cells (Fig. 5J). Later on, during heart stage, the expression is seen in the same domains with a peak in the ground tissue (Fig. 5H and K). In the post-embryonic root, GFP expression in the epidermis was not observed, but in addition to the ground tissue, a very weak expression could be detected in the vasculature, yet the occasional QC expression in J0571 was not observed in these lines (Fig. 5I and L). Fusion of a shorter genomic DNA fragment, adjacent to the insertion point, may help to limit expression to the ground tissue. In addition, thorough examination of the AT4G39900 promoter using deletions should help to elucidate this point.

M0164 is the only line in our selection that harbors two insertions, one on chromosome I and another on chromosome IV. On chromosome I, the T-DNA insert is oriented in the same direction and 4.5 kb upstream of AT1G09520, and 733 bp upstream of the start codon of AT1G09530, but in reverse orientation. AT1G09520 encodes an unknown protein, which possesses Zinc-finger and PHD-type domains; whereas AT1G09530 encodes PHYTOCHROME INTERACTING FACTOR 3 (PIF3), which is bHLH transcription factor that interacts with photoreceptors phyA and phyB (Ni et al., 1998) (Fig. 4; Table S1). On chromosome IV, M0164 enhancer trap T-DNA was inserted between AT4G00440 and AT4G00430, encoding TON RECRUITING MOTIF 15 (TRM15) and PLASMA MEMBRANE INTRINSIC PROTEIN 1; 4 (PIP1; 4), respectively. The GAL4 T-DNA insertion was about 1.8 kb upstream of and with the same orientation as TRM15. Based on the microarray-predicted expression patterns (Brady et al., 2007) and including the fact that M0164 has two insertions on different chromosomes, it is very challenging to predict which gene(s) underlie the specific expression pattern of M0164. Interestingly, we mapped M0167 enhancer trap T-DNA insertion on chromosome I to exactly the same position as one of the insertions identified in M0164 (Fig. 4; Table S1). Thus, because M0164 has a broader expression pattern than M0167 (Fig. 2M-P and Q-T), it is likely that M0167 is a derivative of M0164.

Identification of insert location in the embryonic RAM lines

In line J3281, the T-DNA insertion was positioned 424 bp downstream of the start codon and in reverse orientation of AT5G62220, and 6.8 kb upstream of AT5G62230 (in the same orientation; Fig. 4; Table S1). The latter encodes ERECTA-LIKE 1 (ERL1; Shpak et al., 2004), an ERECTA (ER) family leucine-rich repeat-receptor-like kinase (LRR-RLK), which together with ERECTA (ER) and ERECTA-LIKE 2 (ERL2) specifies aerial organ size by promoting cell proliferation (Shpak et al., 2004). In addition, ERL1 is involved in specification of stomatal stem cell fate and differentiation of guard cells (Shpak et al., 2005). As reported previously, the expression of ERL1 marks the actively proliferating organs like shoot meristem, leaf primordia and

young developing flowers (Shpak et al., 2004). According to microarray data, expression of ERL1 can be also found in the root vascular tissue (Brady et al., 2007) consistent with the observed pattern of expression in J3281.

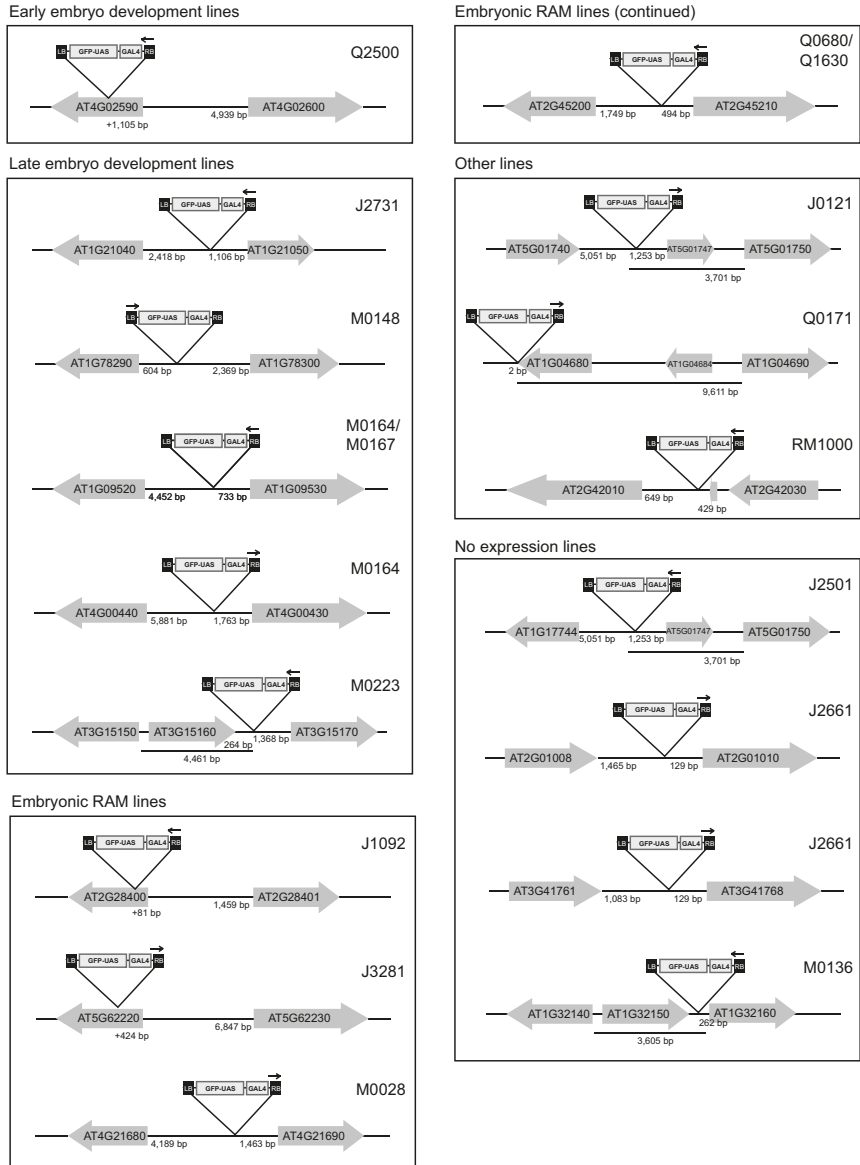


Figure 4: Schematic representation of the GAL4 insertion positions.

Numbers below the T-DNA insertion sites represent the distance (bp) either from the start codon or from the stop codon. Genes are not drawn to scale.

J3411 enhancer trap T-DNA is inserted in an intergenic region between AT2G36370 and AT2G36380 and is coding in the opposite direction to the two flanking genes (Fig. 5C and Table S1). AT2G36370 and AT2G36380 encode ubiquitin-protein ligase and ATP-BINDING CASSETTE G34 (ABCG34, known also as PLEIOTROPIC DRUG RESISTENCE 6), respectively. Both flanking genes have broad expression domains as shown by microarray experiments (Birnbaum et al., 2003) and it is therefore less likely that one of them is responsible for the specific GAL4-GFP expression pattern in J3411. However, the T-DNA was inserted upstream and it was in the same orientation as AT2G36360 (Fig. 5C). The latter encodes Galactose oxidase/kelch repeat superfamily protein with unknown function. To determine if the J3411 expression pattern is due to regulatory DNA sequence in the AT2G36360 promoter, a 3.9kb fragment upstream of the AT2G36360 ATG was fused to n3GFP and was subsequently introduced in wild type Col-0 plants. pAT2G36360-n3GFP was detected in both embryos and roots. In embryos, the GFP signal was first detected at early heart stage, in the future QC and columella cells and expanded to include the ground tissue and the protoderm later on (Fig. 5M, N). In the post-embryonic root, the expression was broader including the vasculature and columella initials, but it was however absent from the distal columella cells (Fig. 5O). Thus, the specific GAL4 expression of J3411 is not replicated by the expression of pAT2G36360-n3GFP. A possible scenario is that the expression pattern observed in line J3411 may be due to regulatory elements located in the genomic DNA instantly upstream of the T-DNA insert. A similar case has been previously reported, where the expression of AT5G65590, a Dof zinc finger transcription factor, which is flanking the insert in the guard cell specific line E1728, did not match the expression of E1728. Alternatively, fusion of a DNA fragment neighboring the T-DNA insert in E1728 and the GAL4 TATA box to *uidA* reporter gene showed strong GUS activity in guard cells (Gardner et al., 2009).

In line M0028, the T-DNA insert was found to be 1.46 kb upstream from the ATG of GIBBERELLIN3-OXIDASE3 (GA3OX3), in the same coding direction (AT4G21690) and 6.2 kb upstream from the ATG of NITRATE TRANSPORTER 1.8 (NRT1.8, known also as NPF7.2), being in the opposite direction of the T-DNA (AT4G21680; Fig. 4). The expression of GA3OX3 has been studied in detail and it was found in heart and torpedo stage embryos, near the junction between the embryo axis and cotyledons (Hu et al., 2008). Although, the T-DNA insert in M0028 is in a very close proximity of GA3OX3 and in the same orientation, GA3OX3 has very different expression pattern. Further analysis on the intergenic region between GA3OX3 and NPF7.2 will provide more information about the elements responsible for the observed columella-specific M0028 expression pattern.

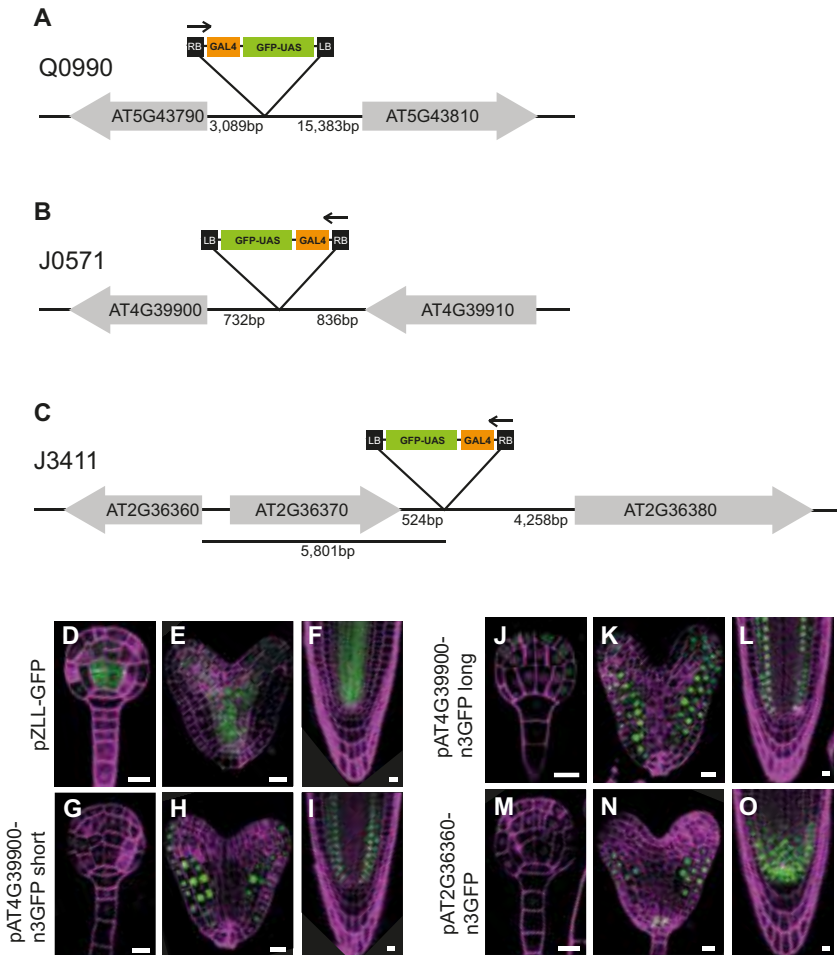


Figure 5: Identification of T-DNA insertion sites and validation of neighboring gene expression. (A-C) Schematics illustrating the insertion site of the enhancer trap T-DNA in Q0990 (A), J0571 (B) and J3411 (C) lines. Neighboring genes are indicated as grey arrows, starting at the gene's ATG, and pointing towards the gene's stop codon. The schematics are not drawn to scale and distances (in bp) are indicated. The orientation of the T-DNA insertions is indicated by an arrow over the T-DNA map (starting from the Right Border). (D-O) Expression of transcriptional fusions of gene promoters to nuclear lamina-localized GFP (D-F) or nuclear 3xGFP (G-O) in globular stage embryos (D, G, J, M), heart stage embryos (E, H, K, N) and root tips (F, I, L, O). (D-F) pZLL-GFP (related to Q0990). (G-L) short version (G-I) or long version (J-L) of the At4G39900 promoter fused to n3GFP (related to J0571). (M-O) pAt2G36360-n3GFP (related to J3411). Magenta counterstaining in (D, E, G, H, J, K, M, N) is Renaissance fluorescence, and Propidium Iodide in (F, I, L, O). All scale bars represent 10 micrometer.

The last two lines in this group, Q0680 and Q1630, report exactly the same expression pattern (Fig. 3M-O and P-R). Moreover, the TAIL-PCR mapping results indicated that insert position of the enhancer traps in both lines is exactly the same, located

between AT2G45200 and AT2G45210 that code for GOLGI SNARE 12 (GOS12) and SMALL AUXIN UPREGULATED 36 (SAUR36), respectively. The GAL4 gene was found 1.75 kb upstream of the start codon of AT2G45200 (GOS12), in the same coding orientation, and only 494 bp upstream of the start codon of AT2G45210 (SAUR36), but in the opposite orientation (Fig. 4; Table S1). GOS12 is predicted to be expressed in the root tip with a peak in the columella initials, while SAUR36 is expected to be more specific to the vasculature (Brady et al., 2007). It is difficult to speculate, which gene is more likely to report the cell-specific GAL4 expression of Q0680/Q1630 based on the predicted expression patterns, since both genes are generally expressed. In this case the T-DNA insert is closer to SAUR36, although it is in the opposite orientation. In line M0223, the T-DNA was also found to be in the opposite orientation, upstream of CUC1 (AT3G15170) (see also Fig. 4; Table S1) and Cary and et al. showed that this is the gene reporting the expression of M0223 (Cary et al., 2002).

Ambiguous insertion sites

In line J2501, the T-DNA insertion is located on chromosome I between a gene encoding 3-Phosphoglycerate dehydrogenase (PGDH; AT1G17745) and a plasma membrane leucine-rich repeat receptor kinase (PERP2; AT1G17750). In this case, the T-DNA is oriented in the opposite direction to the two flanking genes (Fig. 4; Table S1). Both genes are expressed in the root stele (including also the ground tissue), as shown by microarray experiments (Birnbaum et al., 2003) and it is not clear which might be responsible for the GAL4 expression pattern in J2501. The next gene, which is 4.3 kb downstream and in the same orientation, is an unknown gene AT1G17744. In J2661, the T-DNA was located in a duplicated region present in both chromosome II and III. In chromosome II, the insert is found in the very end of the chromosome between the first and the second gene, namely AT2G01008 (unknown gene) and AT2G01010 (18S ribosomal RNA) and the T-DNA was in the same direction as both genes. Since the T-DNA is only 129 bp upstream of the second gene it is possible that this gene is preferentially or exclusively expressed in the pericycle. The situation in chromosome III is similar to the T-DNA inserted between AT3G41768 and AT3G41761, encoding 18S ribosomal RNA and other RNA, respectively. The same as in chromosome II, the T-DNA is 129 bp upstream of the 18S ribosomal RNA. The T-DNA insert in M0136 is positioned, in the opposite direction, between a transcription factor bZIP68 (AT1G32150) and an unknown protein (AT1G32160), but it is in the same direction and 3.6 kb upstream of a F-box family protein (AT1G32140; Fig. 4; Table S1).

Genotyping enhancer trap lines

We have characterized the expression patterns and the insert locations of 21 GAL4 enhancer trap lines. This information can now be used for several purposes, one of which being the development of primer pairs that can identify presence and absence of the insertion, which is useful for PCR genotyping during introgression. As a proof of principle, we designed genotyping primers for four lines. We chose two of the most used GAL4 driver lines, J0121 and J0571; and Q0680 and Q1630, where we can verify the suggestion that these two lines harbor the same enhancer trap insertion.

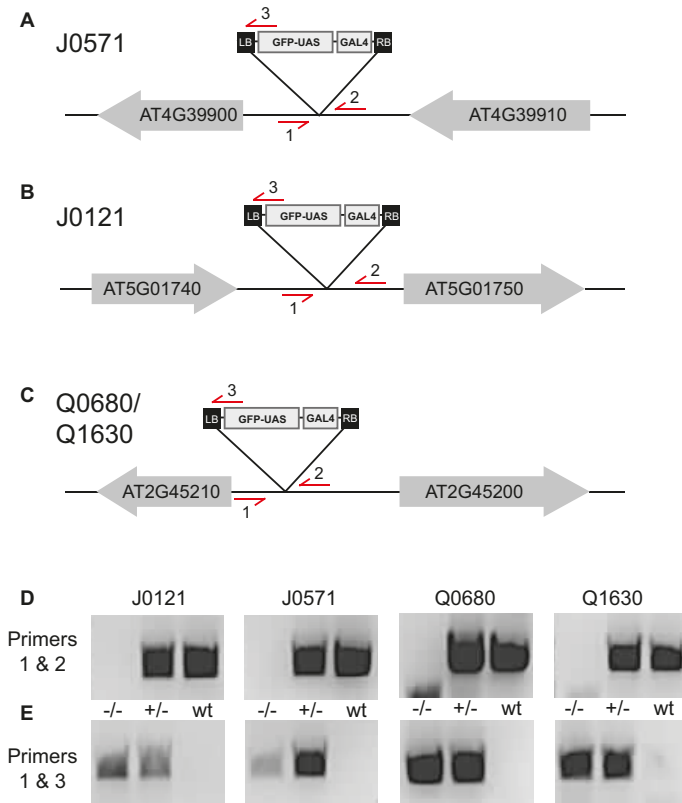


Figure 6: Genotyping GAL4/UAS transgenes.

(A-C) Schematics illustrating the insertion site of the enhancer trap T-DNA in J0571 (A), J0121 (B) and Q0680/Q1630 (C) lines. Neighboring genes are indicated as grey arrows, starting at the gene's ATG, and pointing towards the gene's stop codon. Primers used for genotyping are indicated as red arrows. In each case, the transgene should be amplified using primer 1 and 3, while wild-type is amplified with primer 1 and 2. (D, E) PCR amplification of the wild-type fragment (D; primers 1 & 2) or the insert-specific fragment (E; primers 1 & 3) in genomic DNA from plants homozygous (-/-), heterozygous (+/-) or wild-type (wt) for the J0517, J0121, Q0680 and Q1630 lines. Note that fragments of the appropriate size are amplified according to the expected genotype of each line.

For all four lines, two pairs of primers were generated. One pair binds to the genomic sequence flanking the left and the right border of the T-DNA insert and it should amplify DNA only in absence of insertion (Fig. 6A-C). The other pair of primers uses a primer in the left border of the T-DNA (the same as the third specific primer used for the TAIL-PCRs) and a primer in the genomic sequence in the proximity of the left T-DNA border. Contrary to the first primer pair, this one should amplify DNA only in presence of insertion. We have tested the primers on wild-type plants as well as on homo- and heterozygous plants and the results were as expected (Figure 6D, E).

Concluding remarks

We have selected twenty-one widely used GAL4/*UAS* enhancer trap lines for detailed characterization of expression pattern and genomic insertion site. Firstly, we have systematically documented their expression patterns and grouped them into three groups based on their dominant expression domains. Our microscopic observations demonstrate that the expression patterns of many of the selected GAL4 enhancer trap lines are often expressed in a domain that extends beyond the one that is often reported. This finding should be taken as a cautionary note in two ways. First, the expression may not be taken as evidence for a cell type in its strictest sense, but rather as a regional marker in the local context. Second, when using these GAL4/*UAS* drivers to target local misexpression, one should be aware that the target gene is in fact misexpressed in a broader domain, which might cause more pleiotropic effects that could otherwise be interpreted as non cell-autonomous effects of misexpression. Furthermore, we present an overview of the genomic positions of the T-DNA insertion of all lines, with one exception - M0171. This identified two duplications: line M0167 is likely a derivative of M0164, and lines Q0680 and Q1630 share exactly the same expression pattern as well as the same insert position. In addition, the detailed insertion maps should help to rationalize the cause of local gene expression, and as such will provide a useful tool in studying the genes close to the insertion sites. In addition, the expression patterns can now in principle be migrated to other ecotype backgrounds, or used to generate simpler promoter-reporter fusions based on the genes close to the insertion site.

Materials and Methods

Plant material and growth conditions

The *GAL4-GFP* enhancer trap lines used here (Table S1) are part of a collection of *A. thaliana* lines generated in C24 ecotype (Haseloff, 1999) and are available from the Nottingham Arabidopsis Stock Center (NASC). The new transgenic lines generated in this study are Columbia (Col-0) ecotype.

All *Arabidopsis* seeds were surface-sterilized and dried seeds were subsequently grown on half-strength Murashige and Skoog (MS) plates with or without antibiotics selection at 22°C in standard long-day (16:8 h light:dark) growth conditions. After two weeks of growth, the seedlings were transferred to soil and further grown under the same conditions.

Microscopy

Confocal microscopy was performed as according to (Llavata-Peris et al., 2013) with minor modifications. Briefly, for imaging of embryos, ovules were isolated and mounted in a 4% paraformaldehyde/5% glycerol/1xPBS solution including 1.5% SCRI Renaissance Stain 2200 (R2200; Renaissance Chemicals, UK) for counterstaining of embryos. After applying the coverslip, the embryos were squeezed out of the ovules, and R2200 and GFP fluorescence were visualized by excitation at 405 and 488 nm and detection between 430-470 and 500-535 nm, respectively. For imaging of roots, 5 day-old-seedlings were incubated in 10 µg/ml propidium iodide (PI) solution for 1-2 minutes and GFP and PI were visualized by excitation at 488 nm and detection between 500-535 nm and 630-700 nm, respectively. All confocal imaging was performed on a Leica SP5 II system equipped with Hybrid Detectors.

Mapping of the T-DNA insertion sites

Genomic DNA was isolated from all *GAL4-GFP* enhancer trap line seedlings using CTAB (Cetyltrimethylammonium bromide) extraction buffer (1% CTAB, 100mM Tris-HCL pH8.0, 20mM EDTA pH8.0, 1.5M NaCl, H₂O) and was afterwards precipitated with isopropanol and washed with 70% ethanol. TAIL-PCR was performed as described previously (Liu et al., 1995) with minor modifications using the specific left border primers listed in Table S2. Some PCR products were cloned into pGEM-T vector (Promega) and the PCR products were then sequenced for identification of the T-DNA insertion sites. Database searches were done using the BLAST through the NCBI website (www.ncbi.nlm.nih.gov).

Cloning and plant transformation

For generating transcriptional fusions, up to 4 kb fragments upstream of the ATG were amplified from genomic DNA using Phusion Flash PCR Master Mix (Thermo Scientific) and cloned into vector pPLV104 using ligation-independent cloning (Wendrich et al., 2015a) and primers listed in Table S2. All inserts were completely sequenced and transformed into Col-0 wild-type *Arabidopsis* plants by simplified floral dipping (De Rybel et al., 2011).

Genotyping of the GAL4-GFP enhancer trap lines

Genomic DNA from homozygous and heterozygous GAL4-GFP plants, and C24 wild-type plants was isolated as mentioned above using CTAB extraction buffer. 4 μ l of the isolated DNA was subsequently used for PCR amplification with home-made purified recombinant Taq DNA polymerase. Program: 35 cycles (95°C, 30s; 65°C, 30s; 72°C, 30s). The primers used for genotyping are listed in Table S2.

Acknowledgements

We are very grateful to Jim Haseloff for generating the extremely useful collection of GAL4/UAS lines and sharing these with the *Arabidopsis* community. We acknowledge the Nottingham *Arabidopsis* Stock Center for distributing seeds, Fons de Vogel and Akshita Chordia for assistance in microscopy and insert mapping, and our colleagues Bert De Rybel, Kuan-Ju Lu and Joakim Palovaara for comments on the manuscript. This work was supported by grants from the Netherlands Organization for Scientific Research (NWO; ALW-NSFC Plant Development Collaborative Grant 846.11.001 to D.W. and ALW-VENI grant 863.12.010 to C.t.H.) and the European Research Council (Starting Grant CELLPATTERN; contract number 281573 to D.W.).

Author contributions

T.R. and D.W. conceived the research. T.R. performed expression analysis of all GAL4 enhancer trap lines and mapped and verified all insertion sites. T.R., C.t.H. and S.S. generated promoter-GFP lines for validation. T.R. and D.W. wrote the article with input from all other authors. D.W. supervised the project.

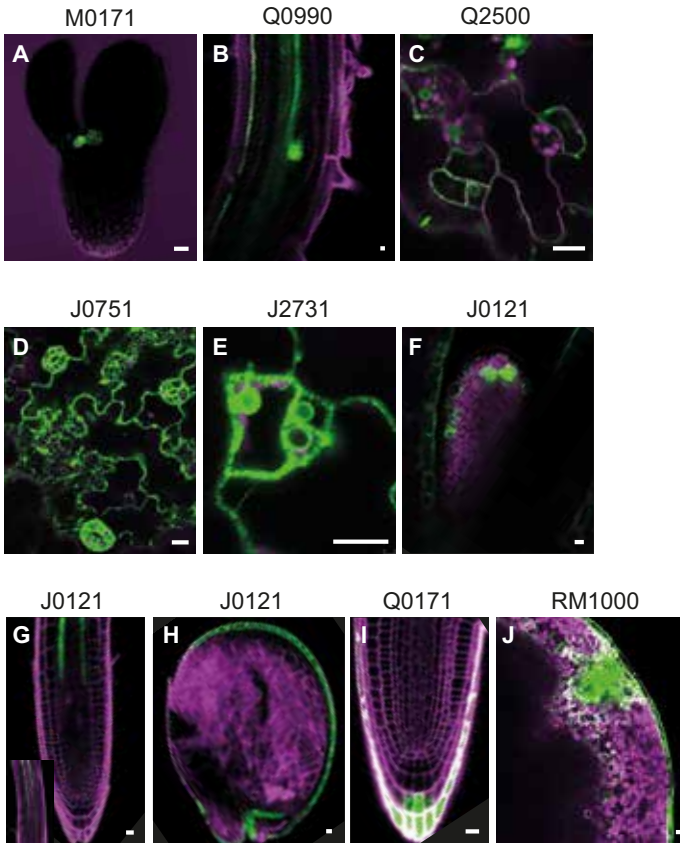
References

- Belmonte, M.F., Kirkbride, R.C., Stone, S.L., Pelletier, J.M., Bui, A.Q., Yeung, E.C., Hashimoto, M., Fei, J., Harada, C.M., Munoz, M.D., Le, B.H., Drews, G.N., Brady, S.M., Goldberg, R.B. and Harada, J.J. (2013). Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *Proc Natl Acad Sci USA* 110, E435-E444.
- Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W. and Benfey, P.N. (2003). A gene expression map of the *Arabidopsis* root. *Science* 302, 1956-1960.
- Brady, S.M., Orlando, D.A., Lee, J.-Y., Wang, J.Y., Koch, J., Dinneny, J.R., Mace, D., Ohler, U. and Benfey, P.N. (2007). A High-Resolution Root Spatiotemporal Map Reveals Dominant Expression Patterns. *Science* 318, 801-806.
- Brand, A.H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401-415.
- Cary, A.J., Che, P. and Howell, S.H. (2002). Developmental events and shoot apical meristem gene expression patterns during shoot development in *Arabidopsis thaliana*. *Plant J* 32, 867-877.
- Ckurshumova, W., Koizumi, K., Chatfield, S.P., Sanchez-Buelna, S.U., Gangaeva, A.E., McKenzie, R. and Berleth, T. (2009). Tissue-Specific GAL4 Expression Patterns as a Resource Enabling Targeted Gene Expression, Cell Type-Specific Transcript Profiling and Gene Function Characterization in the *Arabidopsis* Vascular System. *Plant Cell Physiol* 50, 141-150.
- De Rybel, B., Moller, B., Yoshida, S., Grabowicz, I., Barbier de Reuille, P., 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*. *Dev Cell* 24, 426-437.
- de Vries, S.C., Booi, H., Meyerink, P., Huisman, G., Wilde, H.D., Thomas, T.L. and van Kammen, A. (1988). Acquisition of embryogenic potential in carrot cell-suspension cultures. *Planta* 176, 196-204.
- 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.
- Fendrych, M., Van Hautegeem, T., Van Durme, M., Olvera-Carrillo, Y., Huysmans, M., Karimi, M., Lippens, S., Guérin, Christopher J., Krebs, M., Schumacher, K. and Nowack, M.K. (2014). Programmed Cell Death Controlled by ANAC033/SOMBRERO Determines Root Cap Organ Size in *Arabidopsis*. *Curr Biol* 24, 931-940.
- Gardner, M.J., Baker, A.J., Assie, J.M., Poethig, R.S., Haseloff, J.P. and Webb, A.A. (2009). GAL4 GFP enhancer trap lines for analysis of stomatal guard cell development and gene expression. *J Exp Bot* 60, 213-226.
- Hartley, K.O., Nutt, S.L. and Amaya, E. (2002). Targeted gene expression in transgenic *Xenopus* using the binary Gal4-UAS system. *Proc Natl Acad Sci USA* 99, 1377-1382.
- Haseloff, J. (1999). GFP variants for multispectral imaging of living cells. *Methods Cell Biol* 58, 139-151.
- Hu, J., Mitchum, M.G., Barnaby, N., Ayele, B.T., Ogawa, M., Nam, E., Lai, W.-C., Hanada, A., Alonso, J.M., Ecker, J.R., Swain, S.M., Yamaguchi, S., Kamyia, Y. and Sun, T.P. (2008). Potential Sites of Bioactive Gibberellin Production during Reproductive Growth in *Arabidopsis*. *Plant Cell* 20, 320-336.
- Ikedai-Iwai, M., Umehara, M., Satoh, S. and Kamada, H. (2003). Stress-induced somatic embryogenesis in vegetative tissues of *Arabidopsis thaliana*. *Plant J* 34, 107-114.
- Kang, Y.H., Song, S.K., Schiefelbein, J. and Lee, M.M. (2013). Nuclear Trapping Controls the Position-Dependent Localization of CAPRICE in the Root Epidermis of *Arabidopsis*. *Plant Physiol* 163, 193-204.

- Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N. and Mishina, M. (2004). A Transposon-Mediated Gene Trap Approach Identifies Developmentally Regulated Genes in Zebrafish. *Dev Cell* 7, 133-144.
- Laplaze, L., Parizot, B., Baker, A., Ricaud, L., Martiniere, A., Auguy, F., Franche, C., Nussaume, L., Bogusz, D. and Haseloff, J. (2005). GAL4-GFP enhancer trap lines for genetic manipulation of lateral root development in *Arabidopsis thaliana*. *J Exp Bot* 56, 2433-2442.
- Levesque, M.P., Vernoux, T., Busch, W., Cui, H., Wang, J.Y., Blilou, I., Hassan, H., Nakajima, K., Matsumoto, N., Lohmann, J.U., Scheres, B. and Benfey, P.N. (2006). Whole-Genome Analysis of the SHORT-ROOT Developmental Pathway in *Arabidopsis*. *PLoS Biol* 4, e143.
- Liu, Y.G., Mitsukawa, N., Oosumi, T. and Whittier, R.F. (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* 8, 457-463.
- Miyashima, S., Koi, S., Hashimoto, T. and Nakajima, K. (2011). Non-cell-autonomous microRNA165 acts in a dose-dependent manner to regulate multiple differentiation status in the *Arabidopsis* root. *Development* 138, 2303-2313.
- Moller, I.S., Gilliham, M., Jha, D., Mayo, G.M., Roy, S.J., Coates, J.C., Haseloff, J. and Tester, M. (2009). Shoot Na⁺ exclusion and increased salinity tolerance engineered by cell type-specific alteration of Na⁺ transport in *Arabidopsis*. *Plant Cell* 21, 2163-2178.
- Moussian, B., Schoof, H., Haecker, A., Jürgens, G. and Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during *Arabidopsis* embryogenesis. *EMBO J* 17, 1799-1809.
- Mylona, P., Linstead, P., Martienssen, R. and Dolan, L. (2002). SCHIZORIZA controls an asymmetric cell division and restricts epidermal identity in the *Arabidopsis* root. *Development* 129, 4327-4334.
- Ni, M., Tepperman, J.M. and Quail, P.H. (1998). PIF3, a Phytochrome-Interacting Factor Necessary for Normal Photoinduced Signal Transduction, Is a Novel Basic Helix-Loop-Helix Protein. *Cell* 95, 657-667.
- Ornitz, D.M., Moreadith, R.W. and Leder, P. (1991). Binary system for regulating transgene expression in mice: targeting int-2 gene expression with yeast GAL4/UAS control elements. *Proc Natl Acad Sci USA* 88, 698-702.
- Pagnussat, G.C., Yu, H.J., Ngo, Q.A., Rajani, S., Mayalagu, S., Johnson, C.S., Capron, A., Xie, L.F., Ye, D. and Sundaresan, V. (2005). Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development* 132, 603-614.
- Parizot, B., Laplaze, L., Ricaud, L., Boucheron-Dubuisson, E., Bayle, V., Bonke, M., De Smet, I., Poethig, S.R., Helariutta, Y., Haseloff, J., Chriqui, D., Beeckman, T., Nussaume, L. (2008). Diarch Symmetry of the Vascular Bundle in *Arabidopsis* Root Encompasses the Pericycle and Is Reflected in Distich Lateral Root Initiation. *Plant Physiol* 146, 140-148.
- Peris, C.I., Rademacher, E.H. and Weijers, D. (2010). Green beginnings - pattern formation in the early plant embryo. *Curr Top Dev Biol* 91, 1-27.
- Petersson, S.V., Johansson, A.I., Kowalczyk, M., Makoveychuk, A., Wang, J.Y., Moritz, T., Grebe, M., Benfey, P.N., Sandberg, G. and Ljung, K. (2009). An Auxin Gradient and Maximum in the *Arabidopsis* Root Apex Shown by High-Resolution Cell-Specific Analysis of IAA Distribution and Synthesis. *Plant Cell* 21, 1659-1668.
- Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Freire Rios, A., Borst, J.W., Lukowitz, W., Jurgens, G. and Weijers, D. (2012). Different auxin response machineries control distinct cell fates in the early plant embryo. *Dev Cell* 22, 211-222.
- 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.

- Sabatini, S., Heidstra, R., Wildwater, M. and Scheres, B. (2003). SCARECROW is involved in positioning the stem cell niche in the Arabidopsis root meristem. *Genes Dev* 17, 354-358.
- Sadowski, I., Ma, J., Triezenberg, S. and Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. *Nature* 335, 563-564.
- Scheer, N. and Campos-Ortega, J.A. (1999). Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. *Mech Dev* 80, 153-158.
- 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.
- Sena, G., Wang, X., Liu, H.-Y., Hofhuis, H. and Birnbaum, K.D. (2009). Organ regeneration does not require a functional stem cell niche in plants. *Nature* 457, 1150-1153.
- Shpak, E.D., Berthiaume, C.T., Hill, E.J. and Torii, K.U. (2004). Synergistic interaction of three ERECTA-family receptor-like kinases controls Arabidopsis organ growth and flower development by promoting cell proliferation. *Development* 131, 1491-1501.
- Shpak, E.D., McAbee, J.M., Pillitteri, L.J. and Torii, K.U. (2005). Stomatal Patterning and Differentiation by Synergistic Interactions of Receptor Kinases. *Science* 309, 290-293.
- Sugimoto, K., Jiao, Y. and Meyerowitz, E.M. (2010). Arabidopsis Regeneration from Multiple Tissues Occurs via a Root Development Pathway. *Dev Cell* 18, 463-471.
- ten Hove, C.A., Lu, K.J. and Weijers, D. (2015). Building a plant: cell fate specification in the early Arabidopsis embryo. *Development* 142, 420-430.
- Tsugeki, R., Anicet Ditengou, F., Palme, K. and Okada, K. (2010). NO VEIN facilitates auxin-mediated development in Arabidopsis. *Plant Signal Behav* 5, 1249-1251.
- Tucker, M.R., Hinze, A., Tucker, E.J., Takada, S., Jürgens, G. and Laux, T. (2008). Vascular signalling mediated by ZWILLE potentiates WUSCHEL function during shoot meristem stem cell development in the Arabidopsis embryo. *Development* 135, 2839-2843.
- Waki, T., Miyashima, S., Nakanishi, M., Ikeda, Y., Hashimoto, T. and Nakajima, K. (2013). A GAL4-based targeted activation tagging system in Arabidopsis thaliana. *Plant J* 73, 357-367.
- Webb, M., Jouannic, S., Foreman, J., Linstead, P. and Dolan, L. (2002). Cell specification in the Arabidopsis root epidermis requires the activity of ECTOPIC ROOT HAIR 3 – a katanin-p60 protein. *Development* 129, 123-131.
- Weijers, D., Schlereth, A., Ehrismann, J.S., Schwank, G., Kientz, M. and Jurgens, G. (2006). Auxin triggers transient local signaling for cell specification in Arabidopsis embryogenesis. *Dev Cell* 10, 265-270.
- Wendrich, J.R., Moller, B.K., Uddin, B., Radoeva, T., Lokerse, A.S., De Rybel, B., and Weijers, D. (2015). A set of domain-specific markers in the Arabidopsis embryo. *Plant Reprod* 28, 153-160.
- Wenzel, C.L., Marrison, J., Mattsson, J., Haseloff, J. and Bougourd, S.M. (2012). Ectopic divisions in vascular and ground tissues of Arabidopsis thaliana result in distinct leaf venation defects. *J Exp Bot* 63, 5351-5364.
- Wolters, H., Anders, N., Geldner, N., Gavidia, R. and Jürgens, G. (2010). Coordination of apical and basal embryo development revealed by tissue-specific GNOM functions. *Development* 138, 117-126.
- Xu, P., Yuan, D., Liu, M., Li, C., Liu, Y., Zhang, S., Yao, N. and Yang, C. (2013). AtMMS21, an SMC5/6 Complex Subunit, Is Involved in Stem Cell Niche Maintenance and DNA Damage Responses in Arabidopsis Roots. *Plant Physiol* 161, 1755-1768.
- Zhang, Y., Jiao, Y., Liu, Z. and Zhu, Y.X. (2015). ROW1 maintains quiescent centre identity by confining WOX5 expression to specific cells. *Nat Commun* 6, 6003.

Supplementary information



Supplemental Figure 1: Additional GFP expression patterns.

(A) GFP fluorescence in M0171 torpedo stage embryo, (B) Q0990 lateral root primordia and (C) Q2500 (D) J0571 (E) and J2731 leaf, (F) J0121 SAM, (G) pericycle and (H) and seed coat; (I) Q0171 post-embryonic root cap; (J) RM1000 leaf margin. Magenta counterstaining in (A and H) is Renaissance fluorescence, Propidium Iodide in (B, C, D, E, G, I, J) and chlorophyll autofluorescence in (F). All scale bars represent 10 micrometer.

Table S1: Overview of the GAL4/UAS-GFP expression patterns and GAL4 insertion positions.

Driver line/ ARBC Seed Stock #	Expression pattern			Insertion position (from the ATG)	Flanking Gene(s)	Chromosomal position	Strand	Short gene description	Reference
	Embryo	Root	Other						
Group 1: Early embryo development lines									
Q0990/N9217	pro-vascular cells	vasculature	n. d.	- 15,383 bp	AT5G43810	17,596,556	Forward	Encodes Argonaute10, a member of the EIF2C (elongation initiation factor 2c)/ Argonaute class of proteins	
				- 823 bp	AT5G43800		Forward	transposable element gene	
				- 3,074 bp	AT5G43790		Complement	Pentatricopeptide repeat (PPR) superfamily protein	
	ground tissue, mesophyll	ground tissue, pericycle and QC	cotyledon junction and guard mother cells	+ 1,105 bp	AT4G02590	1,139,201	Complement	unfertilized embryo sac 12 (UNE12), bHLH transcription factor	
Q2500/N9135	precursors, protoderm of cotyledons			- 4,939 bp	AT4G02600		Forward	MLO1; A member of a large family of seven-transmembrane domain proteins specific to plants, homologs of the barley mildew resistance locus o (MLO) protein	
Group 2: Late embryo development lines									
J0571/N9094	ground tissue, basal protoderm cells	ground tissue, some QC cells	guard and pavement cells	- 732 bp	AT4G39900	18,511,021	Complement	unknown protein	Miyashima et al., 2011
J2731	vascular tissue	cells above QC	guard and pavement cells	- 2,418 bp	AT1G21040	7,365,752	Complement	encodes a nuclear ubiquitin-specific protease	Møller et al., 2009
				- 1,106 bp	AT1G21050		Forward	unknown protein	

M0148/N9303	cells surrounding pericycle cells surrounding SAM	AT1G78300	29,459,514	Forward	G-box binding factor GF14 omega encoding a 14-3-3 protein
		- 2,369 bp			encodes a member of SNF1-related protein kinase (SnRK2) family whose activity is activated by ionic (salt) and non-ionic (mannitol) osmotic stress and dehydration.
M0164/N9307	vascular strands in embryonic cotyledon primordia	AT1G09530	3,076,482	Complement	Domains: Zinc finger, PHD-type, conserved site
		- 733 bp		Forward	(PIF3) Transcription factor interacting with photoreceptors phyA and phyB
M0167/N9308	intersection of the two cotyledon primordia	AT4G00440	193,413	Forward	TON1 RECRUITING MOTIF 15, TRM15
		- 1,763 bp		Forward	a member of the plasma membrane intrinsic protein subfamily PIP1
		- 5,881 bp		Complement	Domains: Zinc finger, PHD-type, conserved site
		- 4,452 bp		Complement	(PIF3) Transcription factor interacting with photoreceptors phyA and phyB
		- 733 bp		Forward	(PIF3) Transcription factor interacting with photoreceptors phyA and phyB

M0223/N9336	intersection of the two cotyledon primordia	n. d.	cotyledon junction	- 4,461 bp AT3G15150	5,108,543	Complement	Encodes a SUMO E3 ligase that regulates endocycle onset and meristem maintenance	Cary et al., 2002
				- 264 bp / - 1,368 bp AT3G15160 / AT3G15170		Forward	unknown protein/ CUC1, Encodes a transcription factor involved in shoot apical meristem formation and auxin-mediated lateral root formation	
Group 3: Embryonic RAM lines								
J1092/N9147	future RAM and LRC cells	RAM, LRC	n. d.	+ 81 bp AT2G28400	12,148,920	Complement	unknown protein	
				- 1,459 bp AT2G28401		Forward	unknown protein	
J3281/N9128	future RAM	young vascular tissue; columella cells	n. d.	- 6,847 bp AT5G62230	24,989,586	Forward	Encodes a receptor-like kinase that, together with ER and ERL2 governs the initial decision of protodermal cells to either divide proliferatively to produce pavement cells or divide asymmetrically to generate stomatal complexes	
				+ 424 bp AT5G62220		Complement	Encodes a Golgi apparatus-localized galactosyltransferase involved in galactosyl-substitution of xyloglucan at position 2.	
J3411/N9131	future RAM and LRC cells	RAM, LRC	n. d.	- 5,801 bp AT2G36360	15,253,324	Complement	Galactose oxidase/keich repeat superfamily protein	
				AT2G36370/80		Forward	ubiquitin-protein ligases/pleiotropic drug resistance 6 (PDR6)	

M0028/N9274	future RAM	columella cells	n. d.	- 1,463 bp	AT4G21690	11,525,766	Forward	gibberellin 3-oxidase 3 (GA3OX3)	Encodes a nitrate transporter (NRT1.8). Functions in nitrate removal from the xylem sap
Q0680/N9209	future RAM	columella cells	n. d.	- 6,189 bp	AT4G21680		Complement		
Q1630/N9227	future RAM	columella cells	n. d.	- 1,749 bp	AT2G45200	18,641,389	Complement	Encodes a member of the GOS1 (Golgi SNARE) gene family	
				- 494 bp	AT2G45210		Forward	SAUR36,SAUR-like auxin-responsive protein family	
				- 1,749 bp	AT2G45200	18,641,389	Complement	Encodes a member of the GOS1 (Golgi SNARE) gene family	
				- 494 bp	AT2G45210		Forward	SAUR36,SAUR-like auxin-responsive protein family	
Other lines									
J0121/N9090	seed coat	xylem pole pericycle cells	young leaf primordia	- 3,701 bp	AT5G01750	286,333	Forward	unknown protein	Laplace et al., 2005
				- 1,253 bp	AT5G01747		Forward	Encodes a microRNA that targets several genes containing NAC domains including NAC1	
Q0171/N9207	n. d.	root cap	n. d.	- 5,051 bp	AT5G01740		Forward	Nuclear transport factor 2 (NTF2) family protein	
				- 9,611 bp	AT1G04690	1,304,051	Forward	potassium channel beta subunit 1 (KAB1)	
					AT1G04680		Complement	Pectin lyase-like superfamily protein	

RM1000	n. d.	leaf margins	n. d.	- 649 bp	AT2G42010	17,538,639	Complement	phospholipase D (PLDbeta)
					AT2G42030		Complement	RING/U-box superfamily protein
No expression lines								
J2501/N9121	n. d.	n. d.	n. d.	-4,314 bp	AT1G17744	6,106,109	Complement	unknown protein
					AT1G17745/50		Forward	encodes a 3-Phosphoglycerate dehydrogenase/Encodes PEPR2, a plasma membrane leucine-rich repeat receptor kinase functioning as a receptor for the Pep1 and Pep2 peptides
J2661/N9187	n. d.	n. d.	n. d.	- 129 bp	AT2G01010	3,577	Forward	18S rRNA
					AT2G01008		Forward	unknown protein
				- 129 bp	AT3G41768	14,197,548	Forward	18S rRNA
					AT3G41761		Forward	other RNA
M0136/N9302	n. d.	n. d.	n. d.	- 3,605 bp	AT1G32140	11,568,438	Complement	F-box family protein
					AT1G32150/60		Forward	Encodes a G group bZIP(68) transcription factor family member that can bind cis elements with an ACGT core, such as G-box, Hex, C-box and As-1/ unknown protein

n. d. not detected

Table S2: Primers used in this study.

AT number	Gene name	Primer orientation	Sequence
Primer pairs used for transcriptional fusions			
AT2G36360	-	Sense	TAGTTGGAATGGGTTCCCAACCACACTTATAACT
		Antisense	TTATGGAGTTGGGTTAAGCCGATGATGACGAAATC
AT4G39900 short	unknown	Sense	TAGTTGGAATGGGTTGAAAAGCTGAGAAAGCAGAAAAGAGGAC
		Antisense	TTATGGAGTTGGGTTGAAATGCAGAGAGAATCGTGAACCA
AT4G39900 long	unknown	Sense	TAGTTGGAATGGGTTGAAATTTGGCATTACATGAACITGATATT
		Antisense	TTATGGAGTTGGGTTGAAATGCAGAGAGAATCGTGAACCA
AT5G43810	ZLL	Sense	TAGTTGGAATGGGTTGAAAAGCCGGTGGTTTGCATATC
		Antisense	TTATGGAGTTGGGTTGAAATTTTGTGTTGGATTTTCAAAAACCTC
Primer pairs used for genotyping			
LB 5199 short	Primer 3 in T-DNA	Sense	GCAAAACCAGCGTGGACCGCT
J0571	Primer 1	Antisense	GAGGAAGAAAGCGGGAGGCTTG
	Primer 2	Sense	GCTGGCAATGGCACCGGAGAAAT
J0121	Primer 1 (use with 3)	Antisense	GGAGACAGCATGCTGACCCATGAAC
	Primer 1' (use with 2)	Sense	GCAAGATGGGGAATATAGCA
	Primer 2	Antisense	CTAATTCTAGCCGCTGAACATA
Q0680/Q1630	Primer 1	Antisense	CCGCAACGAACTCTTTGGGG
	Primer 2	Sense	CGTTGGATCCGGGTTTGGCG
Primers used for TAIL-PCR			
LB 5102	Specific 1	Sense	GGAACAACACTCAACCCTATCTCGGG
LB 5147	Specific 2	Sense	GGGATTTGCCGATTTCCGGAACCACC
LB 5199	Specific 3	Sense	GCAAAACCAAGCTGGACCGCTTGCTGC
AD 1	Arbitrary 1		NGTCGASWGANAWGAA
AD 2	Arbitrary 2		WGTGNAGWANCANAGA



Chapter 6

General Discussion

Embryos are one of the most significant innovations in the evolution of land plants and their presence is essential for plant reproduction and dispersal. In seed plants, the embryo is enclosed in a seed and its formation is initiated by a fertilization event, forming the zygote. In *Arabidopsis*, after a precise pattern of several cell divisions, the totipotent zygote is transformed into a mature embryo with the precursors of all the major tissue types of the plant body (ten Hove et al., 2015). In other, larger embryos like in maize the early stages seem to be less precisely organized (Johri et al., 1992), suggesting that the rigid early division pattern seen in *Arabidopsis* embryos represents the minimum set of cells required for early differentiation events.

During *Arabidopsis* embryogenesis, basically every round of cell division is associated with a new cell fate decision whereby different cell identities are installed (Wendrich and Weijers, 2013). Cell identity is usually connected to cell-type specific gene expression and specific transcription factor networks. Although cell identity is normally stable, the fate of a cell can be altered *in vitro* by forced changes in expression of defined transcription factors shown first in *Drosophila* (Schneuwly et al., 1987) and then in mammals (Davis et al., 1987). In mammals, an example of direct reprogramming is the one that confers pluripotency to somatic cells through ectopic expression of four key transcription factors (Oct4, Sox2, Klf4 and Myc (Takahashi and Yamanaka, 2006)), however the efficiency of conversion is shown to be extremely low probably due to chromatin modifications which are highly redundant and may serve as obstacles to the reprogramming event (reviewed in Nashun et al., 2015). In plants, such reprogramming is much more efficient and even a single transcription factor can induce a switch in cell fate (shown in this thesis). Moreover, from non-embryonic (somatic) cells, mature plants can be regenerated, underlining the unique capacity of plant cells to adopt a totipotent state. As discussed in **Chapter 1**, several transcription factors have been reported to induce somatic to embryo cell identity reprogramming in *Arabidopsis* upon their overexpression. The fact that structurally different proteins have the ability to induce embryogenesis raises the important question of whether all somatic cells possess the same genetic information required to change their developmental fate. If so, perhaps the only requirement for embryogenesis is a trigger, which in the above-described case is represented by ectopic gene expression. While attractive, this hypothesis is hard to reconcile with the finding that overexpression of “embryo inducers” does not convert *all* cells to embryonic (Boutilier et al., 2002; Lotan et al., 1998). Furthermore, during differentiation cells acquire different cell identities marked by specific gene expression. Therefore, embryonic and somatic cell lineages may be associated with specific genetic programs. In agreement with the

above-mentioned role of genetic networks in instructing cell identity, a successful reprogramming event may first require a complete deletion of the existing (somatic) transcriptional program and then installation of a new (embryonic) program. In this scenario, “embryo inducers” may act at either step.

Considerable efforts have been made to provide insight into the first steps of the reprogramming process to install embryo identity, but the diverse experimental systems used and lack of systematic analysis make it difficult to assemble the pieces. To mechanistically dissect the embryo initiation process, we exploited the well-known developmental plasticity of the extra-embryonic (suspensor) cells to convert to embryonic cell fate (Yeung and Meinke, 1993). As early as the middle of the 20th century, the potential of suspensor cells to form a second embryo has been appreciated (Haccius, 1955). However, in all but a few cases, the formation of suspensor-derived embryos is a consequence of an incapacitated original embryo. This suggested that the presence of a functional embryo suppresses the embryonic program in suspensor cells (Liu et al., 2015; Schwartz et al., 1994; Weijers et al., 2003; Zhang and Somerville, 1997). A plausible hypothesis is that the embryo sends chemical signals to the suspensor to suppress embryogenesis, and death or arrest of the embryo disrupts signal production or propagation, thus lifting suppression in suspensor cells. One of the few genetic backgrounds in which suspensor embryogenesis occurs in the absence of pro-embryo arrest or death is *twin1*. This mutant generates twin embryos without prior arrest of the pro-embryo (Vernon and Meinke, 1994), and therefore the *TWIN1* gene may be involved in generating or transmitting an embryo-derived signal, or in the actual suppression of embryogenesis in suspensor cells. As the *TWIN1* gene has not yet been identified, this remains an open question. A candidate for an embryo-derived signaling molecule, suppressing suspensor embryogenesis, is the signaling molecule auxin. Components of the auxin response machinery are expressed in suspensor cells (Rademacher et al., 2012; Rademacher et al., 2011) and when these are inhibited, the suspensor proliferates and eventually loses its identity resulting in embryonic cell fate and formation of twin seedlings (Rademacher et al., 2012). Importantly, this response to auxin response inhibition is cell-autonomous, and the inducible switch between cell fates provides an ideal experimental system to study the mechanisms that underlie the initiation phase of embryo formation. Early work in other systems of somatic embryogenesis showed that embryogenic capacity can indeed be induced directly in isolated single cells (Nomura and Komamine, 1985). If this would also hold true for the initial events in suspensor-derived embryogenesis it would obviously simplify interpretation.

In **Chapter 2**, the predictable and uniform response was used to identify genes whose expression changes during the suspensor to embryonic cell fate conversion. We expressed a stabilized mutant Aux/IAA12 (*bdl*) protein – a transcriptional inhibitor of auxin response (Hamann et al., 1999; Rademacher et al., 2012) - exclusively in suspensor cells using the GAL4/*UAS* two-component expression system to locally inhibit auxin response (Rademacher et al., 2012). For that purpose, we used an enhancer trap line (M0171) that expressed GAL4 specifically in suspensor cells. We next performed a genome-wide transcriptomic analysis on embryos harvested when the first aberrant cell divisions started to occur. Despite this narrow window in which the embryos were collected, an unexpectedly large number of genes were differentially expressed in M0171 >> *bdl* embryos, suggesting that reprogramming is a complex transcriptional response. Because the pro-embryo and suspensor are physically connected, even in these isolated embryos, transcriptional changes will likely include both primary effects in suspensor cells and secondary effects in pro-embryo cells. Therefore, a large-scale expression analysis was performed using promoter-GFP reporters for nearly 70 differentially expressed genes. This helped to identify a smaller set of genes whose changes in expression are a likely result of local auxin response inhibition in suspensor cells. Here, we focused on a set of 4 *bHLH* genes. A rationale for choosing these *bHLH* genes for in-depth analysis was that other members of the *bHLH* transcription factor family were previously reported to be involved in auxin-dependent development (Chandler et al., 2009; De Rybel et al., 2013; Schlereth et al., 2010). We showed that all these 4 *bHLH*s are indeed regulated by auxin, in an ARF-dependent manner, although only one, namely *bHLH49*, appeared to be an immediate auxin target. Indeed, when analyzing genome-wide transcriptional changes in *bHLH49* misexpression and *bhlh49* mutant roots, we found that the other *bHLH* genes are among the misregulated genes. This suggests that their auxin regulation is mediated by *bHLH49*. Phenotypic analysis showed that *bHLH49* is also biologically relevant for auxin-dependent suspensor to embryo transformation. Overexpression of *bHLH49* resulted in abnormal divisions and even formation of embryo-like structures in the suspensor, resembling the suspensor-specific *bdl* misexpression. Thus, the auxin-repressed *bHLH49* gene is an important mediator of auxin-dependent suppression of embryo identity in suspensor cells. An important, yet unanswered question is what cellular process is triggered by *bHLH49* to promote embryo development in the suspensor? Our analysis of the *bHLH49*-dependent transcriptome does not pinpoint a key cellular process that can illuminate its ability to target proliferation of suspensor cells. Nonetheless, the overlap of differentially expressed genes in both M0171 and *bHLH49* microarrays hints towards the need of defined set of genetic regulators required to induce the

switch from suspensor to embryonic cell fate. An in-depth analysis of the genes targeted by *bHLH49*, ideally in embryos, will likely help to define a cellular target process. *bHLH49* is normally more strongly expressed in embryo cells, and is repressed by auxin response in the suspensor. An important question is whether the same genes are targeted by *bHLH49* in its normal expression domain and in the ectopic suspensor domain. Another interesting observation is that although *bHLH49* overexpression can trigger formation in embryo-like structures in suspensor cells and despite the severe post-embryonic defects, twin seedlings were never observed. This observation can be associated with dedifferentiation of the suspensor cells (erasure of their initial transcriptional program) into cells that can only divide, but cannot be transformed completely into functional embryo structures (installation of embryonic program).

In the past decades, a range of genes has been identified that are able to trigger embryogenesis. As reviewed in **Chapter 1**, even though some potential convergence points between embryo regulators were suggested, a more rigorous and systematic analysis in a uniform system is needed to establish the relevance of such convergence points. In **Chapter 3**, we used the predictable suspensor-derived embryogenesis to systematically assess the effect of known “embryo inducers” on embryo initiation process. We demonstrated that, apart from *bdl* (Rademacher et al., 2012) and *bHLH49* (**Chapter 2**), from the 15 previously reported “embryo inducers” only RKD1 was capable of generating suspensor-derived twin embryos, reflecting the specificity of suspensor to embryo transition. It is remarkable, how a single transcription factor is able to override the inhibition of the embryonic program in the suspensor cells and to generate viable mature twin plants. In *Arabidopsis*, there are 14 *RWP-RK* genes, divided in two subfamilies NIN-like and RKD (Schäuser et al., 2005). The RKD subfamily is composed of 5 members and except RKD1 two other members (RKD2 and RKD4) were previously also reported to post-embryonically promote embryogenesis (Koszegi et al., 2011; Waki et al., 2011), but none of them could confer embryonic fate to suspensor cells. Yet, the homology between members of the family implies presence of a specific RKD1 domain that might be required for induction of embryo formation in suspensor cells. Further analysis of RKD1 function, possibly including domain swaps between RKD1 and its close homolog RKD2, should help to dissect the unique capacities of RKD1. There are two possible scenarios that could explain the differences in embryo-induction potential between the homologs. Firstly, since the RKD proteins act as transcription factors, the range of targets recognized by each protein might be different. Another interpretation could be that the difference lies, not in their intrinsic potential, but rather in the level to which they should be misexpressed in the suspensor. It is conceivable that RKD1 is

expressed at higher levels, and that RKD2 and RKD4 would also induce suspensor-derived embryogenesis when expressed at higher levels. Another interesting finding is that none of the other embryo inducers tested induced embryogenesis in suspensor cells. These included BBM, LEC1, AGL15 and SERK1. Each of these had previously been shown to induce (aspects of) embryogenesis when overexpressed. Could it be that most of these factors act later in the above-discussed multi-step process that first requires erasure of previous programs and/or installment of competence? Does auxin response, through regulating bHLH49 and perhaps RKD1 establish competence? Clearly, this first systematic comparison of a range of regulators in the same, simple developmental system, already reveals marked differences between the activity of these regulators, possibly depending on the cellular context in which they are expressed.

Based on the results in **Chapters 2** and **3**, the suspensor seems to be a system in which specific genetic regulators can be found that induce embryogenesis. Thus, in **Chapter 3** and **Chapter 4**, our efforts were focused on adaptation and optimization of a genome-wide activation tagging screen to identify more genes that can, by themselves, reprogram the cell fate and install embryonic cell identity in the suspensor. The screen relies on the use of suspensor-specific GAL4 driver line (M0171) where a T-DNA that carries a GAL4-dependent *UAS* promoter pointing outward was randomly inserted. Insertion of this T-DNA next to a gene may cause activation of that gene in suspensor cells. If the gene has the capacity to induce embryogenesis in suspensor cells, this will lead to twin seedlings. The addition of a simple protocol for high-throughput seed sorting and antibiotic-free selection of mutant phenotypes resulted in a very efficient selecting procedure. Remarkably, in our pilot “proof of concept” screen, we found 7 genuine twin seedlings and for one, we confirmed that the true twinning was both genetically heritable and dominant, and was caused by excessive divisions in suspensor cells. This confirms the abilities of the screen to find genes that can do all that is required to ultimately change the suspensor cell fate “single-handed”. With this efficient procedure in hand, we should now be able to saturate the genome with insertions and systematically identify all genes that can trigger embryogenesis in suspensor cells.

A major aim of the work described in this thesis was to mechanistically dissect the embryo initiation process and to try to shed more light on the trigger upon which a non-embryonic cell is induced to change its fate and to enter an embryonic developmental program. By consistently using a uniform experimental system, suspensor-derived embryogenesis, we have identified three independent factors, *bdl* (Rademacher et al., 2012), bHLH49 (**Chapter 2**) and RKD1 (**Chapter 3**) that can

induce embryo formation in the context of the suspensor. However, it is unclear if the pathways and the set of genes involved in the response to each factor are conserved. Comparison of transcriptomes in embryos in which either of the regulators is mutated or misexpressed should uncover if there is a significant overlap of misregulated genes that are part of a core embryo induction program. In addition, as touched upon in this discussion, the *Arabidopsis twin1* recessive mutant has a phenotype similar to the one found upon misexpression of *bdl*, *RKD1* or *bHLH49*, all three of which are dominant traits. In all cases, twin embryos are formed from the suspensor without prior defects in the pro-embryo (**Chapter 2,3**; (Vernon and Meinke, 1994). The *TWIN1* gene has not yet been characterized, but at present we are in the process of mapping the mutation using Next-Generation Sequencing of genomes in bulk segregants of a cross between *twin1* and its wild-type. We therefore hope to identify the molecular basis of the *twin1* mutant phenotype soon, and determine if and how this is connected to *bdl*, *bHLH49* and *RKD1* activity. However, even though the *TWIN1* gene is as yet unknown, the *twin1* mutant is an excellent fourth genetic background in which suspensor cells change their fate to an embryonic one. A future research direction would be to follow the origin and the subsequent development of the secondary twin embryos in the mutant with molecular markers. This will, for instance, reveal when exactly the cell fate conversion is completed and if there is a consistency on which suspensor cell will host the formation of the second embryo. Our initial observations on *RKD1* twin embryos (**Chapter 3**) showed that there is no precise timing and consistency on when the first aberrant division in the suspensor will arise and in which cell that will happen. It appears that growth and development of the two embryos is always asynchronous, consistent with the delay in development of the twin seedlings. Further, the successfully established new generation activation tagging screen is the first step towards generating a complete map of genes that can induce embryogenesis in the suspensor.

Changes in gene expression, in both animals and plants, are coordinated by epigenetic regulation, which involves chromatin modification, including DNA methylation and histone modification. The evolutionary conserved Polycomb Repressive Complex (PRCs) proteins are involved in histone modification as they repress their target genes by catalyzing histone H3 lysine 27 trimethylation (H3K27me3). Recently, transcriptomic analysis on differentiated and undifferentiated tissues in plants revealed hundreds of H3K27me3 target genes (protein coding, transposable element and miRNA genes), suggesting that plant PRC proteins have an essential role in regulating tissue-specific expression patterns of gene families and interestingly proposed a link between epigenetic mechanisms and auxin signaling (Lafos et al., 2011). Moreover, DNA methylation was proposed to influence, either directly

or indirectly, the transcription of genes implicated in cell identity regulation during early embryogenesis (Xiao et al., 2006). In particular, DNA methylation is likely required for normal suspensor development (Xiao et al., 2006). Epigenetic regulation appears to be different between the embryonic and extra-embryonic primary lineages and this fundamental difference is conserved in plants and animals (Yamanaka and Blau, 2010). In *Arabidopsis*, PRC1 and 2 along with other chromatin remodeling factors were shown to suppress embryonic state in differentiated tissues by directly interacting with transcription factors involved in embryogenesis (Bouyer et al., 2011; Bratzel et al., 2010; Ogas et al., 1999; Yang et al., 2013). Hence, it is clear that there is a link between genetic and epigenetic mechanisms, but how these work together to establish a balance between differentiated and undifferentiated cell state is waiting for experimental support.

In conclusion, the results in this thesis indicate that the suspensor to embryo cell-fate transition requires a defined set of genetic regulators and it is a step forward to understanding the crucial first reprogramming step of embryo induction process. The latter is important not only in fundamental point of view, but also when considering engineering embryogenesis for crop propagation. In addition, this work provides essential tools and a basis for further research to unravel the mystery of totipotency in plants.

References

- Boutillier, K., Offringa, R., Sharma, V.K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.-M., van Lammeren, A.A.M., Miki, B.L.A., Custers, J.B. and van Lookeren Campagne, M.M. (2002). Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. *Plant Cell* 14, 1737-1749.
- Bouyer, D., Roudier, F., Heese, M., Andersen, E.D., Gey, D., Nowack, M.K., Goodrich, J., Renou, J.P., Grini, P.E., Colot, V. and Schnittger, A. (2011). Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. *PLoS Genet* 7, e1002014.
- Bratzel, F., López-Torrejón, G., Koch, M., Del Pozo, J.C. and Calonje, M. (2010). Keeping Cell Identity in Arabidopsis Requires PRC1 RING-Finger Homologs that Catalyze H2A Monoubiquitination. *Curr Biol* 20, 1853-1859.
- Chandler, J.W., Cole, M., Flier, A. and Werr, W. (2009). BIM1, a bHLH protein involved in brassinosteroid signalling, controls Arabidopsis embryonic patterning via interaction with DORNROSCHEN and DORNROSCHEN-LIKE. *Plant Mol Biol* 69, 57-68.
- Davis, R.L., Weintraub, H. and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987-1000.
- De Rybel, B., Moller, B., Yoshida, S., Grabowicz, I., Barbier de Reuille, P., 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. *Dev Cell* 24, 426-437.
- Haccius, B. (1955). Experimentally Induced Twinning in Plants. *Nature* 176.
- Hamann, T., Mayer, U. and Jürgens, J. (1999). The auxin-insensitive bodenlos mutation affects primary root formation and apical-basal patterning in the Arabidopsis embryo. *Development* 126, 1387-1395.
- Johri, B.M., Ambegaokar, K.B., and Srivastava, P.S. (1992). Comparative embryology of angiosperms (Berlin; New York: Springer-Verlag).
- Közsegi, D., Johnston, A.J., Rutten, T., Czihal, A., Altschmied, L., Kumlehn, J., Wust, S.E., Kirioukhova, O., Gheyselinck, J., Grossniklaus, U. and Bäumllein, H. (2011). Members of the RKD transcription factor family induce an egg cell-like gene expression program. *Plant J* 67, 280-291.
- Lafos, M., Kroll, P., Hohenstatt, M.L., Thorpe, F.L., Clarenz, O. and Schubert, D. (2011). Dynamic Regulation of H3K27 Trimethylation during Arabidopsis Differentiation. *PLoS Genet* 7, e1002040.
- Liu, Y., Li, X., Zhao, J., Tang, X., Tian, S., Chen, J., Shi, C., Wang, W., Zhang, L., Feng, X., et al. (2015). Direct evidence that suspensor cells have embryogenic potential that is suppressed by the embryo proper during normal embryogenesis. *Proc Natl Acad Sci USA* 112, 12432-12437.
- Lotan, T., Ohto, M., Yee, K.M., West, M.A., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93, 1195-1205.
- Nashun, B., Hill, P.W.S. and Hajkova, P. (2015). Reprogramming of cell fate: epigenetic memory and the erasure of memories past. *EMBO J* 34, 1296-1308.
- Nomura, K. and Komamine, A. (1985). Identification and Isolation of Single Cells that Produce Somatic Embryos at a High Frequency in a Carrot Suspension Culture. *Plant Physiol* 79, 988-991.
- Ogas, J., Kaufmann, S., Henderson, J. and Somerville, C. (1999). PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in Arabidopsis. *Proc Natl Acad Sci USA* 96, 13839-13844.
- Rademacher, E.H., Lokerse, A.S., Schlereth, A., Llavata-Peris, C.I., Bayer, M., Kientz, M., Freire Rios, A., Borst, J.W., Lukowitz, W., Jurgens, G. and Weijers, D. (2012). Different auxin response machineries control distinct cell fates in the early plant embryo. *Dev Cell* 22, 211-222.
- Rademacher, E.H., Moller, B., Lokerse, A.S., Llavata-Peris, C.I., van den Berg, W. and Weijers, D.

- (2011). A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. *Plant J* 68, 597-606.
- Schauser, L., Wieloch, W. and Stougaard, J. (2005). Evolution of NIN-Like Proteins in Arabidopsis, Rice, and *Lotus japonicus*. *J Mol Evol* 60, 229-237.
- Schlereth, A., Möller, B., Liu, W., Kientz, M., Flipse, J., Rademacher, E.H., Schmid, M., Jurgens, G. and Weijers, D. (2010). MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. *Nature* 464, 913-916.
- Schneuwly, S., Klemenz, R. and Gehring, W.J. (1987). Redesigning the body plan of *Drosophila* by ectopic expression of the homoeotic gene *Antennapedia*. *Nature* 325, 816-818.
- Schwartz, B.W., Yeung, E.C., and Meinke, D.W. (1994). Disruption of morphogenesis and transformation of the suspensor in abnormal suspensor mutants of Arabidopsis. *Development* 120, 3235.
- Takahashi, K. and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663-676.
- ten Hove, C.A., Lu, K.J. and Weijers, D. (2015). Building a plant: cell fate specification in the early Arabidopsis embryo. *Development* 142, 420-430.
- Vernon, D. and Meinke, D. (1994). Embryogenic transformation of the suspensor in twin, a polyembryonic mutant of Arabidopsis. *Dev Biol* 165, 566-573.
- Waki, T., Hiki, T., Watanabe, R., Hashimoto, T. and Nakajima, K. (2011). The Arabidopsis RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. *Curr Biol* 21, 1277-1281.
- Weijers, D., Van Hamburg, J.P., Van Rijn, E., Hooykaas, P.J. and Offringa, R. (2003). Diphtheria toxin-mediated cell ablation reveals interregional communication during Arabidopsis seed development. *Plant Physiol* 133, 1882-1892.
- Wendrich, J.R. and Weijers, D. (2013). The Arabidopsis embryo as a miniature morphogenesis model. *New Phytol* 199, 14-25.
- Xiao, W., Custard, K.D., Brown, R.C., Lemmon, B.E., Harada, J.J., Goldberg, R.B. and Fischer, R.L. (2006). DNA Methylation Is Critical for Arabidopsis Embryogenesis and Seed Viability. *Plant Cell* 18, 805-814.
- Yamanaka, S. and Blau, H.M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. *Nature* 465, 704-712.
- Yang, C., Bratzel, F., Hohmann, N., Koch, M., Turck, F. and Calonje, M. (2013). VAL- and AtBMI1-Mediated H2Aub Initiate the Switch from Embryonic to Postgerminative Growth in Arabidopsis. *Curr Biol* 23, 1324-1329.
- Yeung, E.C. and Meinke, D.W. (1993). Embryogenesis in Angiosperms: Development of the Suspensor. *Plant Cell* 5, 1371-1381.
- Zhang, J.Z. and Somerville, C.R. (1997). Suspensor-derived polyembryony caused by altered expression of valyl-tRNA synthetase in the *twn2* mutant of Arabidopsis. *Proc Natl Acad Sci USA* 94, 7349-7355.

Summary

Although the zygotic embryogenesis is the primary mode of embryogenesis, plants possess a remarkable capacity to induce embryogenesis from a range of different cell types. **Chapter 1** introduces a “map” covering the many routes to embryo identity in plants and discusses a framework for sketching the developmental roles and mechanisms of plant embryogenesis regulators.

In **Chapter 2**, we exploited one of the alternative modes of embryogenesis, namely the transformation of extra-embryonic (suspensor) cells into embryo cells. This event can be triggered by inhibition of response to the plant hormone auxin in suspensor cells, and was used here as a model to determine transcriptome reprogramming underlying this dramatic cell fate conversion. We identified a set of 4 *bHLH* genes that are regulated by auxin during suspensor-derived embryogenesis. After further characterization of these genes, emphasis was placed on *bHLH49*, which appeared to be a direct transcriptional target repressed by auxin-dependent transcription factors. Importantly, *bHLH49* overexpression resulted in proliferation and even formation of multiple embryo-like structures in the suspensor, similar to the effect of suspensor-specific auxin response inhibition. Interestingly, a set of misregulated genes was shared upon either *bHLH49* overexpression or inhibition of auxin response in suspensor cells. This suggests that *bHLH49* contributes to the genetic program that is regulated by auxin to suppress embryogenesis in suspensor cells.

Several genes have already been reported as initiators of embryogenesis, but in very different experimental systems. In **Chapter 3**, using the suspensor as a highly predictable and simple model for embryo initiation, we systematically tested the ability of known embryonic regulators to induce suspensor-derived twin embryos. Strikingly, from 15 genes tested, in addition to *bdl* and *bHLH49*, only the *RKD1* transcription factor was able to induce embryogenesis in suspensor cells. This suggests that, rather than being a generic cellular response, suspensor to embryo transition is regulated by a specific genetic network. Starting from this observation, we established a genome-wide suspensor-specific activation tagging screen to identify novel genes that can induce embryogenesis when ectopically expressed in the suspensor. We selected several mutants with twin-like phenotype, but could unfortunately not validate that phenotypes were caused by local activation of identified genes in the suspensor. Nonetheless, this screen revealed that a more efficient method for selection and screening of transformants was needed.

In **Chapter 4**, we therefore re-designed and optimized our activation tagging screen by developing a smaller activation tagging vector harboring a red seed fluorescence

cassette that enables more efficient transformation and antibiotic-free selection of mutant phenotypes. Furthermore, we developed a simple protocol for high-throughput seed sorting generating large number of transgenic seeds within minutes. A next generation activation tagging screen was set up and tested. The pilot screen yielded genuine twin seedlings of which one could be confirmed to be genetically heritable, dominant and originated from suspensor-derived embryogenesis. This initial result shows that optimization was successful and that we now have an efficient procedure that should allow the systematic identification of novel “embryo inducers”.

Although the *GAL4/UAS* enhancer trap lines, including the one used as genetic background in the suspensor-specific activation tagging screen, are widely used to drive misexpression and as markers for cell identities, very few have been molecularly characterized. In **Chapter 5** we provide the molecular basis for the cell-type-specific expression pattern of set of 21 *GAL4/UAS*-GFP enhancer trap lines by mapping their insertion site. In addition, while thoroughly documenting their expression during embryogenesis as well as post-embryonically, we observed that their expression domain is often broader than usually reported. This should be taken as a cautionary note for future use of these lines in misexpression studies. Furthermore, we showed how the knowledge obtained in this Chapter could be used.

Finally, **Chapter 6** discusses how the results from this study contribute to what is known about the embryo initiation process, highlights questions that still remain and formulates challenges for future research. The results in this thesis collectively show that the fate change towards embryogenesis in plants is complex and it is regulated by a specific set of genes.

Samenvatting

Hoewel embryo's meestal vanuit een zygote worden gevormd, hebben planten het opmerkelijke vermogen om embryogenese ook in andere celtypes te induceren.

Hoofdstuk 1 introduceert een "landkaart" van de vele verschillende routes naar embryo-identiteit in planten en bespreekt een raamwerk voor de rol van embryogenese regulatoren in de ontwikkeling, en hoe deze regulatoren hun werk doen.

In **Hoofdstuk 2** staat één van de alternatieve manieren van embryogenese, namelijk de transformatie van cellen buiten het embryo (de suspensor) naar embryo cellen, centraal. Deze overgang kan geactiveerd worden door in suspensor cellen het vermogen om op het plantenhormoon auxine te reageren (auxine respons) te verminderen. De activatie van embryogenese in suspensor cellen is in dit hoofdstuk gebruikt als een model om te bepalen hoe transcriptoom reprogrammering deze dramatische wijziging van cel-identiteit stuurt. We hebben een set van 4 *bHLH* genen geïdentificeerd die gereguleerd worden door auxine tijdens embryogenese vanuit de suspensor. Nadat deze genen beter gekarakteriseerd waren, richtten we ons vooral op *bHLH49*, dat op directe wijze door auxine-afhankelijke transcriptiefactoren wordt onderdrukt. Over-expressie van *bHLH49* resulteerde in vermenigvuldiging van suspensor cellen en zelfs tot de vorming van meerdere embryo-achtige structuren in de suspensor. Dit lijkt sterk op het effect van auxine respons inhibitie in suspensor cellen. Opmerkelijk is, dat een aantal genen zowel misgereguleerd werden tijdens over-expressie van *bHLH49* als tijdens inhibitie van de auxine respons. Dit suggereert dat *bHLH49* bijdraagt aan een auxine-gereguleerd genetisch programma dat onder normale omstandigheden embryogenese in de suspensor onderdrukt.

Er zijn al een aantal genen bekend die embryogenese activeren, maar deze zijn in sterk uiteenlopende modelsystemen beschreven. In **Hoofdstuk 3** hebben we de suspensor gebruikt als model om te testen of deze bekende regulatoren in staat zijn embryogenese te induceren in de suspensor. Naast *bdl* en *bHLH49* hebben we 15 genen getest, waarvan alleen de transcriptiefactor *RKD1* in staat bleek om tweeling embryo's te vormen. Dit suggereert dat de omschakeling van suspensor naar embryo geen specifieke reactie op verstoring van het normale ontwikkelingsprogramma is, maar door specifieke regulatoren wordt gestuurd in de suspensor. Met deze observatie als uitgangspunt zijn we gestart met een genoom-brede "activation tagging" screen om nieuwe genen te vinden die, na overexpressie, embryogenese kunnen induceren in de suspensor. We hebben een aantal mutanten met tweeling fenotypes geselecteerd, maar konden helaas niet bevestigen dat dit fenotype kwam door locale activatie van genen in de suspensor. Desondanks heeft deze screen laten

zien dat er een efficiëntere manier nodig was om transformanten te screenen.

In **Hoofdstuk 4** hebben we daarom de activation tagging screen geoptimaliseerd door een kleinere vector met een rode zaad-fluorescentie selectie merker te gebruiken. Deze vector maakt het mogelijk om mutanten te selecteren zonder gebruik van antibiotica. Daarnaast hebben we een eenvoudig protocol geoptimaliseerd om grote aantallen transgene zaden binnen enkele minuten te sorteren. Deze nieuwe activation tagging screen is getest en resulteerde in de vondst van een aantal tweeling zaailingen. Een van deze mutante fenotypes was genetisch erfelijk, dominant en was afkomstig van embryogenese in de suspensor. Dit eerste resultaat laat zien dat de optimalisatie van de screen succesvol is, en dat we nu een efficiënte procedure hebben om systematisch nieuwe genen te identificeren die embryogenese induceren.

De GAL4/UAS enhancer trap lijnen, waarvan we een aantal in de activation tagging screen hebben gebruikt, worden vaak toegepast in mis-expressie studies en als cel-identiteit merkers. Er zijn echter maar een paar van deze lijnen op moleculair niveau gekarakteriseerd. In **Hoofdstuk 5** laten we zien hoe 21 GAL/UAS-GFP enhancer trap lijnen hun specifieke expressie patroon verkrijgen door hun insertie site te identificeren. Daarnaast hebben we de expressie van deze lijnen in detail bestudeerd tijdens embryogenese en verdere ontwikkeling van de plant. We zagen dat het expressie domein van deze genen vaak breder is dan eerder gerapporteerd is. Dit feit is belangrijk om rekening mee te houden in toekomstige mis-expressie studies.

Tot slot wordt in **Hoofdstuk 6** besproken wat de resultaten van deze studie bijdragen aan onze kennis over initiatie van embryogenese, welke grote vragen er nog zijn en welke uitdagingen er zijn voor verder onderzoek. De resultaten in deze thesis laten zien dat omschakeling naar embryogenese in planten een complex process is dat gereguleerd wordt door een specifieke set genen.

Acknowledgements

When I started my PhD journey, four years ago, I had no idea that I will have so much fun, learn so many things and meet so many fantastic people! Now at the end, I would like to thank all the people who accompanied me during this ‘once in a lifetime’ journey.

First of all, I would like to thank my supervisor and promoter Prof. dr. Dolf Weijers. Thanks Dolf for your support, scientific guidance, discussions and time. You are a great supervisor and a person, an example of an outstanding scientist. You inspired my curiosity to the unknown. I feel lucky to be part of your amazing team!

The next person I want to thank is Prof. dr. Sacco de Vries. Sacco, thank you for providing and improving the learning environment at the Biochemistry Department and for being the best student I have ever had :). I appreciate your help and really enjoy our informal discussions.

Furthermore, I would like to thank Dr. Catherine Albrecht who contributed enormously to my project. Cathy, thank you very much for all the help and fruitful discussions. Without your input and hard work the ‘Twin Project’ could not have progressed smoothly. Your help is priceless!

And thanks to all the members of Weijers’ group. I could not have asked for better colleagues and fellow PhDs. Thank you all for the unforgettable time, discussions, social activities and for the enormous amount of candies and cookies that we shared. I would specially like to thank the former and present tenants of the ‘Ladies office and Kuan-Ju’. I really enjoyed the time spent with you and our scientific and non-scientific discussions. In addition, I want to thank Willy and Bert for all the help in the lab and the scientific advice. Bert, I am looking forward to reading all the amazing papers coming from the ‘Vascular team’ in Ghent. Colette and Jos, bedankt dat jullie altijd Nederlands tegen me wilden praten.

To my paranymphs, Cristina and Maritza, I am glad I met you and that I have you as friends. Cris, thank you very much for your help in the lab and for your patient answers when I was lost in the beginning. I really enjoyed working with you and now enjoy our coffee times and walks. Maritza, you are a great colleague and friend! Thank you for listening to all my complaints and frustrations, and for always being there! I wish you the best with your thesis, I know you have a bright scientific future.

A big thank you to all of the other members of the Biochemistry Department, especially to Laura van Egmond and Dr. Jan Willem Borst for help with bureaucracy

Acknowledgements

and microscopy. Mieke, Joseline, Gudrun and Antsje, thank you for teaching me how to make SDS-PAGE gels and how to handle protein work :).

Another friend who I met during my PhD is Christoph, thank you for being such a good friend in the last four years and for your help. I really enjoyed the coffee breaks and our conversations.

I also want to say thank you to Dr. Juliane Taepal. Jule, thank you for helping me with the seed sorting and for the pleasant conversations we had. I hope we will stay in touch.

I joined the EPS PhD Council during my last year, which was an awesome distraction. Suraj, Setareh, Tom, Amalia, Mannos, Sara, Elysa, Chrisa, Magda and Francesca thank you for the great meetings. I had a lot of fun and learned a lot! Good luck finishing your theses.

I would especially like to thank Dr. Mira Zhiponova. Mira, thank you very much for your help during my stay in Ghent, thank you for your help when I was looking for a PhD position. You were actually the person who brought my attention to Dolf's group and who encouraged me to apply for this position. I wish you all the best in your professional and personal life!

To my Bulgarian friends, thanks for the great time, for your interest in my research, for the motivation and support.

Further, I would like to thank my soulmate Tony. Thank you for always believing in me and for encouraging me to go ahead, for your support, your love, and for making all my dreams come true. I know it was fate that we met and I hope we will get old together. Обичам те с цялото си сърце!

And last but not least, I would like to thank my parents and my sister Sonya. Thank you for your support all these years, for giving me the opportunity to study and for teaching me how to be a good person. You made me the person who I am today. I know you will always be there for me!

Curriculum vitae

Tatyana Radoeva was born on 22th of July 1986, in Devin, Bulgaria. In 2005, she finished her secondary education and afterwards she moved to Sofia to study Biotechnology at Sofia University. After graduating with a Bachelor Degree of Science in 2009, she started studying Industrial Biotechnology. She did her Master thesis in the group of Prof. Dr. Dominique Van Der Straeten, at Ghent University in Belgium. During this thesis, she focused on interactions between brassinosteroids and ethylene during negative gravitropism in *Arabidopsis*. In 2011, she obtained her Master Degree of Science with honors. Shortly thereafter, she moved to the Netherlands and in April 2012 she started a PhD project in the group of Prof. dr. Dolf Weijers at Wageningen University. In this project, she studied how embryo formation can be induced in plants, and the results obtained over the four years of her PhD are presented in this thesis. Following her PhD, she will continue her scientific career in the same group as a post-doc.



Publications

- Radoeva, T.**, ten Hove, C.A., Saiga, S. and Weijers, D. (2016). Molecular characterization of Arabidopsis GAL4/UAS enhancer trap lines identifies novel cell-type-specific promoters. *Plant Physiol* 171, 1-13.
- Radoeva, T.***, Lokerse, A.S.*, Llavata-Peris, C.I., Wendrich, J.R., Xiang, D., Liao, C.-Y., Borst, J.W., Vlaar, L., Boekschoten, M., Hooiveld, G., Datla, R. and Weijers, D. (2016). A novel ARF/bHLH module regulates extra-embryonic identity during Arabidopsis embryogenesis. *Plant Cell*, Under revision. *equal author contribution
- Freire-Rios, A., **Radoeva, T.**, De Rybel, B., Weijers, D. and Borst, J.W. (2016). FRET-FLIM for visualizing and quantifying protein interactions in live plant cells. *Methods Mol Biol*, In press.
- Wendrich, J.R., Moller, B.K., Uddin, B., **Radoeva, T.**, Lokerse, A.S., De Rybel, B., and Weijers, D. (2015). A set of domain-specific markers in the Arabidopsis embryo. *Plant Reprod* 28, 153-160.
- Radoeva, T.** and Weijers, D. (2014). A roadmap to embryo identity in plants. *Trends Plant Sci* 19, 709-716.

Education Statement of the Graduate School Experimental Plant Sciences



Issued to: Tatyana Radoeva
Date: 21 June 2016
Group: Laboratory of Biochemistry
University: Wageningen University & Research

1) Start-up phase	date
▶ First presentation of your project Mechanistic dissection of plant embryo initiation	Nov 22, 2012
▶ Writing or rewriting a project proposal	
▶ Writing a review or book chapter A roadmap to embryo identity in plants, Radoeva, T. and Weijers, D., Trends in Plant Science (2014), Vol 19, issue 11, p709-716. DOI: 10.1016/j.tplants.2014.06.009	2013-2014
▶ MSc courses	
▶ Laboratory use of isotopes	
Subtotal Start-up Phase	7.5 credits*
2) Scientific Exposure	date
▶ EPS PhD student days EPS PhD student day, Amsterdam University	Nov 30, 2012
EPS PhD student day, Leiden University	Nov 29, 2013
EPS PhD student day, Get2Gether, Soest, the Netherlands	Jan 29-30, 2015
EPS PhD student day, Get2Gether, Soest, the Netherlands	Jan 28-29, 2016
▶ EPS theme symposia EPS theme 1 'Developmental Biology of Plants', Leiden University	Jan 17, 2013
EPS theme 1 'Developmental Biology of Plants', Wageningen University	Jan 24, 2014
EPS theme 1 'Developmental Biology of Plants', Leiden University	Jan 08, 2015
EPS theme 1 'Developmental Biology of Plants', Wageningen University	Jan 21, 2016
▶ Lunteren days and other National Platforms Annual Meeting 'Experimental Plant Sciences', Lunteren	Apr 02-03, 2012
Annual Meeting 'Experimental Plant Sciences', Lunteren	Apr 22-23, 2013
Annual Meeting 'Experimental Plant Sciences', Lunteren	Apr 14-15, 2014
Annual Meeting 'Experimental Plant Sciences', Lunteren	Apr 13-14, 2015
▶ Seminars (series), workshops and symposia Thursday Seminars Biochemistry: Masahiko Futurani (NAIST, Nara, Japan)	Apr 20, 2012
Thursday Seminars Biochemistry: Jürgen Kleine-Vehn (BOKU, Vienna, Austria)	Apr 24, 2012
Thursday Seminars Biochemistry: Michael Nodine (Whitehead Institute, Boston, USA)	May 11, 2012
Thursday Seminars Biochemistry: Hong Ma (Fudan University, Shanghai, China)	May 29, 2012
Thursday Seminars Biochemistry: Doris Wagner (University of Pennsylvania, Philadelphia, USA)	Jul 10, 2012

Education Statement

Thursday Seminars Biochemisrty: David Robinson (Heidelberg University, Germany)	Sep 27, 2012
Minisymposium 'Frontiers in Plant Morphogenesis' (Wageningen, The Netherlands)	Nov 13, 2012
Thursday Seminars Biochemisrty: M.G. Smits (Hospital Gelderse Valley, Ede, The Netherlands)	Mar 28, 2013
Thursday Seminars Biochemisrty: Marcus Grebe (Umeá Plant Science Center, Sweden)	May 29, 2013
Thursday Seminars Biochemisrty: Pierre Hilson (LJPB Versailles, France)	Nov 13, 2013
Thursday Seminars Biochemisrty: Marcus Heisler (EMBL, Heidelberg, Germany)	Nov 28, 2013
Thursday Seminars Biochemisrty: Daniël Van Damme (VIB, Ghent, Belgium)	Jan 30, 2014
Thursday Seminars Biochemisrty: Ivonne Stahl (Heinrich Heine University, Düsseldorf, Germany)	Mar 11, 2014
Thursday Seminars Biochemisrty: Goerge Bassel (University of Birmingham, United Kingdom)	Apr 24, 2014
Thursday Seminars Biochemisrty: Cyril Zipfel (Sainsbury Laboratory, United Kingdom)	Jun 04, 2014
EPS Symposium 'Omics Advances for Academia and Industry - Towards True Molecular Plant Breeding'	Dec 11, 2014
Thursday Seminars Biochemisrty: Saijalisa Kangasjärvi, University of Turku, Finland	May 28, 2015
Thursday Seminars Biochemisrty: Martin Jinek, University of Zürich, Switzerland	Jun 06, 2015
Joint Meeting groups prof. Dolf Weijers and prof. Ben Scheres, Wageningen	Sep 25, 2015
Thursday Seminars Biochemisrty: François Parcy, IRTSV, Grenoble, France	Oct 15, 2015
Thursday Seminars Biochemisrty: Julia Santiago, University of Geneva, Switzerland	Nov 12, 2015
Joint Meeting groups prof. Dolf Weijers and prof. Ben Scheres, Wageningen	Jan 22, 2015
▶ Seminar plus	
Seminar + Masterclass: Detlef Weigel (Max Planck Institute for Developmental Biology, Germany)	Feb 27, 2013
Seminar + Masterclass: Siobhan Brady, University of California, USA	Sep 09, 2015
▶ International symposia and congresses	
Embryo meeting 2013, Haigerloch, Germany	May 06-08, 2013
Auxin Sailing 2013, Leiden, The Netherlands	Jun 08-09, 2013
European Frontiers of Plant Reproduction Research, Oslo, Norway	Oct 02-04, 2013
Midterm Meeting NSFC/NWO: Development of Plants, Utrecht, The Netherlands	Oct 24-25, 2013
Auxentric 2014, Norwich, United Kingdom	May 24-25, 2014
Auxins and Cytokinins in Plant Development 2014, Prague, Czech Republic	Jun 29-Jul 04, 2014
Plant Organ Growth Symposium, Ghent, Belgium	Mar 10-12, 2015
FASEBScience Research Conference 'Mechanisms in Plant Development', Saxtons River, Vermont (USA)	Aug 02-07, 2015
▶ Presentations	
Annual Meeting Experimental Plant Sciencies, Lunteren, The Netherlands (Poster)	Apr 22, 2013
Embryo meeting, Haigerloch, Germany (Talk)	May 07, 2013
European Frontiers of Plant Reproduction Research, Oslo, Norway (Talk)	Oct 09, 2013
EPS Theme 1 'Developmental Biology of Plants' (Talk)	Jan 24, 2014
Auxentric, Norwich, United Kingdom (Talk)	May 25, 2014
Auxins and Cytokinins in Plant Development, Prague, Czech Republic (Poster)	Jun 29-Jul 04, 2014
Annual Meeting Experimental Plant Sciencies, Lunteren, The Netherlands (Poster)	Apr 13-14, 2015
▶ IAB interview	

▶ Excursions	
Scientific PhD Excursion - group Sacco de Vries, United Kingdom	Aug 28- Sep 04, 2013
Excursion to Rijk Zwaan	Sep 27, 2013
Excursion to Enza Zaden	Jun 12, 2015

Subtotal Scientific Exposure 25.2 credits*

3) In-Depth Studies		date
▶ EPS courses or other PhD courses		
EPS PhD Course: Bioinformatics: A User's approach		Aug 25-29, 2014
EPS PhD course: Transcription Factors and Transcription Regulation		Dec 17-19, 2013
▶ Journal club		
Participation in weekly journal club (Dolf Weijers Group)		Apr 2012-Apr 2016
▶ Individual research training		

Subtotal In-Depth Studies 5.5 credits*

4) Personal development		date
▶ Skill training courses		
Language course: Dutch B2		Sep 25-Dec 18, 2012
WGS course Information Literacy including Endnote		Oct 28-29, 2014
Educational Staff Development Course: Teaching and Supervising Thesis Students		Mar 18-19,2013
WGS course: Reviewing a Scientific Paper		Nov 08, 2013
WGS course: Last Stretch of the PhD Programme		Feb 26, 2016
▶ Organisation of PhD students day, course or conference		
Organization of Excursion to 'Enza Zaden'		Jun 12, 2015
Organization of EPS 'GET2GETHER' event		2016
▶ Membership of Board, Committee or PhD council		
Secretary of EPS PhD Council		Apr 2015- Apr 2016
Member of WPC Council		2015- 2016

Subtotal Personal Development 8.8 credits*

TOTAL NUMBER OF CREDIT POINTS*	47.0
---------------------------------------	-------------

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

* A credit represents a normative study load of 28 hours of study.

This research presented in this thesis was performed at the Laboratory of Biochemistry, Wageningen University, and was financially supported by the Netherlands Organization for Scientific Research (NWO; ALW-NSFC Plant Development Collaborative Grant 846.11.001).

Cover design: Iliana Boshoven-Gkini || AgileColor.com

Layout design: Tatyana Radoeva and Iliana Boshoven-Gkini || AgileColor.com

Printed by: GVO drukkers & vormgevers, Ede (NL) || gvo.nl

