

BABY BOOM-induced
somatic embryogenesis
in Arabidopsis

Anneke Horstman

Thesis committee

Promotor

Prof. Dr G.C. Angenent

Personal chair at the Laboratory of Molecular Biology

Wageningen University

Co-promotor

Dr K.A. Boutilier

Senior Researcher, PRI Bioscience, WUR

Other members

Prof. Dr S.C. de Vries, Wageningen University

Prof. Dr M.D. Gaj, University of Silesia, Katowice, Poland

Dr V.A. Willemsen, Wageningen University

Dr E. Larsson, Leiden University / Swedish University of Agricultural Sciences, Uppsala, Sweden

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

BABY BOOM-induced somatic embryogenesis in Arabidopsis

Anneke Horstman

Thesis

submitted in fulfillment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus

Prof. Dr M.J. Kropff,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Friday 10 April 2015

at 4 p.m. in the Aula.

Anneke Horstman

BABY BOOM-induced somatic embryogenesis in Arabidopsis

234 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)

With references, with summaries in Dutch and English

ISBN 978-94-6257-231-7

Contents

<i>Chapter 1</i>	7
Regulation of in vitro somatic embryogenesis	
<i>Chapter 2</i>	27
AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks	
<i>Chapter 3</i>	59
AIL and HDG proteins act antagonistically to control cell proliferation	
<i>Chapter 4</i>	99
BABY BOOM and PLETHORA2 induce somatic embryogenesis in a dose- and context-dependent manner via the LAFL pathway	
<i>Chapter 5</i>	131
Genome-wide binding site analysis of BABY BOOM during somatic embryogenesis in Arabidopsis	
<i>Chapter 6</i>	159
Microarray-based identification of transcription factor target genes	
<i>Chapter 7</i>	187
A cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies	
<i>Chapter 8</i>	207
General discussion	
<i>Summary</i>	219
<i>Samenvatting</i>	223
<i>Acknowledgements</i>	227
<i>Curriculum vitae</i>	231
<i>Education statement</i>	232



Chapter 1

Regulation of *in vitro*
somatic embryogenesis

Introduction

During sexual reproduction, a diploid zygote is formed upon fusion of two haploid gametes, an egg cell and a sperm cell, and goes on to form the embryo and eventually a new plant. In flowering plants, the embryo develops together with a second fertilization product, the endosperm, and both the embryo and endosperm are surrounded by the maternal sporophytic tissue, the seed coat, which is derived from the ovule integuments. The seed coat, embryo and endosperm constitute a seed. During seed germination, the embryo breaks out of the seed coat and develops further to produce the different organs that make up the plant body. Thus effectively, the single-celled zygote has the capacity to form a whole plant with different tissues, and is therefore totipotent. Many plant cells other than the zygote have the capacity to become totipotent, including differentiated cells. This remarkable ability is referred to as totipotency and has intrigued scientists for decades (Reinert, 1958; Steward et al., 1958).

Somatic embryogenesis (SE), the development of embryos from somatic or vegetative cells, is one form of plant cell totipotency. SE occurs naturally *in planta* in a small number of plant species, and can also be induced *in vitro*. SE is characterized by the production of a bipolar structure with an apical pole (the future shoot) and a basal pole (the future root) and its own independent provascular system. Somatic embryos can therefore be distinguished from adventitious organs, such as shoots and roots, which are unipolar structures with a vascular connection to the underlying tissue. Somatic embryos also accumulate species-specific storage products that are not found at other stages of plant development and generally lack the trichomes found on vegetative tissues. A classic example of ‘natural’ SE is adventitious embryony or sporophytic apomixis, the formation of embryos from parts of the ovule other than the egg cell, usually from the nucellus or the integuments (Bicknell and Koltunow, 2004). This is an asexual process, but it may require fertilization of the central cell for endosperm production. Adventitious plantlets also form on the leaf margins of some plants e.g. *Kalanchoë daigremontiana*, also known as “mother of thousands”, although the identity of these leaf-derived plantlets has long been under debate. *Kalanchoë* plantlets have a vasculature system that is independent of the maternal tissue, but they only comprise an apical pole and produce adventitious roots later in development. Therefore, this form of plantlet formation is considered a combination of SE and organogenesis (Garces et al., 2007).

SE can also be triggered *in vitro* by exposing explants to stress treatments or exogenous growth regulators. SE was first demonstrated in tissue culture by Steward *et al.* (1958) and Reinert (1958; (Reinert, 1958; Steward *et al.*, 1958) using carrot cell suspension cultures derived from root tissue. A few years later, the importance of the synthetic auxin and herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) for the induction of SE was shown by Halperin and Wetherell (Halperin and Wetherell, 1964). Since then, SE protocols using different stress treatments, growth regulators and explants have been developed for a wide variety of species. These protocols are now routinely used for mass clonal propagation and embryo rescue in plant breeding. For example, SE can be used to asexually propagate male-sterile, open-pollinated, highly heterozygous or sterile polyploid lines. Once induced, somatic embryos can be directly grown into plantlets or stored as artificial/synthetic seeds (Sharma *et al.*, 2013). Somatic embryos can be converted to soil-grown plants in a much shorter time frame and with less labor input than conventional methods that require lengthy periods of shoot and root organogenesis of explants or cuttings. Additional advantages of SE include the unlimited multiplication capacity, possibilities for automation in bioreactors and immortalization of juvenile cultures and mature embryos by cryopreservation. Such advantages have led to significant increases in the production efficiency and -uniformity, and in the quality of crop germplasm, especially in the forestry industry (Fenning and Park, 2014). However, the use of SE for clonal propagation can be limited by the low responsiveness of many species and genotypes and by the production of 'off-types' or 'non true-to-type' embryos, usually resulting from somatic mutations or stable chromatin modifications (Miguel and Marum, 2011).

SE is also used as a research tool to study the mechanism of plant cell totipotency and embryo development, since somatic embryos are easier to access and to generate in bulk than zygotic embryos. However, there are some differences between somatic and zygotic embryogenesis. Zygotic embryogenesis starts with fertilization of the egg cell, after which (in most species) the zygote divides asymmetrically to form a smaller apical cell (the future embryo proper) and a larger basal cell (the future suspensor and hypophysis). The embryo then undergoes coordinated cell division and tissue patterning, maturation and desiccation. In contrast, SE starts from a single somatic cell or a group of cells and cell division and patterning is initially much less regular than during zygotic embryogenesis, yet, somatic embryos go through similar embryo stages and at the end resemble their zygotic counterparts morphologically (Zimmerman, 1993; Mordhorst *et al.*, 1997). At the maturation stage, somatic embryos do not undergo desiccation and dormancy as zygotic embryos do, but maturation can be induced by

application of abscisic acid (ABA) (Gutmann et al., 1996; Tian and Brown, 2000; Vahdati et al., 2008). Interregional communication between the different seed tissues plays a role in guiding the development of the zygotic embryo, not only for the transport of nutrients and hormones from the surrounding tissues, but also for patterning (Weijers et al., 2003; Yang et al., 2008; Costa et al., 2014); however, somatic embryos produced *in vitro* are not surrounded by endosperm or maternal seed tissues. The ability of somatic embryos to develop in the absence of the surrounding seed tissues suggests that either interregional communication between seed tissues is only relevant in the context of seed development *in planta*, or that somatic embryo development relies on the tissue culture medium and/or the original explant for developmental cues. Alternatively, other cell types present in the culture or even the embryo itself takes over the functions of the filial and maternal tissues (Van Hengel et al., 1998; Wiweger et al., 2003).

From explant to somatic embryo

The success of a SE protocol relies heavily on the type of explant that is used as starting material for the culture. The cells of an explant should be 'competent', meaning they can respond to the stimulus and switch to the embryogenic pathway. In most cases, embryonic (e.g. immature zygotic embryos) or juvenile (e.g. cotyledons) tissues are used as explants for *in vitro* SE induction, although in some cases differentiated tissues can be used (e.g., leaf mesophyll protoplasts). The synthetic auxin 2,4-D or a combination of plant hormones is commonly used to induce SE, but abiotic stress factors, including osmotic pressure, pH and low or high temperatures, can also trigger SE (Gaj, 2004).

Somatic embryos can develop directly from the explant, for example, from the epidermal and cortex cells (Dubois et al., 1990). Direct SE starts with cell proliferation and the formation of compact clusters of small, cytoplasm-rich cells (Bassuner et al., 2007; Kurczynska et al., 2007). These embryonic cell clusters can be first recognized by their thick cell walls, which might isolate the embryo from the underlying explant (Kurczynska et al., 2007). Expression of an embryogenesis marker gene (*LEC2*) was observed in few-celled clusters, indicating that these cells have switched to embryo identity (Kurczynska et al., 2007). Somatic embryos can be distinguished more clearly once the protoderm is formed. At later stages, the embryos elongate, undergo apical-basal differentiation and form well-defined shoot and root meristems. Unlike

adventitious shoots, somatic embryos are not connected to the underlying vascular tissue of the explant.

Somatic embryos can also develop indirectly from a proliferating tissue referred to as callus (Reinert, 1958; Steward et al., 1958). Due to its amorphous structure, callus is historically thought to comprise 'undifferentiated' cells, however recent analysis shows that organogenic callus (for adventitious shoot production) has a lateral root identity and is derived from pericycle cells of the root or pericycle-like cells from other organs (Che et al., 2007; Atta et al., 2009; Sugimoto et al., 2010). This suggests that the first step in *de novo* shoot organogenesis first involves direct cell redifferentiation to a distinct cell type. It is not known whether embryogenic callus has a lateral root identity. However, like shoot callus, embryogenic callus can also be derived from pericycle-like cells (Sticklen, 1991; Yang et al., 2010). During indirect SE, aggregates of cytoplasmically-rich proliferating cells, referred to as pro-embryogenic masses (PEMs), are formed within the callus (De Vries et al., 1988). To complete the transition from PEM to a somatic embryo, the apical-basal and bilateral patterning of the embryo needs to be established and the embryo needs to elongate, processes that often require the removal of auxin from the culture medium. Somatic embryos that form via direct or indirect SE are morphologically similar, but changes in the genome (somaclonal variation) due to the longer tissue culture period often occurs in embryos derived from indirect SE (Gaj, 2004).

By far the majority of research on SE is focussed on protocol development, however, a number of different species have also been used to understand the mechanism driving somatic embryo induction including carrot, Norway spruce, alfalfa and cotton. *Arabidopsis thaliana* (Arabidopsis) is perhaps one of the best systems to study SE due to the availability of efficient protocols from different explants, and the wealth of cell biology and functional genomics tools. Therefore, Arabidopsis SE is discussed below in more detail, and data from other model systems are included where relevant.

Arabidopsis as a model system for somatic embryogenesis

Several *in vitro* somatic embryo induction systems have been developed for *Arabidopsis*, using a wide range of explants including leaf protoplasts (Luo and Koop, 1997), immature zygotic embryos (IZEs) (Pillon et al., 1996; Mordhorst et al., 1998; Gaj, 2001), mature zygotic embryos (Kobayashi et al., 2010), shoot apices and flower buds (Ikeda-Iwai et al., 2003) (Figure 1). In general, auxin (2,4-D) is used to trigger SE, although a number of stress treatments are also effective, including osmotic, heavy metal, and dehydration stress (Ikeda-Iwai et al., 2003). In addition, primary somatic embryos can be used as explants to produce secondary somatic embryos via callus (Su et al., 2009). The most extensively used and studied *Arabidopsis* SE system uses 2,4-D –treated IZEs (Figure 1). In this system, the cotyledons of the zygotic embryo first swell due to divisions in the procambium and then, depending on the culture conditions, somatic embryos either develop directly from dividing protodermal and subprotodermal cells or indirectly via callus derived from the same tissue layers (Raghavan, 2004). Somatic embryos also develop side-by-side with adventitious shoots, and it is thought that a minor, loose connection to the explant allows local auxin accumulation and root formation in somatic embryos, while a broader connection to the explant leads to the development of shoots due to auxin flow and the formation of a continuous vascular connection with the explant (Bassuner et al., 2007). The developmental stage of the IZE explant greatly influences the type of SE; early globular/torpedo stage embryos undergo indirect SE, while older cotyledonary stages undergo direct SE (Gaj, 2011). *Arabidopsis* somatic embryos derived from cotyledonary IZEs differ in structure from zygotic embryos, as their division pattern is initially unorganized, and apical-basal and bilateral symmetry is established after a higher number of cell divisions (Bassuner et al., 2007). On the other hand, somatic embryos derived from heart-staged IZEs or from leaf protoplasts follow a division pattern that is very similar to that of zygotic embryos and occasionally even form suspensors (Luo and Koop, 1997). Finally, the maturation (elongation and further patterning) of somatic embryos may require the removal of auxin in case of indirect SE (Raghavan, 2004; Bassuner et al., 2007).

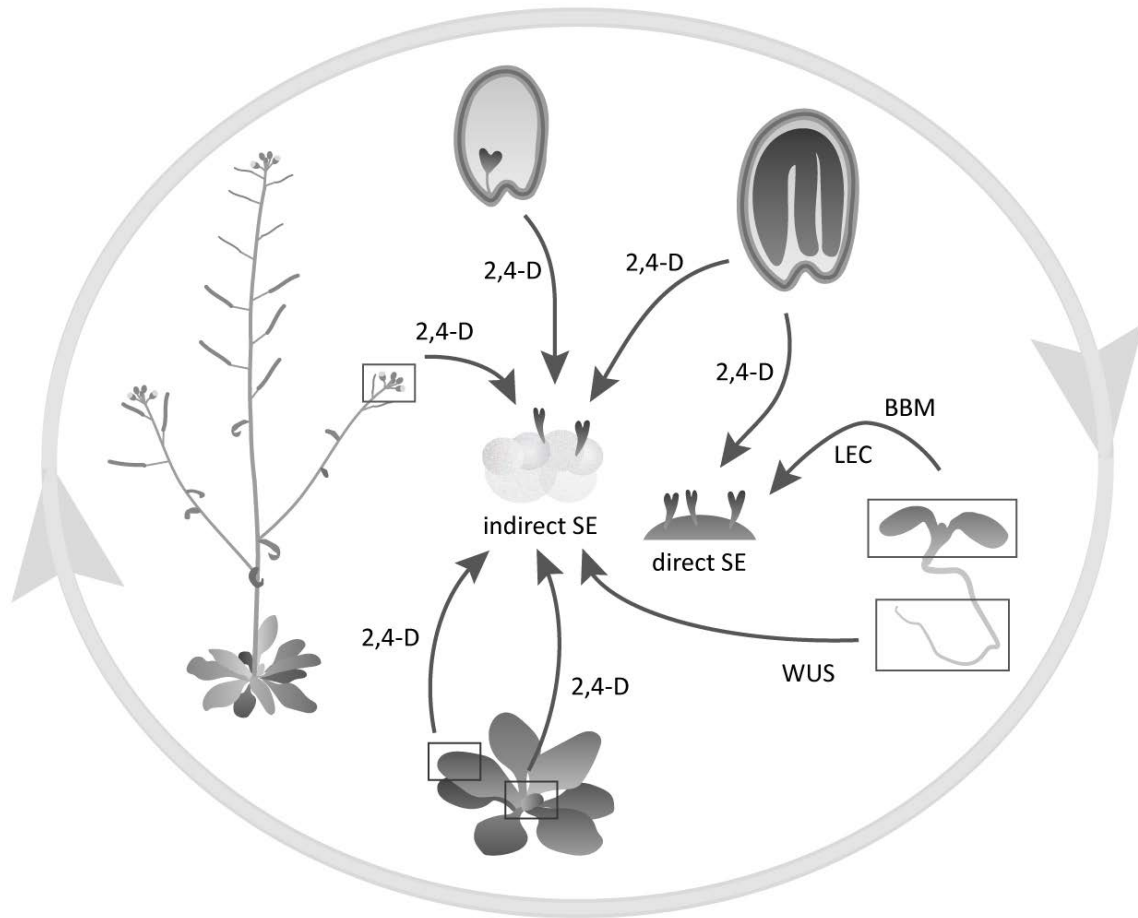


Figure 1. Overview of somatic embryogenesis systems in Arabidopsis.

Arabidopsis SE can be induced from a range of tissues throughout the Arabidopsis life cycle. SE can be induced either directly or indirectly via callus. Most systems use 2,4-D (auxin) as the inducing agent, but overexpression of specific transcription factors is also effective. Only transcription factors that can induce SE without exogenous application of hormones are shown: BABY BOOM (BBM), WUSCHEL (WUS) and LEAFY COTYLEDON (LEC).

Molecular control of somatic embryo induction

Molecular-genetic studies in Arabidopsis have identified genes that are important for SE, and the interactions between these components are becoming increasingly clear. Below, I describe these 'SE factors' and their molecular interactions.

The role of auxin

Most SE protocols rely on the exogenous application of 2,4-D, which is a stable, non-metabolized and poorly transportable auxin (Delbarre et al., 1996; Morris et al., 2004; Hosek et al., 2012). However, it is not known how 2,4-D triggers somatic cells to enter the embryogenesis pathway. It was shown that during indirect SE in *Arabidopsis*, removal of 2,4-D after callus induction leads to an increased expression of auxin biosynthesis genes and to a higher level of endogenous auxin before somatic embryos had formed (Bai et al., 2013). Also in other species, explants that produce somatic embryos accumulate endogenous auxin (Michalczyk et al., 1992; Charriere et al., 1999; Pasternak et al., 2002). In line with this, suppression of a type-2 hemoglobin was shown to promote SE by enhancing auxin biosynthesis (Elhiti et al., 2013). In contrast, knock-out mutants in auxin biosynthesis genes, such as *YUCCA* (*YUC*) genes, show a reduced response in 2,4-D-treated cultures (Bai et al., 2013), indicating the importance of auxin production. In addition, mutants defective in auxin response, such as the *axr* mutants, produce less somatic embryos in 2,4-D cultures (Gaj et al., 2006). Finally, it was shown by Su et al. (Su et al., 2009) that shortly after removal of 2,4-D in an indirect SE system, regions of high and low auxin levels are established on the edge of the callus. The shoot meristem marker *WUSCHEL* (*WUS*) then becomes expressed in areas with low auxin levels, which is followed by auxin accumulation and somatic embryo development at these sites. It was shown that establishment of auxin gradients and polar localization of the auxin efflux carrier PIN1 are required for SE.

Chromatin modification proteins suppress embryo identity

DNA associates with histone proteins to form chromatin. The structural unit of chromatin is the nucleosome, which consists of 147 base pairs of DNA wrapped around an octamer of histones. Modification of histone tails or re-positioning of nucleosomes can alter DNA accessibility and gene transcription. A network of chromatin modifying proteins repress embryo identity during germination. Removing or limiting the function of these chromatin modifying proteins results in the formation of somatic embryos on seedlings through failure to repress the embryo pathway during germination.

Acetylation of histone tails is associated with open chromatin and active gene transcription and this mark can be removed by histone deacetylases (HDACs). The *Arabidopsis*

HDACs HDA6 and HDA19 were shown to control the embryo-to-seedling transition during germination. In the *hda6;hda19* double mutant, the embryo program is not suppressed, which results in the formation of somatic embryos on leaves (Tanaka et al., 2008). During normal germination, the embryo-specific genes *LEC1* and *LEC2* (discussed below) are repressed, while in the *hda19* mutant, *LEC1* and *LEC2* contain more histone acetylation marks, as well as other marks for transcriptionally active chromatin, suggesting that HDA19 directly represses expression of these genes during germination (Zhou et al., 2013).

The embryo-to-seedling transition is also regulated by Polycomb (PcG) protein complexes (Hennig and Derkacheva, 2009). Like HDACs, PcG complexes can also confer transcriptional repression through histone modifications. There are two classes of PcG complexes: Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2), which have histone monoubiquitination and histone methyltransferase activity, respectively (Bratzel et al., 2010; Schuettengruber et al., 2011). RING1 and BMI1 proteins form the Arabidopsis PRC1 complex and *bmi1a;bmi1b* and *ring1a;ring1b* double mutant seedlings have elevated expression of embryo-specific genes and form somatic embryos (Bratzel et al., 2010; Chen et al., 2010). Components of the PRC2 complex in Arabidopsis that repress the embryo pathway during vegetative growth include CURLY LEAF (CLF), SWINGER (SWN), EMBRYONIC FLOWER2 (EMF2), VERNALIZATION2 (VRN2) and FERTILIZATION-INDEPENDENT ENDOSPERM (FIE). The corresponding *vrn2;emf2* (Schubert et al., 2005), *clf;swn* (Chanvivattana et al., 2004) and *fie* (Bouyer et al., 2011) mutant seedlings form somatic embryos.

Another way of changing the chromatin structure and influencing gene expression is through re-positioning of nucleosomes, a process carried out by ATP-dependent chromatin remodelers (Whitehouse et al., 1999). *PICKLE (PKL)*, a CHD3-type ATP-dependent chromatin remodeling factor, also represses *LEC* genes in seedlings after germination (Ogas et al., 1999; Dean Rider et al., 2003; Zhang et al., 2012). The *pkl* mutant shows embryonic traits after germination, including seed storage product accumulation, and produces somatic embryos from a number of seedling tissues (Ogas et al., 1997; Ogas et al., 1999; Henderson et al., 2004).

In conclusion, different chromatin modification and remodeling proteins repress expression of embryo-specific genes after germination to promote the embryo-to-seedling transition. Loss of function mutants for some of these proteins promote SE in seedlings, most likely due to their inability to completely repress the embryo phase of development.

Somatic embryogenesis through ectopic gene expression

A number of genes have been identified that, when overexpressed, enhance the ability of auxin-treated explants to form somatic embryos. Overexpression of the MADS-box transcription factor gene *AGAMOUS-LIKE 15 (AGL15)* facilitates SE formation from IZEs and germinating seeds cultured with 2,4-D, as well as from IZEs in the absence of 2,4-D (Harding et al., 2003). The ability of *AGL15* to promote SE is partly due to a reduction in active gibberellin (GA) levels through activation of a GA2-oxidase (Wang et al., 2002; Wang et al., 2004). In addition, *AGL15* overexpression leads to an increased expression of *SOMATIC EMBRYOGENESIS RECEPTOR KINASE1 (SERK1)*, which also promotes SE induction in 2,4-D cultured IZEs when overexpressed (Hecht et al., 2001). It was proposed that *SERK1*, which is normally expressed in the procambium and in transit amplifying cells, maintains a population of pluripotent cells in these tissues and that treatment with 2,4-D can trigger totipotency in these cells (Kwaaitaal and de Vries, 2007). *AGL15* and *SERK1* were identified in the same protein complex that include components of the brassinosteroid signalling pathway (Karlova et al., 2006), suggesting that brassinosteroid signalling may play a role in SE induction. Other genes that have been shown to promote SE are *Brassica* orthologs of *Arabidopsis SHOOT MERISTEMLESS (STM)*, an important regulator of shoot apical meristem development. Ectopic expression of these genes improves somatic embryo production from 2,4-D treated IZEs, probably due to an enhanced sensitivity to 2,4-D (Elhiti et al., 2010). Finally, the MYB domain transcription factors genes *PGA37/MYB118* and *MYB115* were identified in an activation tagging screen for mutants that triggered a switch to embryogenesis in auxin (IAA)-induced callus in root explants (Wang et al., 2009).

SE can also be induced in *Arabidopsis* seedlings grown without exogenous growth regulators, through ectopic expression of certain transcription factors, including the AP2 domain transcription factor *BABY BOOM* (Boutilier et al., 2002), the CCAAT-box binding factor *LEAFY COTYLEDON1 (LEC1)* (Stone et al., 2001), the B3 domain protein *LEC2* (Lotan et al., 1998) and the homeobox protein *WUSCHEL (WUS)*; (Zuo et al., 2002; Gallois et al., 2004)) (Figure 1).

WUS was identified as *PLANT GROWTH ACTIVATOR 6 (PGA6)* in the same activation tagging screen described above (*MYB115/MYB118*; (Wang et al., 2009)) to find genes that induce SE from root callus (Zuo et al., 2002). *WUS* is a homeodomain transcription factor and its overexpression leads to organogenesis and SE in the shoot and root tips (Zuo et al., 2002; Gallois

et al., 2004). MYB115 and MYB118 did not induce SE via a WUS-dependent pathway (Wang et al., 2009). At present, the mechanism of WUS-induced SE is unknown.

The LEC proteins LEC1 and LEC2, together with ABSCISIC ACID-INSENSITIVE3 (ABI3), FUSCA3 (FUS3) and LEC1-LIKE (L1L), constitute the LAFL network that controls embryo morphogenesis and maturation via complex cross-regulatory interactions (Jia et al., 2013). Loss-of-function mutations in *LAFL* genes result in defects in cotyledon development, storage macromolecule accumulation and desiccation tolerance in zygotic embryos (Keith et al., 1994; Meinke et al., 1994; West et al., 1994; Parcy et al., 1997; Stone et al., 2001), and in a severely reduced somatic embryo induction by 2,4-D (Gaj et al., 2005). In contrast, ectopic expression of *LEC1* and *LEC2* induces SE on Arabidopsis seedlings (Lotan et al., 1998; Stone et al., 2001), while overexpression of *FUS3* leads to the formation of cotyledon-like leaves (Gazzarrini et al., 2004). *LEC2* directly activates the above-mentioned *AGL15* gene, and both *LEC2* and *AGL15* upregulate *INDOLE-3-ACETIC ACID INDUCIBLE 30 (IAA30)* (Braybrook et al., 2006; Zheng et al., 2009). *IAA30* encodes a noncanonical Aux/IAA protein and both 2,4-D- and *AGL15*-induced SE is compromised in the *iaa30* mutant (Zheng et al., 2009). *LEC2* also activates *TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS 1 (TAA1)* and *YUCCA* genes, which encode key enzymes in the auxin biosynthesis pathway, resulting in an increase in endogenous auxin levels (Wojcikowska et al., 2013). As a result, LEC overexpression leads to a reduced exogenous auxin requirement in SE cultures (Wojcikowska et al., 2013). However, overexpression of *LEC2* in combination with a standard 2,4-D concentration is detrimental for somatic embryo production (Ledwon and Gaj, 2011). *LEC* expression is controlled by the action of PKL (see above) and by VP1/ABI3-LIKE (VAL) proteins. VAL proteins are B3 domain-containing transcription factors and mutations in *VAL* genes lead to an increased expression of *LEC* and ectopic embryo formation (Suzuki et al., 2007).

Ectopic expression of the AP2/ERF transcription factor BBM is also sufficient to induce SE on seedlings of different species without exogenous hormone application (Arabidopsis, brassica; (Boutilier et al., 2002), although in tobacco and sweet pepper exogenous cytokinin is required (Srinivasan et al., 2007; Heidmann et al., 2011) (Figure 2). BBM is a member of the *AINTEGUMENTA-LIKE (AIL)* clade of the AP2 subfamily of AP2/ERF genes that includes *AINTEGUMENTA (ANT)*, and other *AIL/PLETHORA (PLT)* genes (Horstman et al., 2014). *BBM* was initially identified as a marker for the induction of haploid embryo development from *B. napus* immature pollen grains (Boutilier et al., 2002). Arabidopsis *BBM* and the other *AIL/PLT* genes are expressed in the embryo and the root and shoot meristems, where they act redundantly to

define and maintain the stem cell niches (Aida et al., 2004; Galinha et al., 2007; Mudunkothge and Krizek, 2012). Despite their redundant functions, very few overlapping overexpression phenotypes have been reported. Only overexpression of *PLT5/AIL5* triggers somatic embryo and organ formation on *Arabidopsis* cotyledons (Tsuwamoto et al., 2010).

Even though *BBM* overexpression leads to very similar phenotypes as those described for the other SE-inducing genes described above, it is unclear how *BBM* induces SE and whether the *BBM* pathway and the other known SE pathways intersect. Microarray analysis after *BBM* activation identified direct *BBM* target genes, but did not reveal any clear links with other SE regulators (Passarinho et al., 2008). In addition to SE, *BBM* also induces other forms of regenerative growth, including callus, shoots, roots (tobacco), which has been exploited to improve regeneration after nuclear transformation in sweet pepper (Heidmann et al., 2011), white poplar (Deng et al., 2009) and chloroplast transformation in *Arabidopsis* (Lutz et al., 2011). At present, it is unclear how the organogenesis and embryogenesis pathways relate to each other.

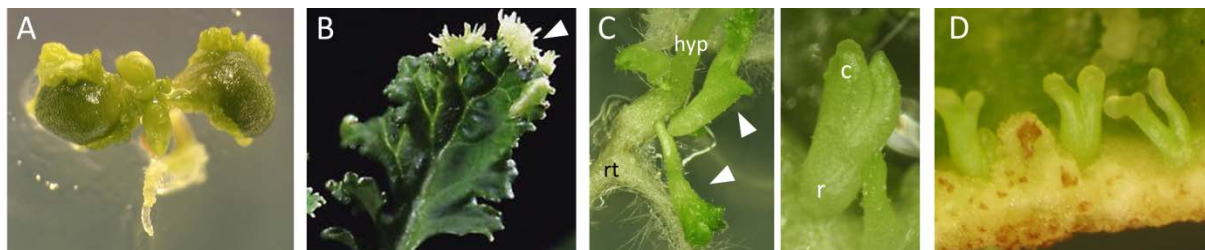


Figure 2. *BBM* overexpression induces somatic embryogenesis in multiple species.

(A) A *35S::BBM* *Arabidopsis* seedling with somatic embryos on its cotyledons and shoot apex (Boutilier et al., 2002).

(B) A *35S::BBM* *Brassica napus* plant with somatic embryos on the leaf margin (arrowhead) (Boutilier et al., 2002).

(C) Left: somatic embryos (arrowheads) developing at the transition zone of *35S::BBM-GR* tobacco seedlings grown on medium containing cytokinin and DEX (Srinivasan et al., 2007). The seedling hypocotyl (*hyp*) and root regions (*rt*) are indicated. Right: a bipolar somatic embryo on a *35S::BBM-GR* tobacco plant. The cotyledons (*c*) and radicle end (*r*) are indicated.

(D) Somatic embryo formation on a cotyledon of a *35S::BBM-GR* sweet pepper plant (Heidmann et al., 2011).

Outline of this thesis

The objective of the research presented in this thesis is to enhance our understanding of BBM-mediated SE and to determine whether the BBM signalling pathway intersects with those of other known SE regulators. Using *Arabidopsis* as a model system, I have studied several aspects of BBM function, including its interaction with HOMEODOMAIN GLABROUS transcription factors, its direct target genes, its functional relationship with other members of the AIL transcription factor family, and the cellular and molecular requirements for BBM (AIL)-mediated SE. The results in this thesis provide one of the first detailed molecular analyses of somatic embryo formation in plants and have allowed us to build up a model that integrates BBM with other genetic and physiological components of SE formation. The fundamental knowledge developed in this thesis can also be used to improve propagation/regeneration protocols in other species.

Chapter 1 introduces the concept of somatic embryogenesis and describes the different SE systems in *Arabidopsis* and the genetic components that play a role in this process.

Chapter 2 reviews the functions of AIL transcription factors during different aspects of plant development, including embryogenesis, stem cell niche specification, meristem maintenance and organ positioning and growth. We outline the gene regulatory networks in which AILs function and how these transcription factors integrate multiple hormonal inputs. Finally, we point out future challenges in AIL research.

Chapter 3 shows that BBM and other AIL transcription factors can interact with HOMEODOMAIN GLABROUS (HDG) transcription factors. We show that overexpression of one of these proteins, *HDG1* leads to root and shoot meristem termination and promotes endoreduplication, while down-regulation of multiple *HDG* genes leads to enhanced proliferation and SE phenotypes. We propose opposite functions for AIL and HDG transcription factors, stimulating and restraining cell proliferation, respectively, and build a model for interaction between BBM and HDG proteins that incorporates their interaction, developmental phenotypes and target genes.

Chapter 4 shows that AIL transcription factors have overlapping cell proliferation phenotypes; all AIL proteins except AIL1 and ANT are able to induce SE from *Arabidopsis* seedlings. Using BBM and PLT2 as representative AIL proteins, we show that their overexpression phenotypes are dosage and context dependent. Analysis of direct BBM targets and subsequent molecular and

genetic analyses link the BBM SE pathway to the *LEAFY COTYLEDON* genes, which are also known to play an important role in Arabidopsis SE.

Chapter 5 describes a genome-wide analysis of BBM DNA binding sites using chromatin immunoprecipitation followed by sequencing (ChIP-seq). Our ChIP-Seq and gene expression analysis reveals that BBM binds and positively regulates genes that are known components of Arabidopsis SE, including the *LEC* genes described in Chapter 4, auxin biosynthesis genes and recently discovered SE components, the *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED (AHL)* genes.

Chapter 6 provides an overview on how microarrays can be used to identify plant transcription factor target genes, describing issues such as construct design for controlled transcription factor activity, experimental setup, statistical analysis and confirmation of candidate target genes.

Chapter 7 is a cautionary note on the use of Bimolecular Fluorescence Complementation (BiFC), or split-YFP, in *in planta* protein-protein interaction studies, that was fuelled by problems I encountered in my research on BBM-HDG interactions. A literature survey revealed that most plant BiFC experiments are carried out in an inappropriate manner, with inappropriate controls and a qualitative rather than quantitative read-out of protein-protein interaction. We present a beginner's guideline for the setup of BiFC experiments, discussing each step of the protocol, including vector choice, plant expression systems, negative controls and signal detection.

Chapter 8 summarizes and discusses the most important results from this thesis and highlights directions for future research.

Funding

This work was funded by a Technology Top Institute - Green Genetics grant

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.
- Atta, R., Laurens, L., Boucheron-Dubuisson, E., Guivarc'h, A., Carnero, E., Giraudat-Pautot, V., Rech, P., and Chriqui, D.** (2009). Pluripotency of Arabidopsis xylem pericycle underlies shoot regeneration from root and hypocotyl explants grown in vitro. *Plant J* **57**, 626-644.
- Bai, B., Su, Y.H., Yuan, J., and Zhang, X.S.** (2013). Induction of Somatic Embryos in Arabidopsis Requires Local YUCCA Expression Mediated by the Down-Regulation of Ethylene Biosynthesis. *Molecular plant* **6**, 1247-1260.
- Bassuner, B.M., Lam, R., Lukowitz, W., and Yeung, E.C.** (2007). Auxin and root initiation in somatic embryos of Arabidopsis. *Plant cell reports* **26**, 1-11.
- Bicknell, R.A., and Koltunow, A.M.** (2004). Understanding apomixis: Recent advances and remaining conundrums. *Plant Cell* **16**, S228-S245.
- Boutillier, K., Offringa, R., Sharma, V.K., Kieft, 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. *The 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 genetics* **7**, e1002014.
- Bratzel, F., Lopez-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. *Current biology : CB* **20**, 1853-1859.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3468-3473.
- Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R., and Goodrich, J.** (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**, 5263-5276.
- Charriere, F., Sotta, B., Miginiac, E., and Hahne, G.** (1999). Induction of adventitious shoots or somatic embryos on in vitro cultured zygotic embryos of *Helianthus annuus*: Variation of endogenous hormone levels. *Plant Physiol Bioch* **37**, 751-757.
- Che, P., Lall, S., and Howell, S.H.** (2007). Developmental steps in acquiring competence for shoot development in Arabidopsis tissue culture. *Planta* **226**, 1183-1194.
- Chen, D., Molitor, A., Liu, C., and Shen, W.H.** (2010). The Arabidopsis PRC1-like ring-finger proteins are necessary for repression of embryonic traits during vegetative growth. *Cell Res* **20**, 1332-1344.
- Costa, L.M., Marshall, E., Tesfaye, M., Silverstein, K.A., Mori, M., Umetsu, Y., Otterbach, S.L., Papareddy, R., Dickinson, H.G., Boutillier, K., VandenBosch, K.A., Ohki, S., and Gutierrez-Marcos, J.F.** (2014). Central cell-derived peptides regulate early embryo patterning in flowering plants. *Science* **344**, 168-172.
- De Vries, S.C., Booij, H., Meyerink, P., Huisman, G., Wilde, H.D., Thomas, T.L., and Vankammen, A.** (1988). Acquisition of Embryogenic Potential in Carrot Cell-Suspension Cultures. *Planta* **176**, 196-204.

- Dean Rider, S., Jr., 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.
- Delbarre, A., Muller, P., Imhoff, V., and Guern, J.** (1996). Comparison of mechanisms controlling uptake and accumulation of 2,4-dichlorophenoxy acetic acid, naphthalene-1-acetic acid, and indole-3-acetic acid in suspension-cultured tobacco cells. *Planta* **198**, 532-541.
- Deng, W., Luo, K.M., Li, Z.G., and Yang, Y.W.** (2009). A novel method for induction of plant regeneration via somatic embryogenesis. *Plant Science* **177**, 43-48.
- Dubois, T., Guedira, M., Dubois, J., and Vasseur, J.** (1990). Direct Somatic Embryogenesis in Roots of *Cichorium*: Is Callose an Early Marker? *Annals of Botany* **65**, 539-545.
- Elhiti, M., Tahir, M., Gulden, R.H., Khamiss, K., and Stasolla, C.** (2010). Modulation of embryo-forming capacity in culture through the expression of Brassica genes involved in the regulation of the shoot apical meristem. *J Exp Bot* **61**, 4069-4085.
- Elhiti, M., Hebelstrup, K.H., Wang, A., Li, C., Cui, Y., Hill, R.D., and Stasolla, C.** (2013). Function of type-2 *Arabidopsis* hemoglobin in the auxin-mediated formation of embryogenic cells during morphogenesis. *Plant J* **74**, 946-958.
- Fenning, T., and Park, Y.-S.** (2014). Conifer Somatic Embryogenesis and Multi-Varietal Forestry. In *Challenges and Opportunities for the World's Forests in the 21st Century* (Springer Netherlands), pp. 425-439.
- Gaj, M.D.** (2001). Direct somatic embryogenesis as a rapid and efficient system for in vitro regeneration of *Arabidopsis thaliana*. *Plant Cell Tiss Org* **64**, 39-46.
- Gaj, M.D.** (2004). Factors influencing somatic embryogenesis induction and plant regeneration with particular reference to *Arabidopsis thaliana* (L.) Heynh. *Plant Growth Regul* **43**, 27-47.
- Gaj, M.D.** (2011). Somatic embryogenesis and plant regeneration in the culture of *Arabidopsis thaliana* (L.) Heynh. immature zygotic embryos. *Methods in molecular biology* **710**, 257-265.
- Gaj, M.D., Zhang, S.B., Harada, J.J., and Lemaux, P.G.** (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of *Arabidopsis*. *Planta* **222**, 977-988.
- Gaj, M.D., Trojanowska, A., Ujczak, A., Medrek, M., Koziol, A., and Garbaciak, B.** (2006). Hormone-response mutants of *Arabidopsis thaliana* (L.) Heynh. impaired in somatic embryogenesis. *Plant Growth Regul* **49**, 183-197.
- 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.
- Gallois, J.L., Nora, F.R., Mizukami, Y., and Sablowski, R.** (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Gene Dev* **18**, 375-380.
- Garces, H.M.P., Champagne, C.E.M., Townsley, B.T., Park, S., Malho, R., Pedroso, M.C., Harada, J.J., and Sinha, N.R.** (2007). Evolution of asexual reproduction in leaves of the genus *Kalanchoe*. *P Natl Acad Sci USA* **104**, 15578-15583.
- Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M., and McCourt, P.** (2004). The transcription factor FUSCA3 controls developmental timing in *Arabidopsis* through the hormones gibberellin and abscisic acid. *Dev Cell* **7**, 373-385.
- Gutmann, M., vonAderkas, P., Label, P., and Lelu, M.A.** (1996). Effects of abscisic acid on somatic embryo maturation of hybrid larch. *J Exp Bot* **47**, 1905-1917.
- Halperin, W., and Wetherell, D.F.** (1964). Adventive Embryony in Tissue Cultures of Wild Carrot *Daucus Carota*. *Am J Bot* **51**, 274-&.

- 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 physiology* **133**, 653-663.
- Hecht, V., Vielle-Calzada, J.P., Hartog, M.V., Schmidt, E.D., 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 physiology* **127**, 803-816.
- Heidmann, I., de Lange, B., Lambalk, J., Angenent, G.C., and Boutilier, K.** (2011). Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. *Plant cell reports* **30**, 1107-1115.
- Henderson, J.T., Li, H.C., Rider, S.D., Mordhorst, A.P., Romero-Severson, J., Cheng, J.C., Robey, J., Sung, Z.R., de Vries, S.C., and Ogas, J.** (2004). PICKLE acts throughout the plant to repress expression of embryonic traits and may play a role in gibberellin-dependent responses. *Plant physiology* **134**, 995-1005.
- Hennig, L., and Derkacheva, M.** (2009). Diversity of Polycomb group complexes in plants: same rules, different players? *Trends Genet* **25**, 414-423.
- 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.
- Hosek, P., Kubes, M., Lankova, M., Dobrev, P.I., Klima, P., Kohoutova, M., Petrasek, J., Hoyerova, K., Jirina, M., and Zazimalova, E.** (2012). Auxin transport at cellular level: new insights supported by mathematical modelling. *J Exp Bot* **63**, 3815-3827.
- Ikeda-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.
- Jia, H., McCarty, D.R., and Suzuki, M.** (2013). Distinct roles of LAFL network genes in promoting the embryonic seedling fate in the absence of VAL repression. *Plant physiology* **163**, 1293-1305.
- Karlova, R., Boeren, S., Russinova, E., Aker, J., Vervoort, J., and de Vries, S.** (2006). The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE1 protein complex includes BRASSINOSTEROID-INSENSITIVE1. *Plant Cell* **18**, 626-638.
- Keith, K., Kraml, M., Dengler, N.G., and McCourt, P.** (1994). fusca3: A Heterochronic Mutation Affecting Late Embryo Development in Arabidopsis. *Plant Cell* **6**, 589-600.
- Kobayashi, T., Nagayama, Y., Higashi, K., and Kobayashi, M.** (2010). Establishment of a tissue culture system for somatic embryogenesis from germinating embryos of Arabidopsis thaliana. *Plant Biotechnol* **27**, 359-364.
- Kurczynska, E.U., Gaj, M.D., Ujczak, A., and Mazur, E.** (2007). Histological analysis of direct somatic embryogenesis in Arabidopsis thaliana (L.) Heynh. *Planta* **226**, 619-628.
- 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.
- Ledwon, A., and Gaj, M.D.** (2011). LEAFY COTYLEDON1, FUSCA3 expression and auxin treatment in relation to somatic embryogenesis induction in Arabidopsis. *Plant Growth Regul* **65**, 157-167.
- Lotan, T., Ohto, M., Yee, K.M., West, M.A.L., 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.
- 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.
- Lutz, K.A., Azhagiri, A., and Maliga, P.** (2011). Transplastomics in Arabidopsis: progress toward developing an efficient method. *Methods Mol Biol* **774**, 133-147.

- Meinke, D.W., Franzmann, L.H., Nickle, T.C., and Yeung, E.C.** (1994). Leafy Cotyledon Mutants of Arabidopsis. *Plant Cell* **6**, 1049-1064.
- Michalczuk, L., Ribnicky, D.M., Cooke, T.J., and Cohen, J.D.** (1992). Regulation of Indole-3-Acetic-Acid Biosynthetic Pathways in Carrot Cell-Cultures. *Plant physiology* **100**, 1346-1353.
- Miguel, C., and Marum, L.** (2011). An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond. *J Exp Bot* **62**, 3713-3725.
- Mordhorst, A.P., Toonen, M.A.J., and deVries, S.C.** (1997). Plant embryogenesis. *Crit Rev Plant Sci* **16**, 535-576.
- Mordhorst, A.P., Voerman, K.J., Hartog, M.V., Meijer, E.A., van Went, J., Koornneef, M., and de Vries, S.C.** (1998). Somatic embryogenesis in Arabidopsis thaliana is facilitated by mutations in genes repressing meristematic cell divisions. *Genetics* **149**, 549-563.
- Morris, D.A., Friml, J., and Zařímalová, E.** (2004). The transport of auxins. (Kluwer Academic Publishers).
- Mudunkothge, J.S., and Krizek, B.A.** (2012). Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function. *Plant J* **71**, 108-121.
- Ogas, J., Cheng, J.C., Sung, Z.R., and Somerville, C.** (1997). Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. *Science* **277**, 91-94.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13839-13844.
- Parcy, F., Valon, C., Kohara, A., Misera, S., and Giraudat, J.** (1997). The ABSCISIC ACID-INSENSITIVE3, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of Arabidopsis seed development. *Plant Cell* **9**, 1265-1277.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M.W., 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 molecular biology* **68**, 225-237.
- Pasternak, T.P., Prinsen, E., Ayaydin, F., Miskolczi, P., Potters, G., Asard, H., Van Onckelen, H.A., Dudits, D., and Feher, A.** (2002). The role of auxin, pH, and stress in the activation of embryogenic cell division in leaf protoplast-derived cells of alfalfa. *Plant physiology* **129**, 1807-1819.
- Pillon, E., Terzi, M., Baldan, B., Mariani, P., and Lo Schiavo, F.L.** (1996). A protocol for obtaining embryogenic cell lines from Arabidopsis. *Plant J* **9**, 573-577.
- Raghavan, V.** (2004). Role of 2,4-dichlorophenoxyacetic acid (2,4-D) in somatic embryogenesis on cultured zygotic embryos of Arabidopsis: Cell expansion, cell cycling, and morphogenesis during continuous exposure of embryos to 2,4-D. *Am J Bot* **91**, 1743-1756.
- Reinert, J.** (1958). Morphogenese Und Ihre Kontrolle an Gewebekulturen Aus Carotten. *Naturwissenschaften* **45**, 344-345.
- Schubert, D., Clarenz, O., and Goodrich, J.** (2005). Epigenetic control of plant development by Polycomb-group proteins. *Current opinion in plant biology* **8**, 553-561.
- Schuettengruber, B., Martinez, A.M., Iovino, N., and Cavalli, G.** (2011). Trithorax group proteins: switching genes on and keeping them active. *Nature reviews. Molecular cell biology* **12**, 799-814.
- Sharma, S., Shahzad, A., and da Silva, J.A.T.** (2013). Synseed technology-A complete synthesis. *Biotechnol Adv* **31**, 186-207.

- Srinivasan, C., Liu, Z., Heidmann, I., Supena, E.D., Fukuoka, H., Joosen, R., Lambalk, J., Angenent, G., Scorza, R., Custers, J.B., and Boutilier, K.** (2007). Heterologous expression of the BABY BOOM AP2/ERF transcription factor enhances the regeneration capacity of tobacco (*Nicotiana tabacum* L.). *Planta* **225**, 341-351.
- Steward, F.C., Mapes, M.O., and Mears, K.** (1958). Growth and Organized Development of Cultured Cells .2. Organization in Cultures Grown from Freely Suspended Cells. *Am J Bot* **45**, 705-708.
- Sticklen, M.B.** (1991). Direct Somatic Embryogenesis and fertile Plants from Rice Root Cultures. *Journal of plant physiology* **138**, 577-580.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11806-11811.
- Su, Y.H., Zhao, X.Y., Liu, Y.B., Zhang, C.L., O'Neill, S.D., and Zhang, X.S.** (2009). Auxin-induced WUS expression is essential for embryonic stem cell renewal during somatic embryogenesis in *Arabidopsis*. *Plant J* **59**, 448-460.
- Sugimoto, K., Jiao, Y.L., and Meyerowitz, E.M.** (2010). *Arabidopsis* Regeneration from Multiple Tissues Occurs via a Root Development Pathway. *Dev Cell* **18**, 463-471.
- Suzuki, M., Wang, H.H., and McCarty, D.R.** (2007). Repression of the LEAFY COTYLEDON 1/B3 regulatory network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. *Plant physiology* **143**, 902-911.
- Tanaka, M., Kikuchi, A., and Kamada, H.** (2008). The *Arabidopsis* histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant physiology* **146**, 149-161.
- Tian, L.N., and Brown, D.C.W.** (2000). Improvement of soybean somatic embryo development and maturation by abscisic acid treatment. *Can J Plant Sci* **80**, 271-276.
- 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 molecular biology* **73**, 481-492.
- Vahdati, K., Bayat, S., Ebrahimzadeh, H., Jariteh, M., and Mirmasoumi, M.** (2008). Effect of exogenous ABA on somatic embryo maturation and germination in Persian walnut (*Juglans regia* L.). *Plant Cell Tiss Org* **93**, 163-171.
- Van Hengel, A.J., Guzzo, F., Van Kammen, A., and De Vries, S.C.** (1998). Expression pattern of the carrot EP3 endochitinase genes in suspension cultures and in developing seeds. *Plant physiology* **117**, 43-53.
- Wang, H., Tang, W., Zhu, C., and Perry, S.E.** (2002). A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. *Plant J* **32**, 831-843.
- 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, X.C., Niu, Q.W., Teng, C., Li, C., Mu, J.Y., Chua, N.H., and Zuo, J.R.** (2009). Overexpression of PGA37/MYB118 and MYB115 promotes vegetative-to-embryonic transition in *Arabidopsis*. *Cell Res* **19**, 224-235.
- 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 physiology* **133**, 1882-1892.

- 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.
- Whitehouse, I., Flaus, A., Cairns, B.R., White, M.F., Workman, J.L., and Owen-Hughes, T.** (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* **400**, 784-787.
- Wiweger, M., Farbos, I., Ingouff, M., Lagercrantz, U., and von Arnold, S.** (2003). Expression of Chia4-Pa chitinase genes during somatic and zygotic embryo development in Norway spruce (*Picea abies*): similarities and differences between gymnosperm and angiosperm class IV chitinases. *J Exp Bot* **54**, 2691-2699.
- Wojcikowska, B., Jaskola, K., Gasiorek, P., Meus, M., Nowak, K., and Gaj, M.D.** (2013). LEAFY COTYLEDON2 (LEC2) promotes embryogenic induction in somatic tissues of Arabidopsis, via YUCCA-mediated auxin biosynthesis. *Planta* **238**, 425-440.
- Yang, J.L., Seong, E.S., Kim, M.J., Ghimire, B.K., Kang, W.H., Yu, C., and Li, C.** (2010). Direct somatic embryogenesis from pericycle cells of broccoli (*Brassica oleracea* L. var. *italica*) root explants. *Plant Cell, Tissue and Organ Culture (PCTOC)* **100**, 49-58.
- Yang, S., Johnston, N., Talideh, E., Mitchell, S., Jeffree, C., Goodrich, J., and Ingram, G.** (2008). The endosperm-specific ZHOUP1 gene of Arabidopsis thaliana regulates endosperm breakdown and embryonic epidermal development. *Development* **135**, 3501-3509.
- Zhang, H., Bishop, B., Ringenberg, W., Muir, W.M., and Ogas, J.** (2012). The CHD3 remodeler PICKLE associates with genes enriched for trimethylation of histone H3 lysine 27. *Plant physiology* **159**, 418-432.
- 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. *The Plant cell* **21**, 2563-2577.
- Zhou, Y., Tan, B., Luo, M., Li, Y., Liu, C., Chen, C., Yu, C.W., Yang, S.G., Dong, S., Ruan, J.X., Yuan, L.B., Zhang, Z., Zhao, L.M., Li, C.L., Chen, H.H., Cui, Y.H., Wu, K.Q., and Huang, S.Z.** (2013). HISTONE DEACETYLASE19 Interacts with HSL1 and Participates in the Repression of Seed Maturation Genes in Arabidopsis Seedlings. *Plant Cell* **25**, 134-148.
- 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

AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks

Anneke Horstman¹, Viola Willemsen², Kim Boutilier¹ and Renze Heidstra²

¹Plant Research International and ²Plant Developmental Biology, Wageningen University and Research centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.

Abstract

Members of the AINTEGUMENTA-LIKE (AIL) family of AP2/ERF domain transcription factors are expressed in all dividing tissues in the plant, where they play central roles in developmental processes such as embryogenesis, stem cell niche specification, meristem maintenance, organ positioning and growth. When overexpressed, AIL proteins induce adventitious growth, including somatic embryogenesis and ectopic organ formation. The *Arabidopsis* (*Arabidopsis thaliana*) genome contains eight *AIL* genes, including *AINTEGUMENTA*, *BABY BOOM* and the *PLETHORA* genes. Studies on these transcription factors have revealed their intricate relationship with auxin, as well as their involvement in an increasing number of gene regulatory networks, in which extensive cross-talk and feedback loops play a major role.

The AIL transcription factor family in Arabidopsis

The eight *AINTEGUMENTA-LIKE* (*AIL*) transcription factor genes within the Arabidopsis genome (Nole-Wilson et al., 2005) include *AINTEGUMENTA* (*ANT*; (Elliott et al., 1996; Klucher et al., 1996), *BABY BOOM* (*BBM*; (Boutillier et al., 2002)) and the *PLETHORA* (*PLT*) genes (Aida et al., 2004; Galinha et al., 2007) (Figure 1; Box 1), which are all expressed in young/dividing tissues in the plant. They play overlapping roles in the establishment and maintenance of meristems, as well as organ initiation and growth (Table 1). A wealth of genetic studies have shown that AIL proteins are master regulators of these developmental processes. Loss-of-function combinations and gain-of-function mutants of this gene family show spectacular phenotypes in which meristems or complete organs are missing or arise at ectopic positions. The central role of this family in meristem and organ development extends as far back as mosses (Karlberg et al., 2011; Aoyama et al., 2012; Rigal et al., 2012). In this review, we provide an overview of recent AIL research in the model plant Arabidopsis and point out future challenges in AIL research.

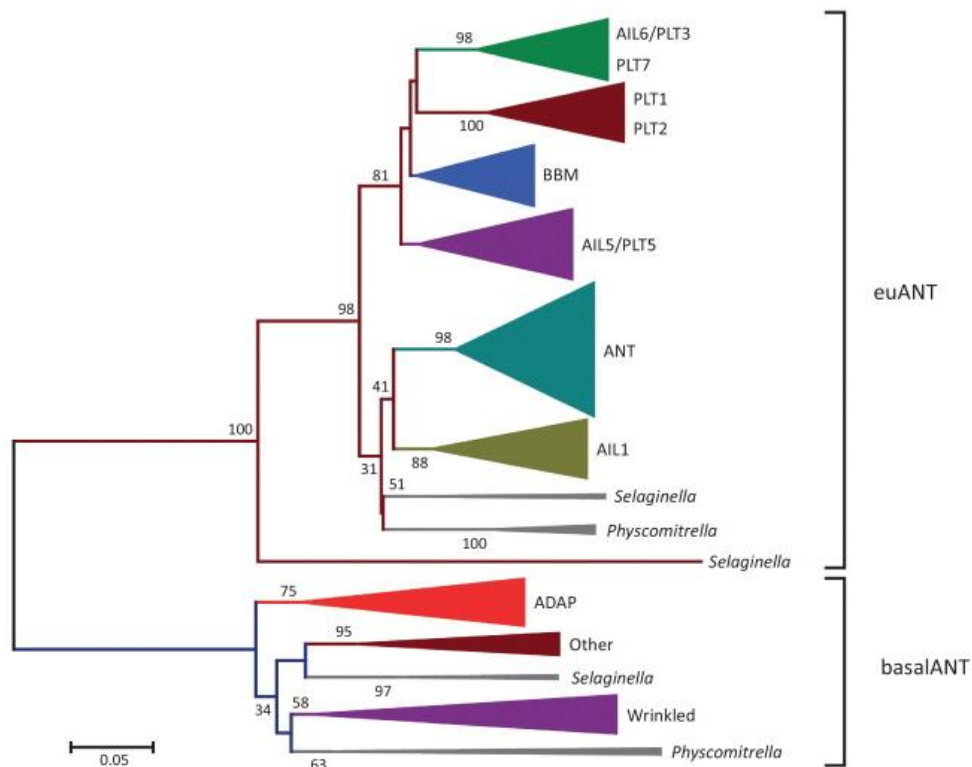


Figure 1. Phylogenetic tree for the AIL family of transcription factors.

Neighbor-joining tree of the euANT and basalANT family using MEGA version 5 (Tamura et al., 2011). A total of 251 sequences were analysed using the AP2/ERF domain region (240 positions). Numbers at the nodes indicate bootstrap support calculated using 100 replicates. Only bootstrap values over 30% are indicated. The “other” basalANT clade corresponds to genes with no apparent ortholog in Arabidopsis. *Selaginella* refers to *Selaginella moellendorffii* and *Physcomitrella* refers to *Physcomitrella patens*.

Box 1. AILs form a subgroup within the large AP2/ERF family

The AIL proteins are members of the AP2/ERF domain family of transcription factors that are found in a phylogenetically-wide group of plants including moss, algae, gymnosperms and angiosperms, and comprise the second largest group of transcription factors in plants, with up to 200 members in a single genome (Riechmann and Meyerowitz, 1998; Nole-Wilson et al., 2005; Kim et al., 2006; Shigyo et al., 2006; Zhuang et al., 2009; Dietz et al., 2010; Licausi et al., 2010; Sharoni et al., 2011; Rashid et al., 2012; Duan et al., 2013). The AP2/ERF domain, a DNA binding domain of approximately 60- to 70-amino acids (Okamuro et al., 1997), was identified initially in the *Arabidopsis* APETALA2 (AP2) protein (Jofuku et al., 1994), and shortly thereafter, in four tobacco (*Nicotiana tabacum*) ETHYLENE RESPONSE FACTORS (ERFs) (Ohme-Takagi and Shinshi, 1995). The AP2/ERF domain was thought to be plant-specific, until the discovery that the same domain exists in HNH homing endonucleases from viruses, cyanobacteria, ciliates and parasitic protists (Magnani et al., 2004; Wuitschick et al., 2004; Balaji et al., 2005). Homing endonucleases are invasive DNA sequences that are usually found in self-splicing introns or inteins, which can transpose and duplicate themselves within and between species (Taylor and Stoddard, 2012).

The AINTEGUMENTA-LIKE (AIL) proteins fall under the AP2 subfamily of proteins, which contain two AP2/ERF domains separated by a linker region (Riechmann and Meyerowitz, 1998). This AP2 subfamily is divided into the euAP2 lineage, which is characterized by a *mir172* binding motif that lies C-terminal to the AP2/ERF domains, and the ANT lineage, which contains signature amino acid insertions in the first and second AP2/ERF domain. The ANT lineage is further divided into the basalANT and euANT/AIL lineages (Figure 1) based on a number of conserved amino acid insertions in the first AP2/ERF domain (euANT1) and the N-terminal region (euANT2-4) of the euANT/AIL proteins (Kim et al., 2006; El Ouakfaoui et al., 2010). These motifs are generally well conserved among AILs, both within and between species (Kim et al., 2006; El Ouakfaoui et al., 2010; Bandupriya et al., 2013).

The three-dimensional (3-D) structure of AP2/ERF proteins has only been resolved for the single AP2 domain-containing protein ERF1. The 3D solution structure of ERF1 showed that its AP2 domain forms a three-stranded anti-parallel β -sheet that lies mostly parallel to a conserved α -helix (Allen et al., 1998). Modeling and DNA binding experiments suggest that the anti-parallel β -sheet is responsible for the DNA binding properties of ERF1 (Allen et al., 1998), which has since been confirmed for other single AP2 domain proteins in the large ERF1 subfamily (Cao et al., 2001; Shoji et al., 2013). Homology modeling of *Arabidopsis* ANT using the ERF1 3-D structure as a template suggests that each AP2 repeat of ANT forms an α -helix, similar to ERF1. However unlike ERF1, the first AP2 repeat of ANT is predicted to contain two β -sheets, at different positions than those in ERF1, while the second AP2/ERF repeat does not appear to form β -sheets (Krizek, 2003). Both AP2 domains of ANT are required for DNA binding and each domain is thought to use different amino acids to contact the DNA (Nole-Wilson and Krizek, 2000; Krizek, 2003). It was proposed that the first AP2 repeat binds to the 5' part of the target sequence and the second AP2 repeat to the 3' part of the target sequence, with the ANT linker region serving as a bridge (Nole-Wilson and Krizek, 2000). The importance of the linker region between the AP2/ERF domains is illustrated by the high conservation of this region within the *AIL* gene family and by the

observation that mutations in the linker abolish DNA binding *in vitro*. These observations lead to the suggestion that the linker region may directly bind DNA or serve to position the AP2 repeats on the DNA (Krizek, 2003).

AP2/ERF domain proteins regulate two major processes in plants: response to stress and control of growth and development. Notably, AIL proteins have been shown to function exclusively in pathways related to development.

Table 1. Arabidopsis AIL genes and their functions^a

AGI number	Gene Names	Function	Refs.
At4g37750	<i>ANT</i>	Shoot and flower meristem maintenance, organ size and polarity, flower initiation, ovule development, floral organ identity, cell proliferation	[2, 3, 51, 78, 83, 84, 92, 93, 97, 103, 104]
At1g72570	<i>AIL1</i>		
At5g17430	<i>AIL2/BBM/PLT4</i>	Embryogenesis, root SCN ^b patterning and meristem maintenance, cell proliferation	[4, 6]
At3g20840	<i>AIL3/PLT1</i>	Embryogenesis, root SCN patterning and meristem maintenance, cell proliferation	[5, 6]
At1g51190	<i>AIL4/PLT2</i>	Embryogenesis, root SCN patterning and meristem maintenance, cell proliferation	[5, 6]
At5g57390	<i>AIL5/CHO1/EMK/PLT5</i>	germination, phyllotaxy, rhizotaxy, cell proliferation, seed maturation	[10, 13, 50, 57, 59, 79, 85]
At5g10510	<i>AIL6/PLT3</i>	Shoot and flower meristem maintenance, organ size, flower initiation, floral organ identity, embryogenesis, root SCN patterning and meristem maintenance, phyllotaxy, rhizotaxy, cell proliferation	[6, 50, 51, 78, 79, 85, 93, 95]
At5g65510	<i>AIL7/PLT7</i>	Shoot meristem maintenance, phyllotaxy, rhizotaxy, cell proliferation	[50, 78, 79, 85]

^aIndividual *AIL* members are given different names. We refer to the genes by their most commonly used name (in bold).

^bSCN, stem cell niche

AIL function during embryogenesis

PLT1, *PLT2*, *AIL6/PLT3* and *BBM* (collectively called *PLT/BBM*) genes play a major role in basal patterning of the embryo (Figure 2A). *PLT1* and *PLT2* gene expression has been described from the octant stage onward, in the lower tier of the embryo proper (Aida et al., 2004). Early embryonic *AIL6/PLT3* and *BBM* expression has not been reported. Later in embryogenesis *PLT1* expression becomes restricted to the quiescent centre (QC) and surrounding stem cells, while the *PLT2*, *AIL6/PLT3* and *BBM* expression domains are slightly expanded to include the ground tissue and provascular cells. Post-embryonically, the *AIL6/PLT3* expression maximum is in the columella stem cells, in contrast to the QC peak expression observed for *PLT1*, *PLT2* and *BBM* (Aida et al., 2004; Galinha et al., 2007). Only combinations of *plt1*, *plt2*, *ail6/plt3* and *bbm*

mutants show embryonic abnormalities (Aida et al., 2004; Galinha et al., 2007). Subtle defects in *plt1;plt2* embryos include enlarged mis-specified QC progenitor cells, and *plt1;plt2* seedlings show defective root development, confirming that *PLT1* and *PLT2* are required for stem cell niche specification (Aida et al., 2004). *plt1;plt2;ail6/plt3* triple mutant embryos show aberrant organization of the embryonic root pole and seedlings are rootless. *plt2;bbm* double mutants fail to develop past the early embryo stage, indicating the importance of *PLT2* and *BBM* for embryogenesis (Galinha et al., 2007).

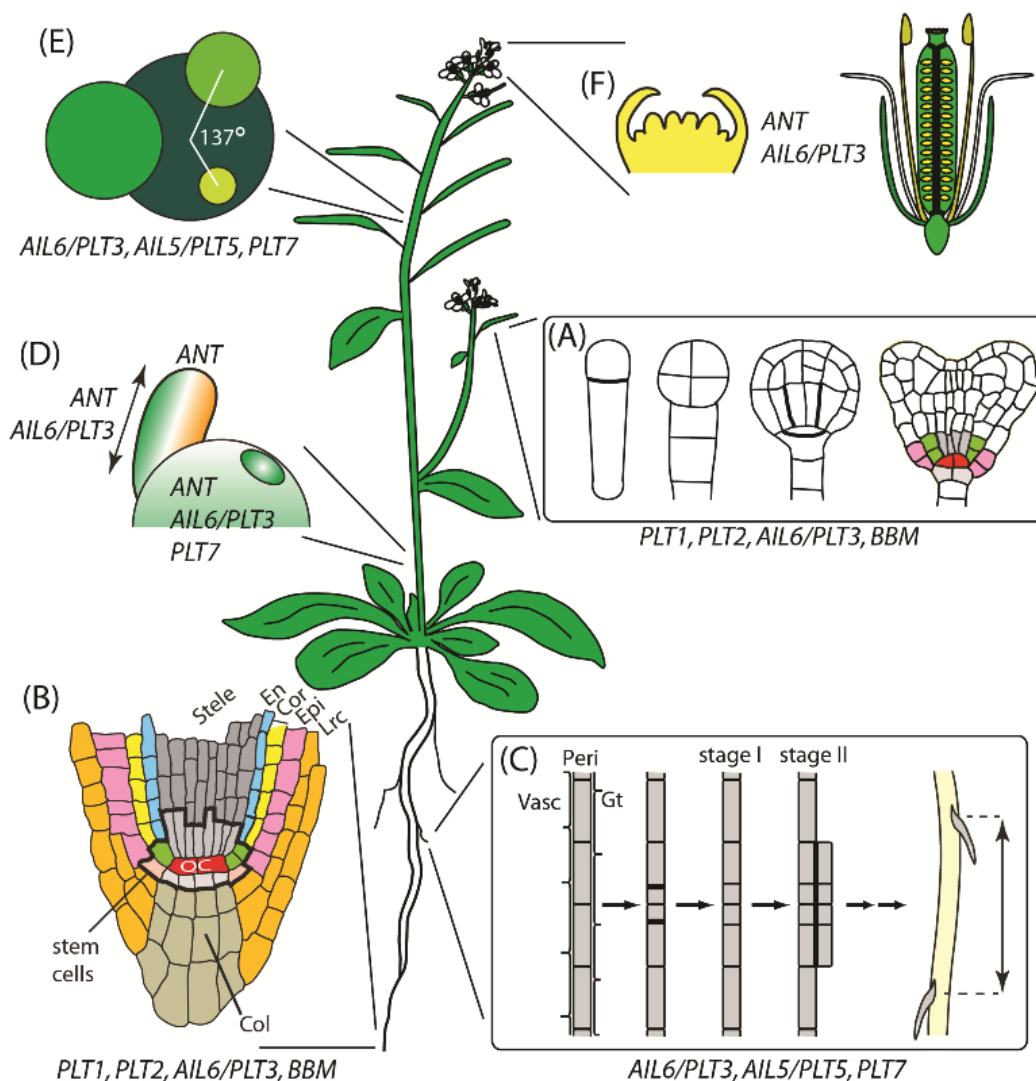


Figure 2. Sites of AIL function during plant development. (A) Embryogenesis. During embryo development, the apical-basal body axis is established by an asymmetric division of the zygote. The apical cell gives rise to most of the embryo, including the cotyledons, shoot apical meristem, hypocotyl and root stem cells. The basal cell forms the hypophyseal cell, which gives rise to the columella and the quiescent center (QC), which is specified at early heart stage (Jürgens and Mayer, 1994; Scheres et al., 1994). Together the QC and surrounding stem cells form the stem cell niche. (B) Root development. The root meristem provides new cells for tissues of the growing root: stele, ground tissue, epidermis, lateral root cap and columella. The QC is essential for maintenance of the surrounding stem cells (van den Berg et al., 1997). Daughter cells produced by the stem

cells traverse the meristematic zone, where they divide to generate a pool of cells that exit the cell cycle when they reach the transition zone, and increase in length in the elongation zone before acquiring their specific characteristics in the differentiation zone. **(C)** Rhizotaxis. In *Arabidopsis*, lateral roots arise from two files of pericycle cells that lie adjacent to the protoxylem within the differentiating root (Dubrovsky et al., 2006; Parizot et al., 2008). A subset of these cells, called founder cells, is stimulated by a local accumulation of auxin to divide and form a lateral root primordium (LRP; (Benková et al., 2003; De Smet et al., 2007; Laskowski et al., 2008; Moreno-Risueno et al., 2010)). LRP spacing correlates with the root curvature (Dubrovsky et al., 2006; Laskowski et al., 2008). **(D-F)** Shoot development. The shoot apical meristem (SAM) gives rise to the leaves. The transition to reproductive growth is marked by the conversion of the vegetative meristem to an inflorescence meristem, which produces floral meristems that differentiate into flowers containing different types of organs with distinct sizes and shapes. Organ size is determined by cell number and cell size. **(D)** Organ polarity. Leaves and floral organs are polarized along their adaxial-abaxial axis and different cell types exist on either side of this axis. The main players controlling organ polarity in *Arabidopsis* correspond to three classes of transcription factors: the adaxial-specifying HD-ZIP III proteins, and the abaxial-specifying KANADI (KAN) and YABBY (YAB) proteins (Szakonyi et al., 2010). The arrow in D represents the role of *ANT* and *AIL6/PLT3* in lateral organ growth. **(E)** Phyllotaxis. In *Arabidopsis* seedlings, the cotyledons and the first pair of leaves are formed in a decussate pattern (opposite to each other), while subsequent leaves develop in a spiral pattern with an angle close to the “golden angle” of 137.5°. This spiral pattern of organ initiation is maintained in the inflorescence meristem (reviewed in (Kuhlemeier, 2007)). **(F)** Floral organ identity and ovule development. Floral organs are specified by the combined activity of the so-called A, B, C and E classes of organ identity genes, which are expressed in overlapping domains. In addition, class A and C activities inhibit each other (reviewed in (Galbiati et al., 2013b)). During gynoecium development, two carpel margin meristems (CMM) form on the adaxial (inner) portion of the medial domain of the gynoecium, and ultimately give rise to the ovules and other organs of the carpel.

BBM or *AIL5/PLT5* overexpression induces the ectopic formation of embryos on the meristem, cotyledons and first leaves of seedlings (Boutillier et al., 2002; Tsuwamoto et al., 2010). Overexpression of *PLT1* or *PLT2* during embryogenesis ectopically induces root stem cell niches and, in the most extreme cases, can lead to a complete transformation of the embryo toward root identity (Aida et al., 2004). In line with this, induced overexpression of *PLT2* in seedlings can also produce roots from the shoot apex (Galinha et al., 2007). Together with the mutant phenotypes, these results suggest that *BBM*, *AIL5/PLT5*, *PLT1* and *PLT2* genes act as master regulators for early embryo and root development. This may reflect a specific function for these genes in embryo initiation or the maintenance of cell potency.

In most plant species, the initial phase of embryo cell proliferation and morphogenesis is followed by the maturation phase in which cell division stops, storage reserves accumulate and the seed becomes desiccation tolerant and dormant. In *Arabidopsis*, the onset of seed maturation is characterized by an increase in the level of the plant hormone abscisic acid (ABA) and by increased expression of an interwoven network of transcription factors, among which ABSCISIC ACID-INSENSITIVE 3 (*ABI3*), that together regulate expression of maturation phase genes. (reviewed in (Braybrook and Harada, 2008; Jia et al., 2013)). Recently, a heterologous system based on activation of the *Phaseolus vulgaris* *ABI3* transcription factor PvALF, and expression of its target gene *PHASEOLIN (PHAS)*, identified *AIL5/PLT5* as being co-expressed with

PHAS (Sundaram et al., 2013). *AIL5/PLT5* is bound by PvALF in the presence of ABA, and required for expression of endogenous seed storage genes in Arabidopsis (Sundaram et al., 2013). It is not known whether *AIL5/PLT5* directly binds to seed storage genes to activate their expression or if it acts upstream of other maturation phase transcription factors, thereby regulating the final phase of embryo development.

AIL-auxin feedback loop

The phytohormone auxin has been shown to play an important role in the formation of the apical–basal axis of the embryo. Auxin binding by its receptors TRANSPORT INHIBITOR RESISTANT1 (TIR1) and TIR1-related proteins promotes degradation of the AUXIN/INDOLE ACETIC ACID (Aux/IAA) family proteins that bind to AUXIN RESPONSE FACTOR (ARF) family transcription factors, inhibiting ARF transcriptional activity on auxin target genes (reviewed in (Mockaitis and Estelle, 2008)).

The expression of *PLT1* and *PLT2* is dependent on the redundant action of ARF5/MONOPTEROS (MP) and that of its close homolog NONPHOTOTROPIC HYPOCOTYL 4 (NPH4/ARF7; (Harper et al., 2000; Aida et al., 2004)). Although the dynamics of auxin-induced *PLT1* and *PLT2* transcription implies that they are (late) auxin response genes, exogenous application of auxin fails to rescue *plt1;plt2* mutants, indicating that auxin cannot bypass the requirement for *PLT1* and *PLT2* (Aida et al., 2004). Expression of the auxin efflux facilitator *PINFORMED4* (*PIN4*) is largely absent in *plt1;plt2* embryos, showing that *PIN4* is downstream of *PLT1* and *PLT2* in the transcriptional network. Reduced transcription of *PIN1* and *PIN3* is also observed in *plt1;plt2;ail6;plt3* triple mutant embryos (Blilou et al., 2005; Galinha et al., 2007). By contrast, the *PLT1* expression domain is expanded to the whole embryo in *pin2;pin3;pin4;pin7* mutants. Explanted *pin2;pin3;pin4;pin7* embryos show impaired cotyledon development, and root hairs emerge at apical positions on the seedling. These findings suggest a feedback loop where PIN proteins, by directing auxin transport and accumulation, restrict *PLT1* and *PLT2* expression to the basal embryo domain to initiate embryonic root specification. In turn, *PLT* activity regulates *PIN* transcription to stabilize the position of the root primordium (Blilou et al., 2005) (Figure 3).

Defects in QC patterning are observed in *RopGEF7* RNAi mutants during embryogenesis that correlate with the reduced expression of *PIN1*, *PLT1* and *PLT2* (Won et al., 2011). *RopGEF7*, which encodes a RAC/ROP GTPase activator, is expressed in the same domain as *PLT1* and *PLT2* and its expression is unaffected in *plt1;plt2* double mutants (Won et al., 2011). It has been suggested that *RopGEF7*, via RAC/ROP GTPases (Hazak and Yalovsky, 2010) regulates correct PIN

endocytosis, thereby affecting local auxin concentrations, and indirectly *PLT1* and *PLT2* expression to mediate root meristem patterning (Won et al., 2011).

Reduced expression of *PLT1*, *PLT2*, *BBM* and several *PIN* genes is observed in *JAGGED LATERAL ORGAN (JLO)* mutant embryos (Bureau et al., 2010). Strong *jlo* mutants are embryo lethal, but plants carrying the hypomorphic *jlo-2* allele produce viable embryos in addition to those resembling *mp* and *bdl* mutant embryos (Borghgi et al., 2007; Bureau et al., 2010). *JLO* may control embryo patterning, either alone or together with its interaction partner *ASYMMETRIC LEAVES 2 (AS2)* (Rast and Simon, 2012), through the auxin-dependent *MP-BDL* pathway or by directly regulating *PLT/BBM* gene expression (Figure 3).

Together these studies highlight the importance of AIL function and its relation to auxin for embryo development.

Embryonic AIL regulatory networks

PLT1 and *PLT2* are directly regulated by the *TOPLESS (TLP)* transcriptional co-repressor (Smith and Long, 2010). *tpl* loss-of-function mutants display ectopic *PLT1* and *PLT2* expression and form a secondary root pole (Long et al., 2006; Smith and Long, 2010). A mutation in the *miR165/166* binding site of *PHABULOSA (PHB)* suppresses the formation of the *tpl* double root. *PHB*, like *REVOLUTA (REV)*, is a member of the *miR165/166*-regulated *HD-ZIP III* gene family of transcription factors that promote apical fate during early embryogenesis (Mallory et al., 2004). *HD-ZIP III* gene expression is absent in the apical region of *tpl* mutants, but a *miR165/166* sensor still accumulates indicating a mechanism for control of *HD-ZIP III* gene expression that is independent of *miR165/166* action. Apical expression of *PHB* and *REV* is restored in *tpl;plt1;plt2* triple mutants suggesting that *PLT1* and *PLT2* act as repressors of *HD-ZIP III* expression during embryogenesis (Smith and Long, 2010). In turn, genetic- and gain-of-function studies have shown that *HD-ZIP III* proteins repress the *PLT1-PLT2* pathway. Thus, the antagonistic action of the *HD-ZIP III* and *PLT1-PLT2* proteins may facilitate proper apical–basal patterning (Smith and Long, 2010) (Figure 3).

Ectopic root formation was also observed at the apical region of *ANGUSTIFOLIA3/GRF INTERACTING FACTOR1 (AN3/GIF1)* and *HANABA TARANU (HAN)* double loss-of-function mutant embryos (*an3/gif;han*; (Kanei et al., 2012)). *AN3/GIF1* encodes a putative transcriptional co-activator regulating various aspects of shoot development (Kim and Kende, 2004; Horiguchi et al., 2005; Rodriguez et al., 2010; Wang et al., 2011). *HAN* encodes a GATA-type transcription

factor required to position the proembryo boundary in the early *Arabidopsis* embryo (Zhao et al., 2004; Nawy et al., 2010). *PLT1* expression is expanded to the apical region of the globular embryo in *an3/gif1;han* double mutants, and ectopic root formation was suppressed by the *plt1* mutation. These results suggest that AN3/GIF1 and HAN repress *PLT1*, possibly via TPL, to establish apical identity (Kanei et al., 2012).

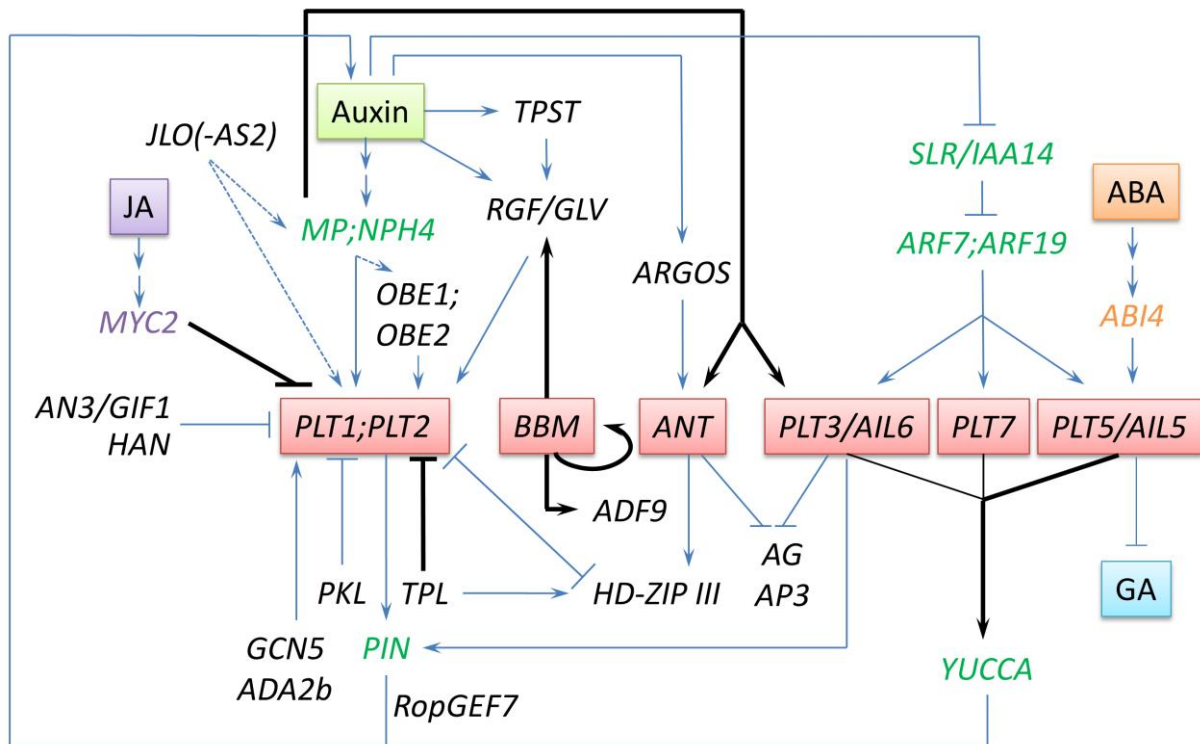


Figure 3. *AIL* gene regulatory networks. Genes in purple, green or orange are directly associated with the JA, auxin or ABA pathways, respectively. Solid black lines represent direct interactions, dashed lines show possible interactions.

Root development and meristem maintenance

PLT/*BBM* expression patterns set up during embryogenesis are maintained during post-embryonic root development (Figure 2B), where they show partly overlapping gradients of protein accumulation, with the highest proteins levels in the stem cell niche (Aida et al., 2004; Galinha et al., 2007). Concomitantly, the *SHORT ROOT* (*SHR*) transcription factor is expressed in the stele and moves to the surrounding cells, including the QC, where it activates *SCARECROW* (*SCR*) (Helariutta et al., 2000; Nakajima et al., 2001). *plt1;plt2*, *shr* and *scr* mutants display loss of (different) QC identity markers and premature termination of root growth (Sabatini et al., 2003; Aida et al., 2004). *PLT1* transcription is not affected in *shr* and *scr* mutants and *PLT1-PLT2* activity

is not required for *SHR* or *SCR* expression. Thus, QC specification requires the parallel action of *PLT1-PLT2* and *SHR-SCR* pathways (Aida et al., 2004).

Additive activities of AIL genes

The QC is the site of the root auxin maximum (Sabatini et al., 1999; Petersson et al., 2009; Brunoud et al., 2012), generated by local biosynthesis (Ljung et al., 2005; Stepanova et al., 2005; Brady et al., 2007; Stepanova et al., 2008; Ikeda et al., 2009), and through the activity of directional auxin transporters such as the PINs (Galweiler et al., 1998; Friml et al., 2002b; Friml et al., 2002a; Blilou et al., 2005). The PIN-mediated rootward-directed auxin transport in the vasculature and a shootward-directed transport in the lateral root cap and epidermis is thought to generate an auxin gradient in the root that is crucial for root meristem size and maintenance (Friml et al., 2002b; Blilou et al., 2005; Wisniewska et al., 2006; Grieneisen et al., 2007). The proposed root auxin gradient correlates with an expression gradient of PLT/BBM proteins that act in a largely additive fashion (Aida et al., 2004; Galinha et al., 2007): Firstly, stepwise lowering *PLT/BBM* gene dosage results in increasingly enhanced root phenotypes. Secondly, complementing *plt1;plt2* mutants with a shortened gradient of *PLT2* expression results in severely decreased root and meristem sizes, but rescues the stem cell niche. Thirdly, expression of *PLT2* in the transition zone of the *plt1;plt2* mutant meristem prolongs meristem activity, but fails to maintain the stem cell niche. Finally, inducible *PLT2* over-expression results in continuous growth of the meristematic zone while inhibiting cell expansion at the elongation zone (Galinha et al., 2007). Simultaneously lowering *RETINOBLASTOMA-RELATED* (*RBR*) levels, as an independent stem cell input (Wildwater et al., 2005), and inducing overexpression of *PLT2* increases stem cell numbers and activity (Galinha et al., 2007). Thus, high *PLT/BBM* levels in the QC are required to specify and maintain the root stem cell niche, intermediate levels are required for proliferation of root meristem cells, and exit from the meristem to differentiation requires *PLT/BBM* levels to drop below a certain threshold (Galinha et al., 2007). Together, the auxin-*PLT/BBM* gradient can provide a concentration-based mechanism for specification and maintenance of root stem cells, as well as for regulating proliferation, elongation, and differentiation of cells.

Translating auxin to AIL expression

Several studies implicate ARFs as general upstream mediators of the auxin-AIL pathway during embryogenesis (MP and NPH4), flower primordia initiation and ovule development (MP) and lateral root formation (ARFs 7 and 19; (Aida et al., 2004; Galbiati et al., 2013b; Hofhuis et al., 2013; Yamaguchi et al., 2013) and see below).

Auxin also positively regulates tyrosylprotein sulfotransferase (*TPST*) and several tyrosine-sulfated peptide encoding *ROOT GROWTH FACTOR (RGF)/GOLVEN (GLV)* genes that are highly expressed in the root stem cell niche (Matsuzaki et al., 2010; Zhou et al., 2010; Fernandez et al., 2013). *tpst* mutants have pleiotropic phenotypes, including short roots, additional QC cells and differentiation of columella stem cells. The observed root defects correlate with impaired expression of auxin biosynthesis genes and *PLT1* and *PLT2* transcript and protein levels. *tpst* root defects were partially restored in the presence of RGF1 peptide. Induced overexpression of *PLT2* also partially rescued *tpst* mutant root defects, whereas *plt1;plt2* mutant roots could not be complemented by addition of RGF1 peptide (Matsuzaki et al., 2010; Zhou et al., 2010). In addition, RGF1-treated roots show an expanded *PLT2* protein expression domain without associated changes in transcription (Matsuzaki et al., 2010). Together these results suggest TPST-sulfated RGF peptides link the auxin pathway to transcriptional and posttranscriptional regulation of *PLT/BBM* in root stem cell maintenance (Figure 3).

AIL genes integrate multiple hormonal inputs

Germination results in emergence of the radicle (hypocotyl + embryonic root) from the seed through cell enlargement and activation of the embryonic root meristem (Balaji et al., 2005). This process is regulated by environmental factors, such as nutrients, light and temperature, as well as by the endogenous plant hormones ABA and gibberellic acid (GA) (Bewley, 1997). The *ail5/plt5* mutant was first identified (and named *chotto1*, Table 1) as a mutant showing reduced sensitivity to ABA during germination (Nambara et al., 2002). Elevated *AIL5/PLT5* expression in the radicle of imbibed seeds requires the transcription factor ABA INSENSITIVE 4 (*ABI4*), a key component of ABA-mediated signaling in seeds (Finkelstein et al., 1998), and genetic analyses confirms that *AIL5/PLT5* acts downstream of *ABI4* (Yamagishi et al., 2009) (Figure 3). Freshly harvested *ail5/plt5* mutant seeds show reduced primary seed dormancy, which was recently also shown for *abi4* (Yamagishi et al., 2009; Yano et al., 2009; Shu et al., 2013). However, neither ABA nor GA seems to regulate *AIL5/PLT5* expression (Yamagishi et al., 2009; Yano et al., 2009). Genetic analysis further shows that the *ail5/plt5* mutant background partially restores GA

biosynthesis and the associated delayed germination phenotype in an ABA over-accumulating mutant. However, *ail5/plt5* is not able to restore germination in a GA-deficient mutant (Yano et al., 2009). Finally, *ail5/plt5* seedlings display partial insensitivity to excess nitrate independent of *ABI4*. These results indicate that AIL5/PLT5 integrates multiple and independent inputs towards regulation of germination and seedling growth.

Jasmonate (JA) negatively affects root growth in a dose-dependent manner resulting in reduced root cell sizes, a reduced meristem size and activity, and QC divisions followed by columella stem cell differentiation (Chen et al., 2011). The JA growth inhibitory effect is mediated via its receptor COI1 (Yan et al., 2009; Sheard et al., 2010). The downstream MYC2 transcription factor (Boter et al., 2004; Lorenzo et al., 2004) acts, at least in part, through direct transcriptional downregulation of the *PLT1* and *PLT2* genes, independent of the auxin pathway (Chen et al., 2011) (Figure 3). The effects of JA application are only partly reduced in *plt1;plt2* mutant and *PLT2* overexpression roots, suggesting that other PLT/BBM members also play a role in mediating JA effects (Chen et al., 2011).

Regulation of PLT expression at the chromatin level

Chromatin factors affect key regulators of cell proliferation and stem cell maintenance such as *WUS*, *WOX5*, *PLT1* and *PLT2* genes (Kaya et al., 2001; Bertrand et al., 2003; Kwon et al., 2005; Kornet and Scheres, 2009; Anzola et al., 2010; Aichinger et al., 2011). The histone acetyltransferase GCN5 and its associated factor ADA2b were found to regulate *PLT1* and *PLT2* gene expression levels and genetic analysis confirmed that GCN5 and ADA2b act in the *PLT1-PLT2* pathway (Kornet and Scheres, 2009) (Figure 3). In addition, the root meristem of *gcn5* mutants, which is gradually lost due to compromised stem cell niche maintenance, is partially rescued by induced overexpression of *PLT2* (Kornet and Scheres, 2009).

Mutation of the chromatin remodeling factor *PICKLE* (*PKL*) causes short primary roots with a reduced meristem size due to impaired stem cell niche activity. By contrast, mutation of the Polycomb-group (PcG) gene *CURLY LEAF* (*CLF*) results in longer roots with an increased meristem size which is associated with increased stem cell activity (Aichinger et al., 2011). These mutants do not show altered auxin accumulation. Rather, their phenotypes correlated with, and were shown genetically to be caused by, decreased *PLT1-PLT2* expression or increased *WOX5* expression, respectively. Decreased expression levels of root stem cell and meristem marker genes in *pkl* correlated with increased levels of (PcG-mediated) trimethylation of histone H3 on

lysine 27 (H3K27me3) at their promoter regions, indicating that root stem cell niche activity is regulated by the antagonistic activities of PcG proteins and PKL (Aichinger et al., 2011). The similar phenotypes of *gcn5* and *pk1* suggest linked gene activities. PKL-mediated remodelling may allow the recruitment of a GCN5-ADA2b containing histone acetyltransferase complex, thereby generating a suitable chromatin state for the accurate interpretation of auxin-dependent PLT signals and reinforcing stem cell fate (Aichinger et al., 2011). Such a stem cell-specific chromatin state would be analogous to that observed in animal systems (Meshorer and Misteli, 2006; Serrano et al., 2013).

The *OBERON1* (*OBE1*) and *OBE2* genes encode plant homeodomain (PHD)-finger proteins, so-called epigenetic readers that recognize and bind to both methylated and unmodified histone H3 tails (Saiga et al., 2008; Musselman and Kutateladze, 2011). The OBE proteins act redundantly in MP-dependent embryonic root initiation (Saiga et al., 2008; Thomas et al., 2009). *obe1;obe2* double mutants resemble *mp* mutants and lack *PLT1* and *PLT2* expression, whereas *MP* expression is unaffected (Saiga et al., 2008). Since *PLT1* and *PLT2* expression is dependent on MP (Aida et al., 2004), OBE1 and OBE2 may act to control embryonic root meristem formation downstream or at the level of MP, and in this way control *PLT1* and *PLT2* expression (Saiga et al., 2008) (Figure 3).

Shoot and flower meristem maintenance

ANT, *AIL6/PLT3* and *PLT7* are expressed in distinct, but overlapping domains in the inflorescence and vegetative shoot meristems. *AIL6/PLT3* and *PLT7* are expressed throughout the meristem, but their expression is elevated in the peripheral zones (PZ) and in the central zone (CZ), respectively (Prasad et al., 2011; Mudunkothge and Krizek, 2012). By contrast, *ANT* is expressed exclusively in the PZ and marks the cryptic bract region of the floral meristem (Long and Barton, 1998; Long and Barton, 2000). The *ant;ail6/plt3;plt7* triple mutant shoot stops growing after the production of a few leaves due to reduced cell divisions in the meristem and differentiation of the meristematic cells, showing that *ANT*, *AIL6/PLT3* and *PLT7* genes are required for shoot apical meristem (SAM) maintenance (Mudunkothge and Krizek, 2012) (Figure 2D). The *ant* and *ail6/plt3* mutations enhance, whereas *plt7* partially rescues *wus* and *shoot meristemless* (*stm*) phenotypes, indicating that *ANT*, *AIL6/PLT3* and *PLT7* do not function in a strictly redundant fashion (Mudunkothge and Krizek, 2012).

Both the *ant* and *ap2* single mutants have fewer cells in their floral meristems (Figure 2F), resulting in a reduced number of floral organs (Bowman et al., 1991; Krizek, 1999). When *ap2* and *ant* mutations are combined, sepal, petal and stamen formation is completely abolished (Elliott et al., 1996). This organ loss is likely to be a result of reduced proliferation due to ectopic expression of the growth-repressing class C MADS-box gene *AGAMOUS* (*AG*) (Elliott et al., 1996; Krizek et al., 2000).

Organ development

Phyllotaxy and rhizotaxy

Cotyledons and the first pair of leaves are formed opposite to each other in Arabidopsis, while subsequent leaves develop in a spiral pattern (Figure 2E). Unlike *AIL6/PLT3* and *PLT7*, *AIL5/PLT5* is expressed in a largely uniform manner throughout the entire SAM (Prasad et al., 2011; Mudunkothge and Krizek, 2012). In the *ail5/plt5;ail6/plt3;plt7* triple mutant, the switch to the spiral phyllotactic pattern is delayed by a few leaf pairs, with double mutant combinations showing lower penetrance (Prasad et al., 2011). The spiral pattern in the inflorescence meristem (Figure 2E) is also affected in this triple mutant, which shows a tendency to develop successive siliques at angles of $\sim 90^\circ$ or $\sim 180^\circ$ instead of at 137.5° in wild-type Arabidopsis. Reducing PIN1 levels leads to the same defects in the phyllotactic pattern as reduced *AIL/PLT* expression (Prasad et al., 2011). It was shown recently by *ail5/plt5;ail6/plt3;plt7* complementation experiments that *AIL/PLT*-induced auxin biosynthesis genes, *YUCCA1* (*YUC1*) and *YUC4*, mediate high auxin levels in the CZ that are required for the correct spiral phyllotactic pattern (Pinon et al., 2013).

Lateral root primordia (LRP) are distributed along the Arabidopsis root and their spacing correlates with the root curvature (Dubrovsky et al., 2006; Laskowski et al., 2008) (Figure 2C). *ARF7* and *ARF19*, together with the gene encoding SOLITARY ROOT (*SLR*)/*IAA14*, which represses these ARFs, are all expressed in the vascular tissue and lateral root initiation sites (Fukaki et al., 2002; Okushima et al., 2005; Vanneste et al., 2005; Wilmoth et al., 2005). *arf7;arf19* double mutant roots are defective in the first asymmetric division leading to LRP formation and almost completely lack lateral roots. *AIL6/PLT3*, *AIL5/PLT5*, and *PLT7* are expressed prior to the first founder cell division and their expression is absent in *arf7;arf19* mutants (Hofhuis et al., 2013). Outgrowth of LRP is severely impaired in *ail6/plt3;plt7* double and *ail6/plt3;ail5/plt5;plt7* triple

mutants, and LRP are frequently grouped in longitudinal or radial clusters. Complementation experiments reveal that expression of *AIL6/PLT3*, *AIL5/PLT5* or *PLT7* from their endogenous promoters rescues LRP emergence more readily than the clustering phenotype, suggesting distinct targets and/or dosage dependency reminiscent of PLT/BBM function in the primary root (Galinha et al., 2007; Hofhuis et al., 2013). Overexpression of *AIL6/PLT3*, *AIL5/PLT5* or *PLT7* in *arf7;arf19* can overcome the block in LRP formation. These results indicate that the *AIL6/PLT3*, *AIL5/PLT5* and *PLT7* genes are downstream components of the ARF7 and ARF19-mediated auxin response module (Hofhuis et al., 2013) (Figure 3), although they do not appear to be direct targets of ARF7 and ARF19 (Okushima et al., 2005; Wilmoth et al., 2005).

Despite the differences in the processes of shoot and root primordia initiation, the involvement of the same three AIL proteins in both processes suggests that similar mechanisms regulate plant architecture in different organs.

Organ size

Loss-of-function *ant* mutants have smaller leaves and flowers as a result of having fewer cells, while overexpression of *ANT* prolongs the cell proliferation period, leading to larger leaves and flowers with more and/or larger cells (Krizek, 1999; Mizukami and Fischer, 2000) (Figure 2D). The reduced cell proliferation observed in *ant* is enhanced in the *ant;ail6/plt3* double mutant, which has even smaller leaves (Krizek, 2009). Overexpression of the auxin-inducible gene *AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE (ARGOS)* prolongs *ANT* expression and leads to the formation of larger organs (Hu et al., 2003). This effect was lost in the *ant* mutant, suggesting that *ARGOS* functions downstream of auxin and upstream of *ANT* in organ size control (Hu et al., 2003) (Figure 3).

An increase in floral organ size due to increased cell number was also observed in *AIL5/PLT5* and *AIL6/PLT3* overexpression lines (Nole-Wilson et al., 2005; Krizek and Eaddy, 2012). However, very high expression levels of *AIL6/PLT3* block cellular differentiation resulting in floral organs with altered morphology (Krizek and Eaddy, 2012). The very small and aberrantly shaped sepals of these *AIL6/PLT3* overexpression lines are unable to cover the entire flower bud, a phenotype that was also reported in *BBM* overexpression lines (Boutillier et al., 2002; Krizek and Eaddy, 2012). Together with the sustained meristem cell proliferation observed upon *PLT2* overexpression (Galinha et al., 2007), these results indicate that AIL transcription factors regulate the balance between cell proliferation and differentiation.

Organ polarity

Double and triple mutant combinations of *ant* with loss-of-function alleles of the polarity genes *FILAMENTOUS FLOWER (FIL)* and *YABBY3* produce dwarfed seedlings that develop small, narrow leaves with loss of both adaxial and abaxial epidermal cell characteristics, which are not observed in single mutants (Siegfried et al., 1999; Nole-Wilson and Krizek, 2006). These defects are also present in some floral organs. The expression of the adaxial specifying *HD-ZIP III* gene *PHB* is reduced in *fil;ant* flowers, suggesting that ANT and FIL act together to promote organ polarity (Figure 2D) by activating *PHB* (Nole-Wilson and Krizek, 2006) (Figure 3).

Flower and floral organ development

Floral initiation

LEAFY (LFY) encodes a transcriptional regulator that promotes the transition to flowering and also specifies floral meristem identity (Schultz and Haughn, 1991; Weigel et al., 1992; Weigel and Meyerowitz, 1993; Weigel and Nilsson, 1995). Auxin treatment induces higher *LFY* gene expression, and results in precocious flower formation. Auxin response elements in the promoter of *LFY* are bound by the MP and BDL proteins (Yamaguchi et al., 2013). Initiation of flower primordia from the inflorescence meristem (Figure 2F) is known to rely on MP; *mp* mutants form naked inflorescence pins lacking flowers, phenocopying loss-of-function mutants in the PIN1 auxin efflux carrier (Przemeck et al., 1996; Galweiler et al., 1998). *MP* and *LFY* expression overlaps with that of the *ANT* and *AIL6/PLT3* genes, and MP was shown to bind to the promoters- and activate expression of *ANT* and *AIL6/PLT3* (Yamaguchi et al., 2013). The pin-like inflorescences formed in the *ant;ail6/plt3;lfy* triple mutant and the additive ability of *ANT* and *LFY* overexpression to rescue flower primordium initiation in the hypomorphic *mp-S319* mutant indicates that ANT and LFY function in parallel pathways downstream of MP in the initiation of flower primordia (Yamaguchi et al., 2013) (Figure 3). As with AIL proteins, LFY feeds back positively into the auxin pathway (Prasad et al., 2011; Pinon et al., 2013; Yamaguchi et al., 2013). MP-dependent upregulation of *LFY*, *ANT* and *AIL6/PLT3*, together with their positive feedback to the auxin pathway, may commit cells to flower primordium formation (Yamaguchi et al., 2013).

Floral organ identity

The petals of the *ant* single mutant develop stomata, which are absent in wild-type petals, indicating partial loss of petal identity in this mutant (Krizek et al., 2000). A more severe phenotype in which the petals become partially stamens was observed in the hypomorphic *ap2-1;ant* double mutant (Krizek et al., 2000). Expression of *AG* was extended to the second whorl in the *ap2-1;ant* mutant where it caused the partial homeotic transformation (Krizek et al., 2000). Single *ail6/plt3* mutants develop normally, but flowers of the *ant;ail6/plt3* double mutant have reduced petal, stamen and carpel identities and consist primarily of small sepals, filamentous organs and unfused carpel valves, showing that *ANT* and *AIL6/PLT3* function redundantly in determining floral organ identity (Krizek, 2009) (Figure 2F). This phenotype correlates with altered expression of the floral organ identity genes *APETALA3* (*AP3*, B-class) and *AG* (C-class) (Krizek, 2009) and suggests their regulation by *ANT* and *AIL6/PLT3* (Figure 3).

Ovule development

The number of ovule primordia in the *ant* mutant is reduced due to carpel margin meristem (CMM) defects. Also the integuments of *ant* ovules do not initiate properly and female gametophyte development is disrupted leading to female-sterility (Elliott et al., 1996; Klucher et al., 1996; Baker et al., 1997; Schneitz et al., 1997). The *CUP SHAPED COTYLEDON 1* (*CUC1*) and *CUC2* genes act in parallel to *ANT* in controlling the number of developing ovule primordia (Galbiati et al., 2013a). *ANT*, *CUC1* and *CUC2* are downregulated in hypomorphic *mp-5319* mutant pistils that do not develop ovules due to absence of carpel margin tissue. Binding of MP to the *ANT*, *CUC1* and *CUC2* genomic regions was confirmed in this tissue (Galbiati et al., 2013a). Thus, MP acts as an upstream regulator of *ANT* during ovule development (Figure 3), reminiscent of its role in regulating *AIL* gene expression during embryogenesis and floral meristem initiation (Aida et al., 2004; Galbiati et al., 2013b; Yamaguchi et al., 2013).

HUELLENLOS (*HLL*) and *ANT* play overlapping roles in ovule primordium outgrowth (Schneitz et al., 1998). The transcriptional co-regulators *SEUSS* (*SEU*) and *LEUNIG* (*LUG*) show partial redundancy with *ANT* in the control of medial domain development (Liu et al., 2000; Azhakanandam et al., 2008), and these defects are associated with reduced expression levels of the *HD-ZIP III* genes *REV* and *PHB* (Azhakanandam et al., 2008). Double mutant *ant;rev* gynoecia exhibit loss of CMM-derived tissue, which is correlated with lower expression of *PHB* (Nole-Wilson et al., 2010). These studies show that *ANT* plays a central role in ovule development (Figure 2F) through activation of *HD-ZIP III* genes (Figure 3).

AIL target genes

Although many genetic interactions have been described for *AIL* genes, our knowledge about target genes that are directly regulated by AIL binding remains limited. Both AP2/ERF domains of ANT are required for DNA binding and each domain is thought to use different amino acids to contact the DNA (Nole-Wilson and Krizek, 2000; Krizek, 2003) (Box1). An ANT DNA binding motif was determined using SELEX, an *in vitro* oligonucleotide selection method, which yielded the consensus sequence gCAC(A/G)N(A/T)TcCC(a/g)ANG(c/t) (Nole-Wilson and Krizek, 2000). AIL5/PLT5 also binds to the ANT consensus sequence *in vitro* (Yano et al., 2009), suggesting that the observed overlapping functions of AIL proteins is realized through a set of common target genes. The ANT DNA binding motif is different from that of other single AP2/ERF proteins, which bind to short GC-rich boxes (Hao et al., 1998; Fujimoto et al., 2000; Shoji et al., 2013).

Currently, only the targets of BBM have been identified. Using a microarray approach in which the BBM protein was inducibly activated in the presence of the translational inhibitor cycloheximide to identify direct targets, it was shown that the vast majority of target genes are upregulated upon DNA binding, suggesting that BBM acts mainly as a transcriptional activator (Passarinho et al., 2008). BBM was reported to interact with TPL-related (TPR) corepressors in yeast (Causier et al., 2012), suggesting downregulation of the limited number of target genes requires recruitment of repressor proteins. AP2 has also been shown to form complexes with TOPLESS (TPL) to repress the floral organ identity genes *AG*, *AP3*, *PI* and *SEP3* (Krogan et al., 2012), suggesting that similar protein-protein complexes are co-opted by different groups of AP2-family proteins.

The functional diversity of BBM targets suggests that this protein functions in a wide range of developmental pathways, which is consistent with the large number of mutant studies in which *BBM* and other *AIL* genes were shown to be regulated incorrectly. One of the few functionally characterized BBM targets is *RGF8/GLV6*, which encodes a homolog of the RGF1 signalling peptide shown to regulate *PLT1* and *PLT2* expression (Matsuzaki et al., 2010). This suggests a positive feedback loop between *RGF/GLV* and *PLT/BBM* genes (Figure 3) that ensures sufficiently high *PLT/BBM* levels to preserve the meristem. Another gene that is directly upregulated by BBM is *ACTIN DEPOLYMERIZING FACTOR9 (ADF9)* (Passarinho et al., 2008), which encodes an ADF/cofilin cytoskeletal protein (Carlier et al., 1997; Gungabissoon et al., 1998). ADF9 is required for hormone-mediated cell proliferation *in vitro* (Burgos-Rivera et al., 2008). Notably, only a few minutes of inducible BBM overexpression was sufficient to stimulate

reorganization of the actin cytoskeleton. ADF9 therefore provides a potential link between the actin cytoskeleton and BBM-mediated meristematic growth. Finally, it was shown that BBM also stimulates its own expression, suggesting it functions in a positive auto-regulatory feedback loop to control its own activity (Passarinho et al., 2008). BBM targets support a role for AIL proteins in the regulation of cell proliferation and differentiation.

Conclusions and future prospects

Almost twenty years of AIL research in *Arabidopsis* has shown that these transcription factors are crucial for stem cell niche specification, meristem maintenance, organ initiation and growth, and function as master regulators of embryogenesis and root formation. At present, functions have been ascribed to all *Arabidopsis* AIL genes except *AIL1*. Research on AIL expression and function in other species is still at an early stage. Analysis of the expression patterns of ANT-like genes from different species suggest that they may (Shigyo et al., 2006; Yamada et al., 2008; Mizumoto et al., 2009), or may not (Rieu et al., 2005; Mizumoto et al., 2009) be conserved with the *Arabidopsis* ANT gene. Nevertheless, functional studies on AIL proteins from different species indicate that they also regulate stem cell specification and plant developmental processes, as in *Arabidopsis*. For example, in poplar (*Populus* spp.), ANT-like AILs have been shown to control meristem activity during adventitious rooting (Rigal et al., 2012) and to directly feed into the cell cycle machinery to establish growth cessation upon exposure to short days (Karlberg et al., 2011). Similarly, in the moss *Physcomitrella patens*, four AIL genes are indispensable for the formation of one type of moss stem cell, the apical gametophore cell (Aoyama et al., 2012).

AIL proteins interact with auxin pathways throughout plant development and at multiple levels, including via ARFs and through PIN feedback loops (Figure 3). Future studies will undoubtedly reveal additional auxin-AIL relationships, as well as new relationships with other hormones, signaling pathways and chromatin-level changes. Factors involved in regulating AIL expression during the earliest stages of embryogenesis have not been identified, even though at least two AILs, BBM and PLT2, are required for progression through early embryogenesis.

Specific AIL family members appear to be expressed and function in either the root or shoot meristem. This division could reflect protein sub-functionalisation, which is also suggested by the different AIL-HD-ZIP III genetic interactions that appear as a general module in the regulation of several developmental processes (Figure 3). Alternatively, the regional expression and interaction of AIL proteins with locally expressed transcriptional co-activators or repressors

may be responsible for this differential effect. Identification of the protein complexes in which AIL proteins act may shed light on this matter. In addition, promoter swapping to express root AILs in the shoot and *vice versa* may reveal how much of this apparent subfunctionalisation is due to differences in protein characteristics versus gene expression patterns.

One aspect of AIL function is their dosage-dependent regulation of developmental processes, as observed during root development. This dosage-dependency could be mediated by quantitative differences in the activation of target genes containing AIL-specific DNA binding motifs. Alternatively, AIL proteins could exert their dosage-dependent regulation via differences in binding affinity for DNA motifs present in common target gene promoters. The fact that both ANT and AIL5/PLT5 can bind the same DNA motif suggests a common set of downstream targets. Elucidating the targets and DNA binding sites of additional AIL proteins may clarify either of the above scenarios.

A poorly understood aspect of AIL function is their ability to induce totipotency and pluripotency when overexpressed. This property offers new possibilities to improve the efficiency of plant regeneration for a range of biotechnology applications. For example, the ability of BBM to induce somatic embryogenesis was used to improve Chinese white poplar and sweet pepper (*Capsicum annuum*) transformation (Deng et al., 2009; Heidmann et al., 2011). More importantly, a better understanding of how AIL proteins function will provide a framework for understanding how regeneration is regulated *in vitro* and during normal plant development.

Acknowledgements

We thank Gabino Sanchez-Perez for constructing an updated phylogenetic tree for the AIL family, Ben Scheres, Gerco Angenent and Stephen Grigg for valuable comments on the manuscript. AH was supported by a Technology Top Institute Green Genetics project. RH was supported by a Netherlands Genomics Initiative Horizon grant (no. 93516054) from The Netherlands Organization for Scientific Research.

References

- Aichinger, E., Villar, C.B.R., Di Mambro, R., Sabatini, S., and Köhler, C.** (2011). The CHD3 Chromatin Remodeler PICKLE and Polycomb Group Proteins Antagonistically Regulate Meristem Activity in the Arabidopsis Root. *Plant Cell* **23**, 1047-1060.
- 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.
- Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M.** (1998). A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO journal* **17**, 5484-5496.
- Anzola, J.M., Sieberer, T., Ortbauer, M., Butt, H., Korbei, B., Weinhofer, I., Müllner, A.E., and Luschig, C.** (2010). Putative Arabidopsis Transcriptional Adaptor Protein (PROPORZ1) is required to modulate histone acetylation in response to auxin. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10308-10313.
- Aoyama, T., Hiwatashi, Y., Shigyo, M., Kofuji, R., Kubo, M., Ito, M., and Hasebe, M.** (2012). AP2-type transcription factors determine stem cell identity in the moss *Physcomitrella patens*. *Development* **139**, 3120-3129.
- Azhakanandam, S., Nole-Wilson, S., Bao, F., and Franks, R.G.** (2008). SEUSS and AINTEGUMENTA mediate patterning and ovule initiation during gynoecium medial domain development. *Plant Physiol* **146**, 1165-1181.
- Baker, S.C., Robinson-Beers, K., Villanueva, J.M., Gaiser, J.C., and Gasser, C.S.** (1997). Interactions among genes regulating ovule development in *Arabidopsis thaliana*. *Genetics* **145**, 1109-1124.
- Balaji, S., Babu, M.M., Iyer, L.M., and Aravind, L.** (2005). Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* **33**, 3994-4006.
- Bandupriya, H.D.D., Gibbings, J.G., and Dunwell, J.M.** (2013). Isolation and characterization of an *AINTEGUMENTA*-like gene in different coconut (*Cocos nucifera* L.) varieties from Sri Lanka. *Tree Genetics & Genomes* **9**, 813-827.
- Benková, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertová, D., Jürgens, G., and Friml, J.** (2003). Local, Efflux-Dependent Auxin Gradients as a Common Module for Plant Organ Formation. *Cell* **115**, 591-602.
- Bertrand, C., Bergounioux, C., Domenichini, S., Delarue, M., and Zhou, D.-X.** (2003). Arabidopsis Histone Acetyltransferase AtGCN5 Regulates the Floral Meristem Activity through the WUSCHEL/AGAMOUS Pathway. *J. Biol. Chem.* **278**, 28246-28251.
- Bewley, J.D.** (1997). Seed Germination and Dormancy. *Plant Cell* **9**, 1055-1066.
- Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K., and Scheres, B.** (2005). The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. *Nature* **433**, 39-44.
- Borghi, L., Bureau, M., and Simon, R.** (2007). Arabidopsis JAGGED LATERAL ORGANS is expressed in boundaries and coordinates KNOX and PIN activity. *Plant Cell* **19**, 1795-1808.
- Boter, M., Ruiz-Rivero, O., Abdeen, A., and Prat, S.** (2004). Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and Arabidopsis. *Genes Dev.* **18**, 1577-1591.
- Boutillier, K., Offringa, R., Sharma, V.K., Kieft, 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. *The Plant cell* **14**, 1737-1749.
- Bowman, J.L., Drews, G.N., and Meyerowitz, E.M.** (1991). Expression of the Arabidopsis floral homeotic gene AGAMOUS is restricted to specific cell types late in flower development. *Plant Cell* **3**, 749-758.
- 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.
- Braybrook, S.A., and Harada, J.J.** (2008). LECs go crazy in embryo development. *Trends Plant Sci.* **13**, 624-630.
- Brunoud, G., Wells, D.M., Oliva, M., Larrieu, A., Mirabet, V., Burrow, A.H., Beeckman, T., Kepinski, S., Traas, J., Bennett, M.J., and Vernoux, T.** (2012). A novel sensor to map auxin response and distribution at high spatio-temporal resolution. *Nature* **482**, 103-106.
- Bureau, M., Rast, M., Illmer, J., and Simon, R.** (2010). JAGGED LATERAL ORGAN (JLO) controls auxin dependent patterning during development of the Arabidopsis embryo and root. *Plant Mol. Biol.* **74**, 479-491.
- Burgos-Rivera, B., Ruzicka, D.R., Deal, R.B., McKinney, E.C., King-Reid, L., and Meagher, R.B.** (2008). ACTIN DEPOLYMERIZING FACTOR9 controls development and gene expression in Arabidopsis. *Plant Mol. Biol.* **68**, 619-632.
- Cao, Z.F., Li, J., Chen, F., Li, Y.Q., Zhou, H.M., and Liu, Q.** (2001). Effect of two conserved amino acid residues on DREB1A function. *Biochemistry (Mosc.)* **66**, 623-627.
- Carlier, M.F., Laurent, V., Santolini, J., Melki, R., Didry, D., Xia, G.X., Hong, Y., Chua, N.H., and Pantaloni, D.** (1997). Actin depolymerizing factor (ADF/cofilin) enhances the rate of filament turnover: implication in actin-based motility. *J. Cell Biol.* **136**, 1307-1322.
- Causier, B., Ashworth, M., Guo, W., and Davies, B.** (2012). The TOPLESS interactome: a framework for gene repression in Arabidopsis. *Plant Physiol* **158**, 423-438.
- Chen, Q., Sun, J., Zhai, Q., Zhou, W., Qi, L., Xu, L., Wang, B., Chen, R., Jiang, H., Qi, J., Li, X., Palme, K., and Li, C.** (2011). The basic helix-loop-helix transcription factor MYC2 directly represses PLETHORA expression during jasmonate-mediated modulation of the root stem cell niche in Arabidopsis. *Plant Cell* **23**, 3335-3352.
- De Smet, I., Tetsumura, T., De Rybel, B., Frey, N.F.d., Laplaze, L., Casimiro, I., Swarup, R., Naudts, M., Vanneste, S., Audenaert, D., Inzé, D., Bennett, M.J., and Beeckman, T.** (2007). Auxin-dependent regulation of lateral root positioning in the basal meristem of Arabidopsis. *Development* **134**, 681-690.
- Deng, W., Luo, K.M., Li, Z.G., and Yang, Y.W.** (2009). A novel method for induction of plant regeneration via somatic embryogenesis. *Plant Science* **177**, 43-48.
- Dietz, K.J., Vogel, M.O., and Viehhauser, A.** (2010). AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. *Protoplasma* **245**, 3-14.
- Duan, C., Argout, X., Gebelin, V., Summo, M., Dufayard, J.F., Leclercq, J., Kuswanhadi, Piyatrakul, P., Pirrello, J., Rio, M., Champion, A., and Montoro, P.** (2013). Identification of the Hevea brasiliensis AP2/ERF superfamily by RNA sequencing. *BMC Genomics* **14**, 30.
- Dubrovsky, J.G., Gambetta, G.A., Hernandez-Barrera, A., Shishkova, S., and Gonzalez, I.** (2006). Lateral root initiation in Arabidopsis: developmental window, spatial patterning, density and predictability. *Ann Bot* **97**, 903-915.
- El Ouakfaoui, S., Schnell, J., Abdeen, A., Colville, A., Labbe, H., Han, S., Baum, B., Laberge, S., and Miki, B.** (2010). Control of somatic embryogenesis and embryo development by AP2 transcription factors. *Plant Mol. Biol.* **74**, 313-326.

- Elliott, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q., Gerentes, D., Perez, P., and Smyth, D.R.** (1996). AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. *Plant Cell* **8**, 155-168.
- Fernandez, A., Drozdzecki, A., Hoogewijs, K., Nguyen, A., Beeckman, T., Madder, A., and Hilson, P.** (2013). Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation. *Plant Physiol* **161**, 954-970.
- Finkelstein, R.R., Li Wang, M., Lynch, T.J., Rao, S., and Goodman, H.M.** (1998). The Arabidopsis Abscisic Acid Response Locus ABI4 Encodes an APETALA2 Domain Protein. *Plant Cell* **10**, 1043-1054.
- Friml, J., Wisniewska, J., Benkova, E., Mendgen, K., and Palme, K.** (2002a). Lateral relocation of auxin efflux regulator PIN3 mediates tropism in Arabidopsis. *Nature* **415**, 806-809.
- Friml, J., Benková, E., Blilou, I., Wisniewska, J., Hamann, T., Ljung, K., Woody, S., Sandberg, G., Scheres, B., Jürgens, G., and Palme, K.** (2002b). AtPIN4 Mediates Sink-Driven Auxin Gradients and Root Patterning in Arabidopsis. *Cell* **108**, 661-673.
- Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., and Ohme-Takagi, M.** (2000). Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell* **12**, 393-404.
- Fukaki, H., Tameda, S., Masuda, H., and Tasaka, M.** (2002). Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis. *Plant J.* **29**, 153-168.
- Galbiati, F., Sinha Roy, D., Simonini, S., Cucinotta, M., Ceccato, L., Cuesta, C., Simaskova, M., Benkova, E., Kamiuchi, Y., Aida, M., Weijers, D., Simon, R., Masiero, S., and Colombo, L.** (2013a). An integrative model of the control of ovule primordia formation. *The Plant journal : for cell and molecular biology*.
- Galbiati, F., Sinha Roy, D., Simonini, S., Cucinotta, M., Ceccato, L., Cuesta, C., Simaskova, M., Benkova, E., Kamiuchi, Y., Aida, M., Weijers, D., Simon, R., Masiero, S., and Colombo, L.** (2013b). An integrative model of the control of ovule primordia formation. *Plant J.* **76**, 446-455.
- 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.
- Galweiler, L., Guan, C., Muller, A., Wisman, E., Mendgen, K., Yephremov, A., and Palme, K.** (1998). Regulation of polar auxin transport by AtPIN1 in Arabidopsis vascular tissue. *Science* **282**, 2226-2230.
- Grieneisen, V.A., Xu, J., Maree, A.F., Hogeweg, P., and Scheres, B.** (2007). Auxin transport is sufficient to generate a maximum and gradient guiding root growth. *Nature* **449**, 1008-1013.
- Gungabissoon, R.A., Jiang, C.J., Drobak, B.K., Maciver, S.K., and Hussey, P.J.** (1998). Interaction of maize actin-depolymerising factor with actin and phosphoinositides and its inhibition of plant phospholipase C. *Plant J.* **16**, 689-696.
- Hao, D., Ohme-Takagi, M., and Sarai, A.** (1998). Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. *J. Biol. Chem.* **273**, 26857-26861.
- Harper, R.M., Stowe-Evans, E.L., Luesse, D.R., Muto, H., Tatematsu, K., Watahiki, M.K., Yamamoto, K., and Liscum, E.** (2000). The NPH4 locus encodes the auxin response factor ARF7, a conditional regulator of differential growth in aerial Arabidopsis tissue. *Plant Cell* **12**, 757-770.

- Hazak, O., and Yalovsky, S.** (2010). An auxin regulated positive feedback loop integrates Rho modulated cell polarity with pattern formation. *Plant Signal Behav* **5**, 709-711.
- Heidmann, I., de Lange, B., Lambalk, J., Angenent, G.C., and Boutilier, K.** (2011). Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. *Plant Cell Rep.* **30**, 1107-1115.
- Helariutta, Y., Fukaki, H., Wysocka-Diller, J., Nakajima, K., Jung, J., Sena, G., Hauser, M.T., and Benfey, P.N.** (2000). The SHORT-ROOT gene controls radial patterning of the Arabidopsis root through radial signaling. *Cell* **101**, 555-567.
- Hofhuis, H., Laskowski, M., Du, Y., Prasad, K., Grigg, S., Pinon, V., and Scheres, B.** (2013). Phyllotaxis and Rhizotaxis in Arabidopsis Are Modified by Three PLETHORA Transcription Factors. *Curr. Biol.* **23**, 956-962.
- Horiguchi, G., Kim, G.T., and Tsukaya, H.** (2005). The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. *Plant J.* **43**, 68-78.
- Hu, Y., Xie, Q., and Chua, N.H.** (2003). The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. *Plant Cell* **15**, 1951-1961.
- Ikeda, Y., Men, S., Fischer, U., Stepanova, A.N., Alonso, J.M., Ljung, K., and Grebe, M.** (2009). Local auxin biosynthesis modulates gradient-directed planar polarity in Arabidopsis. *Nat. Cell Biol.* **11**, 731-738.
- Jia, H., McCarty, D.R., and Suzuki, M.** (2013). Distinct Roles of LAFL Network Genes in Promoting the Embryonic Seedling Fate in the Absence of VAL Repression. *Plant Physiol.*
- Jofuku, K.D., den Boer, B.G., Van Montagu, M., and Okamoto, J.K.** (1994). Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell* **6**, 1211-1225.
- Jürgens, G., and Mayer, U.** (1994). EMBRYOS: Colour atlas of development. (London, UK: Wolfe Publishing).
- Kanei, M., Horiguchi, G., and Tsukaya, H.** (2012). Stable establishment of cotyledon identity during embryogenesis in Arabidopsis by ANGUSTIFOLIA3 and HANABA TARANU. *Development* **139**, 2436-2446.
- Karlberg, A., Bako, L., and Bhalerao, R.P.** (2011). Short day-mediated cessation of growth requires the downregulation of AINTEGUMENTALIKE1 transcription factor in hybrid aspen. *PLoS Genet* **7**, e1002361.
- Kaya, H., Shibahara, K.-i., Taoka, K.-i., Iwabuchi, M., Stillman, B., and Araki, T.** (2001). FASCIATA Genes for Chromatin Assembly Factor-1 in Arabidopsis Maintain the Cellular Organization of Apical Meristems. *Cell* **104**, 131-142.
- Kim, J.H., and Kende, H.** (2004). A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13374-13379.
- Kim, S., Soltis, P.S., Wall, K., and Soltis, D.E.** (2006). Phylogeny and domain evolution in the APETALA2-like gene family. *Mol. Biol. Evol.* **23**, 107-120.
- Klucher, K.M., Chow, H., Reiser, L., and Fischer, R.L.** (1996). The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. *Plant Cell* **8**, 137-153.
- Kornet, N., and Scheres, B.** (2009). Members of the GCN5 histone acetyltransferase complex regulate PLETHORA-mediated root stem cell niche maintenance and transit amplifying cell proliferation in Arabidopsis. *Plant Cell* **21**, 1070-1079.
- Krizek, B.** (2009). AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. *Plant physiology* **150**, 1916-1929.

- Krizek, B.A.** (1999). Ectopic expression of AINTEGUMENTA in Arabidopsis plants results in increased growth of floral organs. *Dev. Genet.* **25**, 224-236.
- Krizek, B.A.** (2003). AINTEGUMENTA utilizes a mode of DNA recognition distinct from that used by proteins containing a single AP2 domain. *Nucleic Acids Res.* **31**, 1859-1868.
- Krizek, B.A., and Eaddy, M.** (2012). AINTEGUMENTA-LIKE6 regulates cellular differentiation in flowers. *Plant Mol. Biol.* **78**, 199-209.
- Krizek, B.A., Prost, V., and Macias, A.** (2000). AINTEGUMENTA promotes petal identity and acts as a negative regulator of AGAMOUS. *Plant Cell* **12**, 1357-1366.
- Krogan, N.T., Hogan, K., and Long, J.A.** (2012). APETALA2 negatively regulates multiple floral organ identity genes in Arabidopsis by recruiting the co-repressor TOPLESS and the histone deacetylase HDA19. *Development* **139**, 4180-4190.
- Kuhlemeier, C.** (2007). Phyllotaxis. *Trends Plant Sci.* **12**, 143-150.
- Kwon, C.S., Chen, C., and Wagner, D.** (2005). WUSCHEL is a primary target for transcriptional regulation by SPLAYED in dynamic control of stem cell fate in Arabidopsis. *Genes Dev.* **19**, 992-1003.
- Laskowski, M., Grieneisen, V.A., Hofhuis, H., Hove, C.A., Hogeweg, P., Maree, A.F., and Scheres, B.** (2008). Root system architecture from coupling cell shape to auxin transport. *PLoS Biol.* **6**, e307.
- Licausi, F., Giorgi, F.M., Zenoni, S., Osti, F., Pezzotti, M., and Perata, P.** (2010). Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. *BMC Genomics* **11**, 719.
- Liu, Z.C., Franks, R.G., and Klink, V.P.** (2000). Regulation of gynoecium marginal tissue formation by LEUNIG and AINTEGUMENTA. *Plant Cell* **12**, 1879-1891.
- Ljung, K., Hull, A.K., Celenza, J., Yamada, M., Estelle, M., Normanly, J., and Sandberg, G.** (2005). Sites and regulation of auxin biosynthesis in Arabidopsis roots. *Plant Cell* **17**, 1090-1104.
- Long, J., and Barton, M.K.** (2000). Initiation of axillary and floral meristems in Arabidopsis. *Dev. Biol.* **218**, 341-353.
- Long, J.A., and Barton, M.K.** (1998). The development of apical embryonic pattern in Arabidopsis. *Development* **125**, 3027-3035.
- Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M.** (2006). TOPLESS regulates apical embryonic fate in Arabidopsis. *Science* **312**, 1520-1523.
- Lorenzo, O., Chico, J.M., Sanchez-Serrano, J.J., and Solano, R.** (2004). JASMONATE-INSENSITIVE1 encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in Arabidopsis. *Plant Cell* **16**, 1938-1950.
- Magnani, E., Sjolander, K., and Hake, S.** (2004). From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell* **16**, 2265-2277.
- Mallory, A.C., Reinhart, B.J., Jones-Rhoades, M.W., Tang, G., Zamore, P.D., Barton, M.K., and Bartel, D.P.** (2004). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J* **23**, 3356-3364.
- Matsuzaki, Y., Ogawa-Ohnishi, M., Mori, A., and Matsubayashi, Y.** (2010). Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. *Science* **329**, 1065-1067.
- Meshorer, E., and Misteli, T.** (2006). Chromatin in pluripotent embryonic stem cells and differentiation. *Nat. Rev. Mol. Cell Biol.* **7**, 540-546.
- Mizukami, Y., and Fischer, R.L.** (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 942-947.
- Mizumoto, K., Hatano, H., Hirabayashi, C., Murai, K., and Takumi, S.** (2009). Altered expression of wheat AINTEGUMENTA homolog, WANT-1, in pistil and pistil-like transformed stamen of an alloplasmic line with *Aegilops crassa* cytoplasm. *Dev. Genes Evol.* **219**, 175-187.

- Mockaitis, K., and Estelle, M.** (2008). Auxin receptors and plant development: a new signaling paradigm. *Annu. Rev. Cell Dev. Biol.* **24**, 55-80.
- Moreno-Risueno, M.A., Van Norman, J.M., Moreno, A., Zhang, J., Ahnert, S.E., and Benfey, P.N.** (2010). Oscillating gene expression determines competence for periodic Arabidopsis root branching. *Science* **329**, 1306-1311.
- Mudunkothge, J.S., and Krizek, B.A.** (2012). Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function. *Plant J.* **71**, 108-121.
- Musselman, C.A., and Kutateladze, T.G.** (2011). Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.* **39**, 9061-9071.
- Nakajima, K., Sena, G., Nawy, T., and Benfey, P.N.** (2001). Intercellular movement of the putative transcription factor SHR in root patterning. *Nature* **413**, 307-311.
- Nambara, E., Suzuki, M., Abrams, S., McCarty, D.R., Kamiya, Y., and McCourt, P.** (2002). A screen for genes that function in abscisic acid signaling in Arabidopsis thaliana. *Genetics* **161**, 1247-1255.
- Nawy, T., Bayer, M., Mravec, J., Friml, J., Birnbaum, K.D., and Lukowitz, W.** (2010). The GATA factor HANABA TARANU is required to position the proembryo boundary in the early Arabidopsis embryo. *Dev. Cell* **19**, 103-113.
- Nole-Wilson, S., and Krizek, B.A.** (2000). DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Res.* **28**, 4076-4082.
- Nole-Wilson, S., and Krizek, B.A.** (2006). AINTEGUMENTA contributes to organ polarity and regulates growth of lateral organs in combination with YABBY genes. *Plant Physiol* **141**, 977-987.
- Nole-Wilson, S., Tranby, T.L., and Krizek, B.A.** (2005). AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant Mol. Biol.* **57**, 613-628.
- Nole-Wilson, S., Azhakanandam, S., and Franks, R.G.** (2010). Polar auxin transport together with AINTEGUMENTA and REVOLUTA coordinate early Arabidopsis gynoecium development. *Dev. Biol.* **346**, 181-195.
- Ohme-Takagi, M., and Shinshi, H.** (1995). Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* **7**, 173-182.
- Okamoto, J.K., Caster, B., Villarreal, R., Van Montagu, M., and Jofuku, K.D.** (1997). The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7076-7081.
- 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.
- Parizot, B., Laplace, L., Ricaud, L., Boucheron-Dubuisson, E., Bayle, V., Bonke, M., De Smet, I., Poethig, S.R., Helariutta, Y., Haseloff, J., Chriqui, D., Beeckman, T., and 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.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M.W., 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.
- 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.
- Pinon, V., Prasad, K., Grigg, S.P., Sanchez-Perez, G.F., and Scheres, B.** (2013). Local auxin biosynthesis regulation by PLETHORA transcription factors controls phyllotaxis in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1107-1112.
- Prasad, K., Grigg, S.P., Barkoulas, M., Yadav, R.K., Sanchez-Perez, G.F., Pinon, V., Blilou, I., Hofhuis, H., Dhonukshe, P., Galinha, C., Mahonen, A.P., Muller, W.H., Raman, S., Verkleij, A.J., Snel, B., Reddy, G.V., Tsiantis, M., and Scheres, B.** (2011). Arabidopsis PLETHORA transcription factors control phyllotaxis. *Curr. Biol.* **21**, 1123-1128.
- Przemeck, G.K., Mattsson, J., Hardtke, C.S., Sung, Z.R., and Berleth, T.** (1996). Studies on the role of the Arabidopsis gene MONOPTEROS in vascular development and plant cell axialization. *Planta* **200**, 229-237.
- Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J., and Xu, Y.** (2012). AP2/ERF Transcription Factor in Rice: Genome-Wide Canvas and Syntenic Relationships between Monocots and Eudicots. *Evol Bioinform Online* **8**, 321-355.
- Rast, M.I., and Simon, R.** (2012). Arabidopsis JAGGED LATERAL ORGANS Acts with ASYMMETRIC LEAVES2 to Coordinate KNOX and PIN Expression in Shoot and Root Meristems. *Plant Cell* **24**, 2917-2933.
- Riechmann, J.L., and Meyerowitz, E.M.** (1998). The AP2/EREBP family of plant transcription factors. *Biol. Chem.* **379**, 633-646.
- Rieu, I., Bots, M., Mariani, C., and Weterings, K.A.** (2005). Isolation and expression analysis of a tobacco AINTEGUMENTA ortholog (NtANTL). *Plant Cell Physiol.* **46**, 803-805.
- Rigal, A., Yordanov, Y.S., Perrone, I., Karlberg, A., Tisserant, E., Bellini, C., Busov, V.B., Martin, F., Kohler, A., Bhalerao, R., and Legue, V.** (2012). The AINTEGUMENTA LIKE1 homeotic transcription factor PtAIL1 controls the formation of adventitious root primordia in poplar. *Plant Physiol* **160**, 1996-2006.
- Rodriguez, R.E., Mecchia, M.A., Debernardi, J.M., Schommer, C., Weigel, D., and Palatnik, J.F.** (2010). Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. *Development* **137**, 103-112.
- 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.
- 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.
- Saiga, S., Furumizu, C., Yokoyama, R., Kurata, T., Sato, S., Kato, T., Tabata, S., Suzuki, M., and Komeda, Y.** (2008). The Arabidopsis OBERON1 and OBERON2 genes encode plant homeodomain finger proteins and are required for apical meristem maintenance. *Development* **135**, 1751-1759.
- 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.
- Schneitz, K., Hulskamp, M., Kopczak, S.D., and Pruitt, R.E.** (1997). Dissection of sexual organ ontogenesis: a genetic analysis of ovule development in Arabidopsis thaliana. *Development* **124**, 1367-1376.
- Schneitz, K., Baker, S.C., Gasser, C.S., and Redweik, A.** (1998). Pattern formation and growth during floral organogenesis: HUELLENLOS and AINTEGUMENTA are required for the formation of the proximal region of the ovule primordium in Arabidopsis thaliana. *Development* **125**, 2555-2563.

- Schultz, E.A., and Haughn, G.W.** (1991). LEAFY, a Homeotic Gene That Regulates Inflorescence Development in Arabidopsis. *Plant Cell* **3**, 771-781.
- Serrano, L., Vazquez, B.N., and Tischfield, J.** (2013). Chromatin structure, pluripotency and differentiation. *Exp. Biol. Med. (Maywood)* **238**, 259-270.
- Sharoni, A.M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., Choi, I.R., Omura, T., and Kikuchi, S.** (2011). Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol.* **52**, 344-360.
- Sheard, L.B., Tan, X., Mao, H., Withers, J., Ben-Nissan, G., Hinds, T.R., Kobayashi, Y., Hsu, F.F., Sharon, M., Browse, J., He, S.Y., Rizo, J., Howe, G.A., and Zheng, N.** (2010). Jasmonate perception by inositol-phosphate-potentiated CO11-JAZ co-receptor. *Nature* **468**, 400-405.
- Shigyo, M., Hasebe, M., and Ito, M.** (2006). Molecular evolution of the AP2 subfamily. *Gene* **366**, 256-265.
- Shoji, T., Mishima, M., and Hashimoto, T.** (2013). Divergent DNA-binding specificities of a group of ETHYLENE RESPONSE FACTOR transcription factors involved in plant defense. *Plant Physiol* **162**, 977-990.
- Shu, K., Zhang, H., Wang, S., Chen, M., Wu, Y., Tang, S., Liu, C., Feng, Y., Cao, X., and Xie, Q.** (2013). ABI4 Regulates Primary Seed Dormancy by Regulating the Biogenesis of Abscisic Acid and Gibberellins in Arabidopsis. *PLoS Genet* **9**, e1003577.
- Siegfried, K.R., Eshed, Y., Baum, S.F., Otsuga, D., Drews, G.N., and Bowman, J.L.** (1999). Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. *Development* **126**, 4117-4128.
- Smith, Z.R., and Long, J.A.** (2010). Control of Arabidopsis apical-basal embryo polarity by antagonistic transcription factors. *Nature* **464**, 423-426.
- Stepanova, A.N., Hoyt, J.M., Hamilton, A.A., and Alonso, J.M.** (2005). A Link between ethylene and auxin uncovered by the characterization of two root-specific ethylene-insensitive mutants in Arabidopsis. *Plant Cell* **17**, 2230-2242.
- Stepanova, A.N., Robertson-Hoyt, J., Yun, J., Benavente, L.M., Xie, D.Y., Dolezal, K., Schlereth, A., Jurgens, G., and Alonso, J.M.** (2008). TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell* **133**, 177-191.
- Sundaram, S., Kertbundit, S., Shakirov, E.V., Iyer, L.M., Juricek, M., and Hall, T.C.** (2013). Gene networks and chromatin and transcriptional regulation of the phaseolin promoter in Arabidopsis. *Plant Cell* **25**, 2601-2617.
- Szakonyi, D., Moschopoulos, A., and Byrne, M.E.** (2010). Perspectives on leaf dorsoventral polarity. *J. Plant Res.* **123**, 281-290.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S.** (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* **28**, 2731-2739.
- Taylor, G.K., and Stoddard, B.L.** (2012). Structural, functional and evolutionary relationships between homing endonucleases and proteins from their host organisms. *Nucleic Acids Res.* **40**, 5189-5200.
- Thomas, C.L., Schmidt, D., Bayer, E.M., Dreos, R., and Maule, A.J.** (2009). Arabidopsis plant homeodomain finger proteins operate downstream of auxin accumulation in specifying the vasculature and primary root meristem. *Plant J.* **59**, 426-436.
- 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.
- van den Berg, C., Willemssen, V., Hendriks, G., Weisbeek, P., and Scheres, B.** (1997). Short-range control of cell differentiation in the Arabidopsis root meristem. *Nature* **390**, 287-289.

- Vanneste, S., De Rybel, B., Beemster, G.T., Ljung, K., De Smet, I., Van Isterdael, G., Naudts, M., Iida, R., Gruissem, W., Tasaka, M., Inze, D., Fukaki, H., and Beeckman, T. (2005). Cell cycle progression in the pericycle is not sufficient for SOLITARY ROOT/IAA14-mediated lateral root initiation in *Arabidopsis thaliana*. *Plant Cell* **17**, 3035-3050.
- Wang, L., Gu, X., Xu, D., Wang, W., Wang, H., Zeng, M., Chang, Z., Huang, H., and Cui, X. (2011). miR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in *Arabidopsis*. *J. Exp. Bot.* **62**, 761-773.
- Weigel, D., and Meyerowitz, E.M. (1993). Activation of floral homeotic genes in *Arabidopsis*. *Science* **261**, 1723-1726.
- Weigel, D., and Nilsson, O. (1995). A developmental switch sufficient for flower initiation in diverse plants. *Nature* **377**, 495-500.
- Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F., and Meyerowitz, E.M. (1992). LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* **69**, 843-859.
- Wildwater, M., Campilho, A., Perez-Perez, J.M., Heidstra, R., Blilou, I., Korthout, H., Chatterjee, J., Mariconti, L., Gruissem, W., and Scheres, B. (2005). The RETINOBLASTOMA-RELATED gene regulates stem cell maintenance in *Arabidopsis* roots. *Cell* **123**, 1337-1349.
- Wilmoth, J.C., Wang, S., Tiwari, S.B., Joshi, A.D., Hagen, G., Guilfoyle, T.J., Alonso, J.M., Ecker, J.R., and Reed, J.W. (2005). NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation. *Plant J.* **43**, 118-130.
- Wisniewska, J., Xu, J., Seifertova, D., Brewer, P.B., Ruzicka, K., Blilou, I., Rouquie, D., Benkova, E., Scheres, B., and Friml, J. (2006). Polar PIN localization directs auxin flow in plants. *Science* **312**, 883.
- Won, C., Shen, X.L., Mashiguchi, K., Zheng, Z.Y., Dai, X.H., Cheng, Y.F., Kasahara, H., Kamiya, Y., Chory, J., and Zhao, Y.D. (2011). Conversion of tryptophan to indole-3-acetic acid by TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS and YUCCAs in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 18518-18523.
- Wuitschick, J.D., Lindstrom, P.R., Meyer, A.E., and Karrer, K.M. (2004). Homing endonucleases encoded by germ line-limited genes in *Tetrahymena thermophila* have APETELA2 DNA binding domains. *Eukaryot. Cell* **3**, 685-694.
- Yamada, T., Hirayama, Y., Imaichi, R., and Kato, M. (2008). AINTEGUMENTA homolog expression in *Gnetum* (gymnosperms) and implications for the evolution of ovulate axes in seed plants. *Evol Dev* **10**, 280-287.
- Yamagishi, K., Tatematsu, K., Yano, R., Preston, J., Kitamura, S., Takahashi, H., McCourt, P., Kamiya, Y., and Nambara, E. (2009). CHOTTO1, a double AP2 domain protein of *Arabidopsis thaliana*, regulates germination and seedling growth under excess supply of glucose and nitrate. *Plant Cell Physiol.* **50**, 330-340.
- 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.
- Yan, J., Zhang, C., Gu, M., Bai, Z., Zhang, W., Qi, T., Cheng, Z., Peng, W., Luo, H., Nan, F., Wang, Z., and Xie, D. (2009). The *Arabidopsis* CORONATINE INSENSITIVE1 protein is a jasmonate receptor. *Plant Cell* **21**, 2220-2236.
- Yano, R., Kanno, Y., Jikumaru, Y., Nakabayashi, K., Kamiya, Y., and Nambara, E. (2009). CHOTTO1, a putative double APETALA2 repeat transcription factor, is involved in abscisic acid-mediated repression of gibberellin biosynthesis during seed germination in *Arabidopsis*. *Plant physiology* **151**, 641-654.

- Zhao, Y., Medrano, L., Ohashi, K., Fletcher, J.C., Yu, H., Sakai, H., and Meyerowitz, E.M.** (2004). HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in Arabidopsis. *Plant Cell* **16**, 2586-2600.
- Zhou, W., Wei, L., Xu, J., Zhai, Q., Jiang, H., Chen, R., Chen, Q., Sun, J., Chu, J., Zhu, L., Liu, C.M., and Li, C.** (2010). Arabidopsis Tyrosylprotein sulfotransferase acts in the auxin/PLETHORA pathway in regulating postembryonic maintenance of the root stem cell niche. *Plant Cell* **22**, 3692-3709.
- Zhuang, J., Peng, R.H., Cheng, Z.M., Zhang, J., Cai, B., Zhang, Z., Gao, F., Zhu, B., Fu, X.Y., Jin, X.F., Chen, J.M., Qiao, Y.S., Xiong, A.S., and Yao, Q.H.** (2009). Genome-wide analysis of the putative AP2/ERF family genes in *Vitis vinifera*. *Scientia Horticulturae* **123**, 73-81.

Chapter 3

AIL and HDG proteins act antagonistically to control cell proliferation

Anneke Horstman¹, Hiroyuki Fukuoka², Jose M. Muino³, Lisette Nitsch⁴,
Changhua Guo¹, Paul Passarinho^{1,†}, Gabino Sanchez-Perez^{1,6}, Richard Immink¹,
Gerco Angenent^{1,5}, Kim Boutilier^{1*}

¹ Wageningen University and Research Centre, Bioscience, Droevendaalsesteeg 1, 6708 PB
Wageningen, The Netherlands

² NARO Institute of Vegetable and Tea Science (NIVTS), 360 Kusawa, Ano, Tsu, Mie 514-2392,
Japan

³ Max Planck Institute for Molecular Genetics, D-14195 Berlin, Germany

⁴ Laboratory of Biochemistry, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

⁵ Laboratory of Molecular Biology, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

⁶ Laboratory of Bioinformatics, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands

[†] Current address: Genetwister Technologies B.V., Nieuwe Kanaal 7b, 6709 PA, Wageningen, The
Netherlands

Abstract

AINTEGUMENTA-LIKE (AIL) transcription factors are key regulators of cell proliferation and meristem identity. Although AIL functions have been well described, the direct signalling components of this pathway are largely unknown. We show that BABY BOOM (BBM) and other AIL proteins physically interact with multiple members of the L1-expressed HOMEODOMAIN GLABROUS (HDG) transcription factor family, including HDG1, HDG11 and HDG12. Overexpression of *HDG1*, *HDG11* and *HDG12* restricts growth due to root and shoot meristem arrest, which is associated with reduced expression of genes involved in meristem development and cell proliferation pathways, while down-regulation of multiple *HDG* genes promotes cell overproliferation. These results suggest a role for HDG proteins in promoting cell differentiation. We also reveal a transcriptional network in which BBM and HDG1 regulate several common targets genes, and where BBM/AIL and HDG regulate each other's expression. Taken together, these results suggest opposite roles for AIL and HDG proteins, with AILs promoting cell proliferation and HDGs stimulating cell differentiation, and that these functions are mediated at both the protein-protein interaction and transcriptional level.

INTRODUCTION

Plant growth is driven by stem cells within the meristems, which are maintained throughout the plant's lifespan to ensure continued growth. At the same time, stem cell proliferation has to be kept in balance and contained within the meristem to prevent neoplastic growth. The AINTEGUMENTA-LIKE (AIL) subfamily of the APETALA2/ETHYLENE RESPONSE ELEMENT-BINDING FACTOR (AP2/ERF) family of transcription factors play an important role in defining the meristematic competence of plant cells (Horstman et al., 2014a). The *Arabidopsis thaliana* (L.) Heyhn (Arabidopsis) AIL clade comprises AINTEGUMENTA (ANT), AIL1, PLETHORA1 (PLT1), PLT2, AIL6/PLT3, PLT7, BABY BOOM (BBM) and AIL5/PLT5 (Nole-Wilson et al., 2005). Arabidopsis *AIL* genes are expressed in the embryo, the flower and the root and shoot meristems, where they act redundantly to define and/or maintain the stem cell niches (Galinha et al., 2007; Mudunkothge and Krizek, 2012; Horstman et al., 2014a). Although mutations in single *AIL* genes only lead to minor developmental defects, double and triple *ail* mutants exhibit more severe phenotypes, such as rootlessness (*plt1;plt2;plt3*), embryo lethality (*bbm;plt2*) (Galinha et al., 2007) or shoot meristem arrest (*ant;ail6/plt3;plt7*) (Mudunkothge and Krizek, 2012). The overexpression phenotypes of AIL proteins also support the notion of a role for these proteins in promoting meristematic competence. Overexpression of *Brassica napus* *BBM* or Arabidopsis *AIL5/PLT5* induces formation of somatic embryos (Boutilier et al., 2002; Tsuwamoto et al., 2010), while overexpression of *PLT1* and *PLT2* induces ectopic root identity (Aida et al., 2004). In addition, the increased cell divisions due to *AIL* overexpression can also lead to increased floral organ size, as shown for both *AIL5/PLT5* and *ANT* overexpression (Krizek, 1999; Nole-Wilson et al., 2005).

Although it is clear that AILs are key regulators of meristem function and cell proliferation, how *AIL* overexpression can trigger ectopic organ formation or embryogenesis is poorly understood. Transcription factor function is mediated in the context of multi-protein complexes. To provide insight into the mode of action of *BBM* and the signalling network in which it functions during cell proliferation, we identified and characterized *BBM*-interacting proteins. Here, we show that *BBM* and other AIL proteins interact with members of the HD-ZIP class IV/HOMEODOMAIN GLABROUS (HDG) transcription factor family. Sixteen HD-ZIP IV proteins have been identified in Arabidopsis, including MERISTEM LAYER1 (ATML1), GLABRA2 (GL2), ANTHOCYANINLESS2 (ANL2) and PROTODERMAL FACTOR2 (PDF2) (Nakamura et al., 2006). *HD-ZIP IV/HDG* genes are expressed in the L1 layer throughout the plant, where they function to

specify epidermis identity and control development of its associated structures, such as trichomes, stomata or giant cells (Abe et al., 2003; Nakamura et al., 2006; Roeder et al., 2012; Peterson et al., 2013; Takada et al., 2013). Our results show that BBM and HDG proteins have antagonistic functions, with BBM stimulating cell proliferation and HDGs stimulating cell differentiation. In addition, we found evidence for transcriptional cross-regulation between *BBM/AIL* and *HDG* genes, suggesting a complex regulatory network for cell proliferation control.

RESULTS

BBM interacts with HDG proteins

We employed the yeast two-hybrid system to identify BBM-interacting proteins. Due to the strong and extensive transcriptional autoactivation activity of *B. napus* BBM1 (supplementary material Figs S1, S2, Table S1), we used the CytoTrap system (Aronheim, 1997) to screen a library made from *B. napus* embryos for interactions with BBM. We identified 10 HDG transcription factors as interacting partners (supplementary material Table S2), nine of which are most similar to Arabidopsis HDG11 and one to HDG1 (supplementary material Table S2). Our subsequent studies focussed on the Arabidopsis BBM and HDG orthologs.

The interaction between BBM and HDG1 and HDG11 proteins was verified *in planta* using Förster Resonance Energy Transfer detected via Fluorescence Lifetime Imaging Microscopy (FRET-FLIM; Fig. 1A) (Bucherl et al., 2014), as was the interaction with HDG12, which functions redundantly with HDG11 in trichome development (Nakamura et al., 2006). These results indicate that BBM also interacts with HDG proteins *in planta*. Using the Gal4 yeast two-hybrid system we found that BBM also interacts with HDG2, HDG3, HDG10, ANL2, ATML1 and PDF2 (Fig. 1B). Next, we determined whether these interactions also extend beyond BBM, by testing for HDG interactions with AIL proteins from the two major AIL clades, the ANT clade (ANT and AIL1) and the BBM/PLT clade (PLT7) (Horstman et al., 2014a). All three AIL proteins interacted with multiple HDG proteins (Fig. 1B). Our results show that BBM and other AIL proteins interact with phylogenetically distinct members of the HDG family (Nakamura et al., 2006).

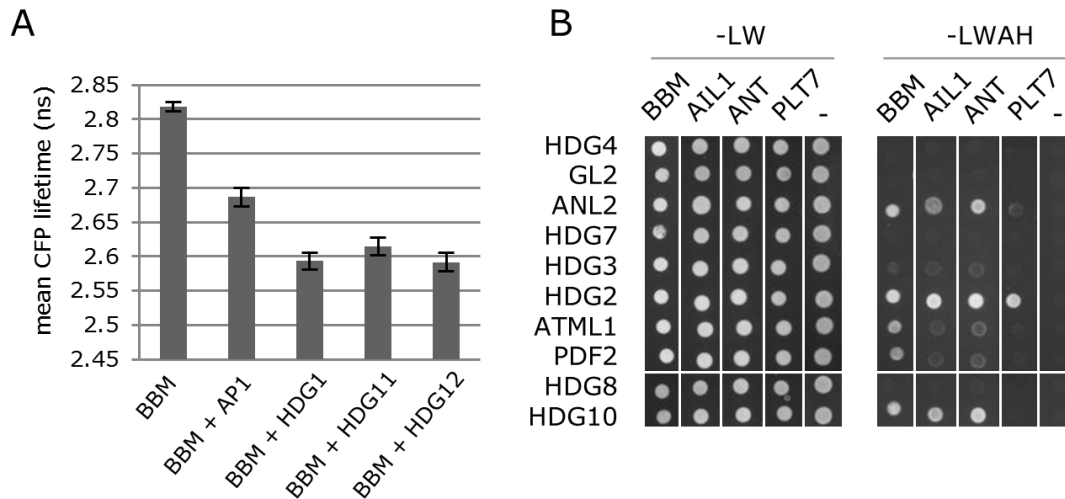


Fig. 1. AIL and HDG proteins interact.

(A) FRET-FLIM measurements with BBM-CFP, YFP-HDG and YFP-AP1 as a negative control. A statistically-significant decrease of BBM-CFP lifetime ($p < 0.05$, Student's t -test) was observed when BBM-CFP was co-expressed with all YFP-HDGs, compared to YFP-AP1. Error bars indicate standard deviations ($n = 26-30$).

(B) Yeast two-hybrid analysis of AD-AIL and BD-HDG interactions. Yeast were grown on double selection synthetic dropout medium (SD-LW) to determine the success of the mating, and tested for protein-protein interaction on SD-LWAH medium. "-", empty-AD vector control.

BBM and HDG expression patterns overlap

Previous studies showed that *BBM*, *HDG1*, *HDG11* and *HDG12* are expressed in embryos and roots (Boutillier et al., 2002; Nakamura et al., 2006; Galinha et al., 2007). We examined their expression patterns in more detail using translational GFP fusion reporters. *BBM* expression was observed throughout the embryo from the 4-celled stage until the globular stage (Fig. 2A, E, I) and became basally localized at the heart stage, as previously reported (Fig. 2M) (Galinha et al., 2007). *HDG11* and *HDG12* expression was observed in all cells at the 4- and 16-cell embryo stages (Fig. 2C, D, G, H), and became restricted to the protoderm from the globular stage onward (Fig. 2K, L, O, P). *HDG1* was weakly expressed, and its expression was first observed in the embryo protoderm starting at the late globular stage (Fig. 2J).

BBM was expressed in the stem cell niche and the provascular tissue of mature roots (Fig. 2Q). Expression of all three *HDG* genes was observed in the epidermis, the outer layer of columella cells and lateral root cap of mature primary and lateral roots (Fig. 2R, S, T). These *HDG* genes were also expressed in the L1 layer of the floral meristems (Fig. 2U, V, W), shoot apical meristem (SAM) and leaf primordia (Fig. 2X, Y, Z). *HDG1* expression was also observed in the subepidermal layers of the flower meristem, SAM, and leaf primordia (Fig. 2U, X).

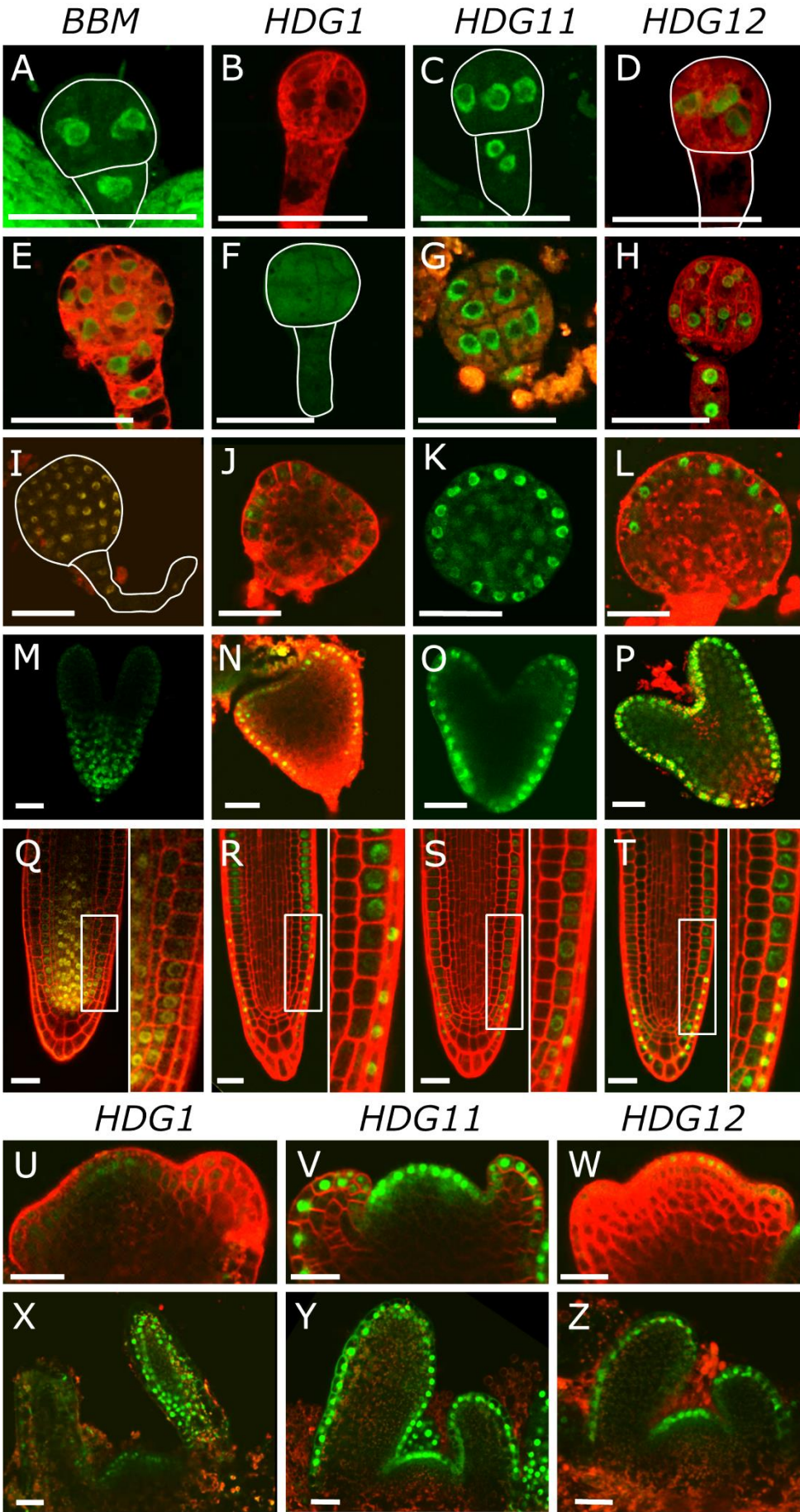


Fig. 2. Expression patterns of *BBM*, *HDG1*, *HDG11* and *HDG12* overlap in embryos and roots.

Confocal microscope images of *BBM* and *HDG* reporter lines in 4-cell (**A-D**), 16-cell (**E-H**), globular (**I-L**) and heart/torpedo embryos (**M-P**), and in roots (**Q-T**), floral meristems (**U-W**) and shoot apical meristems (**X-Z**) expressing YFP (*BBM*) or GFP (*HDG*). The insets in (**Q-T**) show magnifications of the region of root epidermis indicated by the boxes. Scale bars, 25 μ m.

BBM expression was not observed in the shoot or flower. In summary, *BBM* expression overlapped with *HDG11* and *HDG12* expression during early embryo development, and later with all three *HDG* genes in progressively smaller regions of the protoderm. Post-embryonically, there was only a small overlap in expression of *BBM* and the three *HDG* genes, in a few epidermal cells close to the root stem cell niche.

Overexpression of *HDG* genes induces meristem arrest and leaf defects

To determine the functions of HDG proteins, we generated Arabidopsis *HDG1*, *HDG11* and *HDG12* overexpression lines (*p35S::HDG*). Approximately 10% of the primary transformants ($n \geq 200$ per construct) showed (similar) mutant phenotypes, with *HDG1* resulting in the most severely altered phenotypes. Most of the affected seedlings were small and showed increased anthocyanin production (Fig. 3B). Seedlings of the most severe lines had a short primary root lacking lateral roots and stopped growing after producing a few leaves that were narrow and curled upward (Fig. 3B). Leaf fusions and leaves with holes were occasionally observed (Fig. 3C, D). The majority of seedlings with these phenotypes either died or was sterile, complicating further analysis of the lines. Therefore, we created dexamethasone (DEX)-inducible GR-*HDG1*, GR-*HDG11* and GR-*HDG12* proteins (*p35S::GR-HDG*) and selected the primary transformants directly on DEX-containing medium. We observed the same mutant phenotypes as described above. Again, the *p35S::GR-HDG1* ($n=468$) mutant phenotypes were most pronounced, and in some cases more severe than the *p35S::HDG1* mutant phenotypes. The most severely affected *p35S::GR-HDG1* seedlings developed very narrow, gutter-shaped or radialized leaves (Fig. 3E-G, inset K), with occasional leaf ruptures (Fig. 3H). No aberrant phenotypes were observed when *p35S::GR-HDG1* seedlings were grown without DEX (Fig. 3E, left) and overexpression of the *HDG1* transgene was confirmed (supplementary material Fig. S3), indicating that the observed phenotypes were due to ectopic expression of HDG1.

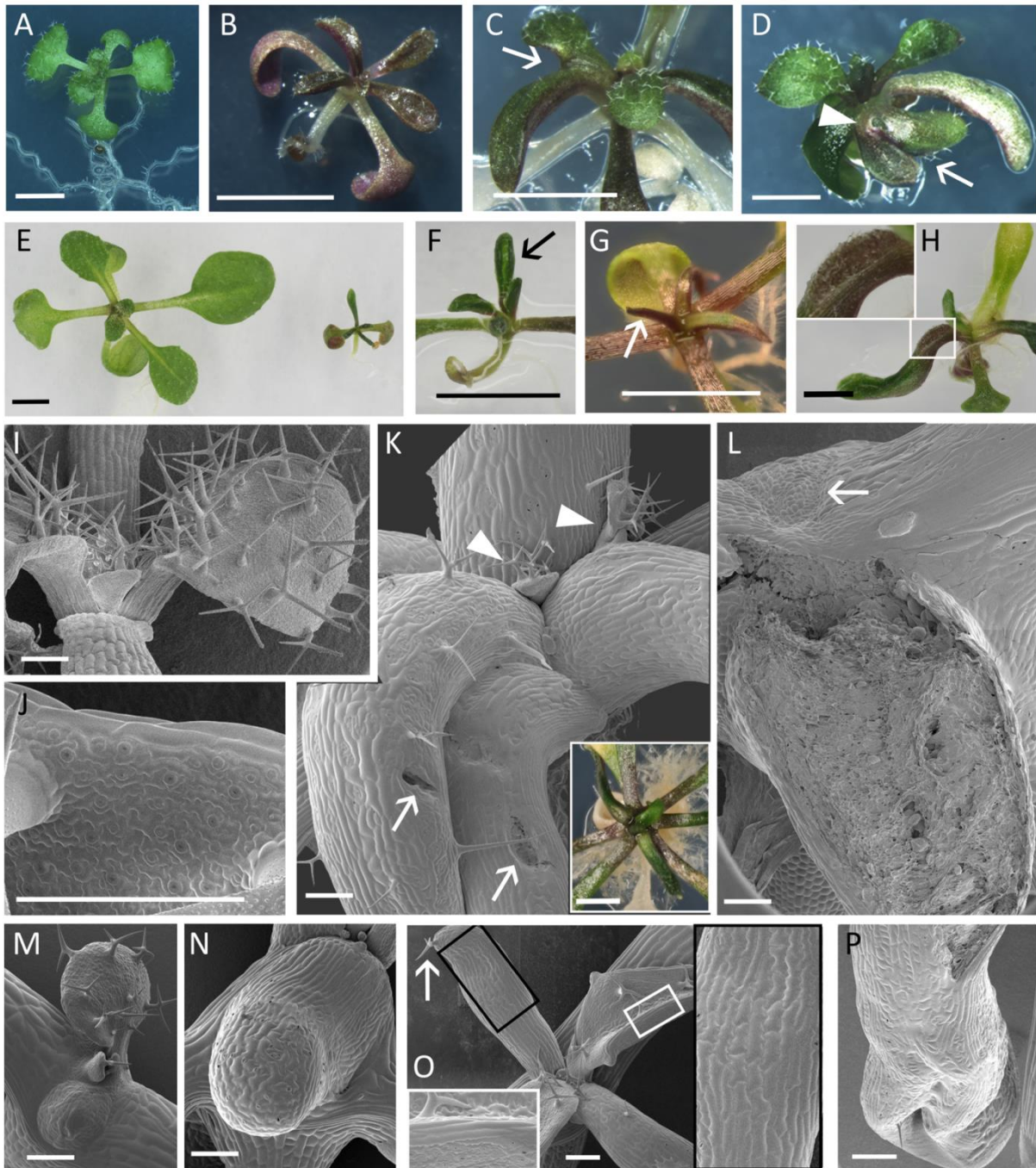


Fig. 3. *HDG1* overexpression affects leaf and shoot meristem development

(A) Eight-day-old wild-type Col-0 seedling.

(B) Six-week-old *p35S::HDG1* seedling showing narrow, upward-curling leaves and anthocyanin production.

(C, D) Three-week-old *p35S::HDG1* seedlings with leaf fusions (arrows) and an epidermal hole (arrowhead).

(E) Two-week-old *p35S::GR-HDG1* seedling grown on DEX (right) showing narrow leaves and retarded growth compared to a two-week-old seedling grown on medium lacking DEX (left).

(F-H) Two- (F) and three- (G, H) week-old DEX-induced *p35S::GR-HDG1* seedlings showing gutter-shaped leaves (F, arrow), pin-shaped leaves (G, arrow) and a leaf rupture (H). The inset in (H) shows a magnification of the region indicated by the box.

(I-P) SEM images. (I) The shoot apex of a wild-type seedling. (J) A wild-type adaxial leaf surface. (K) A *p35S::GR-HDG1* seedling with radialized leaves, as in the inset (arrows, epidermal holes; arrowheads, adventitious leaves). (L) A *p35S::GR-HDG1* seedling with an arrested shoot meristem (arrow) and a ruptured leaf epidermis.

(M, N) *p35S::GR-HDG1* seedlings with arrested leaf growth. (O) A *p35S::HDG1* seedling with a radialized leaf carrying a single trichome at the distal end (arrow), as well as leaves with a clear adaxial/abaxial identity. The insets show magnifications of the leaf areas indicated by the boxes. (P) The gutter-shaped distal region of the ruptured leaf shown in (L). The seedlings shown in (I-J) and (K-P) are five and 14 days old, respectively. Scale bars in light images, 5 mm; in SEM images, 200 μ m.

The phenotypes of *p35S::HDG1* and DEX-induced *p35S::GR-HDG1* seedlings were examined in more detail using scanning electron microscopy (SEM). SEM analysis showed that the leaves of some *HDG1* overexpression lines were radialized, but without an obvious abaxial or adaxial identity (Fig. 3N-P). The leaf surface comprised large numbers of smooth, elongated cells (Fig. 3O, P), reminiscent of the cells found on sepals and leaf margins (Fig. 3J) (Roeder et al., 2012), rather than the puzzle-shaped cells found in wild-type leaves (Fig. 3J). Some of the *HDG1* overexpression seedlings developed a first set of radialized leaves, but did not grow further (Fig. 3M, N), while other *HDG1* overexpression seedlings developed a radialized first leaf pair and a second leaf pair (Fig. 3O) with normal adaxial/abaxial patterning, although with a larger number of elongated cells (Fig. 3O). In addition, we observed holes (Fig. 3K) and large ruptures in the leaf epidermis (Fig. 3L). The altered leaf shape of the *HDG1* overexpression seedlings complicated a general comparison of epidermal characteristics in these lines with those of wild-type seedlings (Fig. 3J). We did not observe any changes in trichome morphology, however the radialized leaves of *HDG1* overexpression seedlings contained less trichomes (Fig. 3M-P), and these were often positioned on the distal end of the leaf (Fig. 3O).

The shoot and root meristem were also affected by ectopic *HDG1* overexpression. In the most severe cases, the shoot meristem was absent (Fig. 3L). Small leaves were observed occasionally in meristem-arrested seedlings (Fig. 3K). These were visible at a later stage, when the surrounding leaves were fully developed. It was not clear whether these leaves developed from axillary meristems or through adventitious growth, but they never developed further. When these seedlings were transferred to medium lacking DEX prior to complete meristem arrest, they recovered and developed into wild-type looking seedlings.

The loss of root meristem function due to *HDG1* overexpression was confirmed by the reduced growth rate and shortened root meristem of DEX-induced *p35S::GR-HDG1* seedlings compared to wild-type seedlings (Fig. 4). Despite these root growth defects, DEX-induced *p35S::GR-HDG1* seedlings were able to produce lateral roots (Fig. 4A). The root meristem defect observed here resembles the phenotype of *bbm;plt* loss-of-function mutants (Galinha et al., 2007), suggesting opposite roles for BBM and HDG1 proteins, with BBM promoting root meristem activity and HDG1 stimulating meristem differentiation.

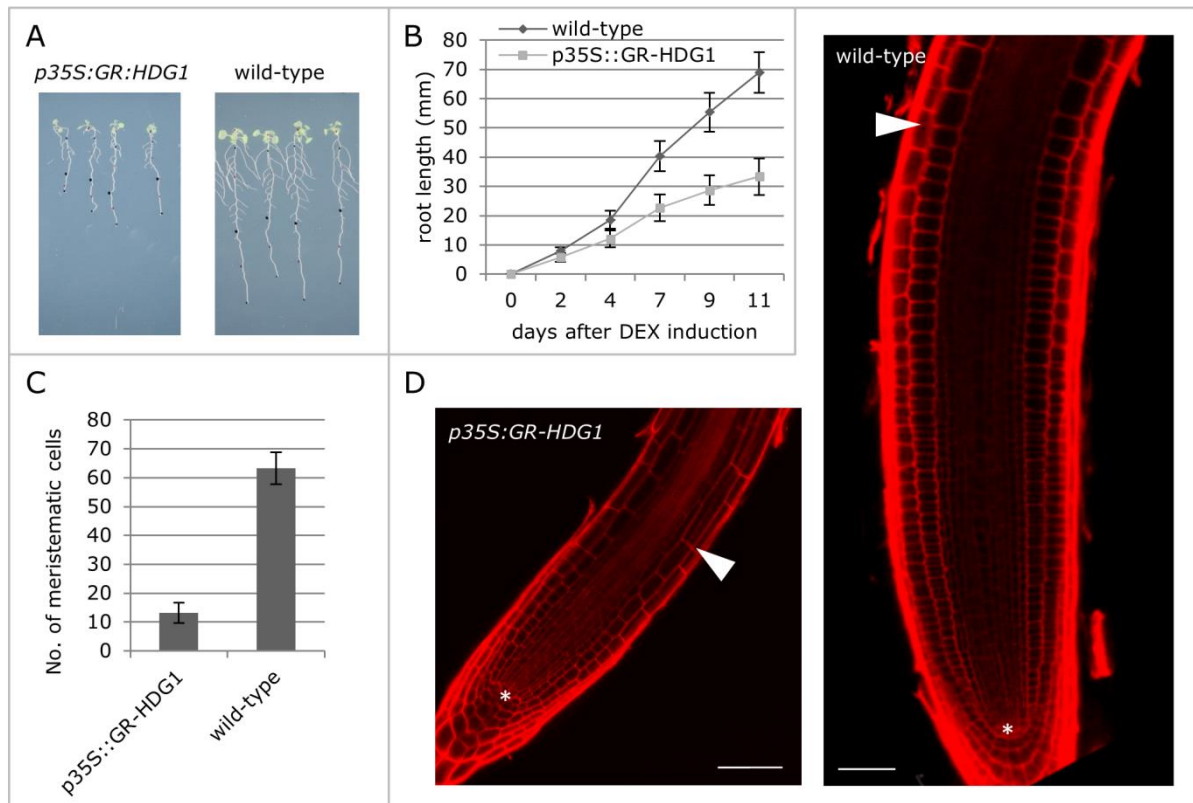


Fig. 4. *HDG1* overexpression reduces root length and meristem size

(A) 15 day-old *p35S::GR-HDG1* and wild-type Col-0 seedlings after transfer to DEX-containing medium, five days after plating. The root tips were marked to facilitate growth rate calculations.

(B) Root growth rate of the DEX-treated *p35S::GR-HDG1* and wild-type Col-0 seedlings shown in (A). Root length was measured starting 5 days after plating and transfer to DEX. Error bars indicate standard deviations ($n=15$).

(C) The number of meristematic cortex cells (from the quiescent centre to the first elongated cell) in *p35S::GR-HDG1* and wild-type Col-0 seedlings grown continuously on DEX for 10 days. Error bars indicate standard deviations ($n=22$ for wild-type, $n=33$ for *p35S::GR-HDG1*).

(D) Confocal microscope images showing roots of *p35S::GR-HDG1* and wild-type Col-0 seedlings that were grown continuously on DEX for 10 days. Asterisks indicate the quiescent centre, arrowheads indicate the first elongated cortex cell. Scale bars, 50 μm .

***HDG1* overexpression promotes giant cell identity**

The elongated cells found in *HDG1* overexpression lines are reminiscent of giant cells, which are differentiated, endoreduplicated cells found in the sepal epidermis (Roeder et al., 2012). Similarly elongated cells are found along the margin of cotyledons and leaves (Fig. 5A, B), and in the root. We used the enhancer trap line YJ158, which reports GUS activity in giant/elongated cells (Eshed et al., 2004; Roeder et al., 2012) (Fig. 5C) to determine whether *HDG1* overexpression seedlings show enhanced giant/leaf margin cell production.

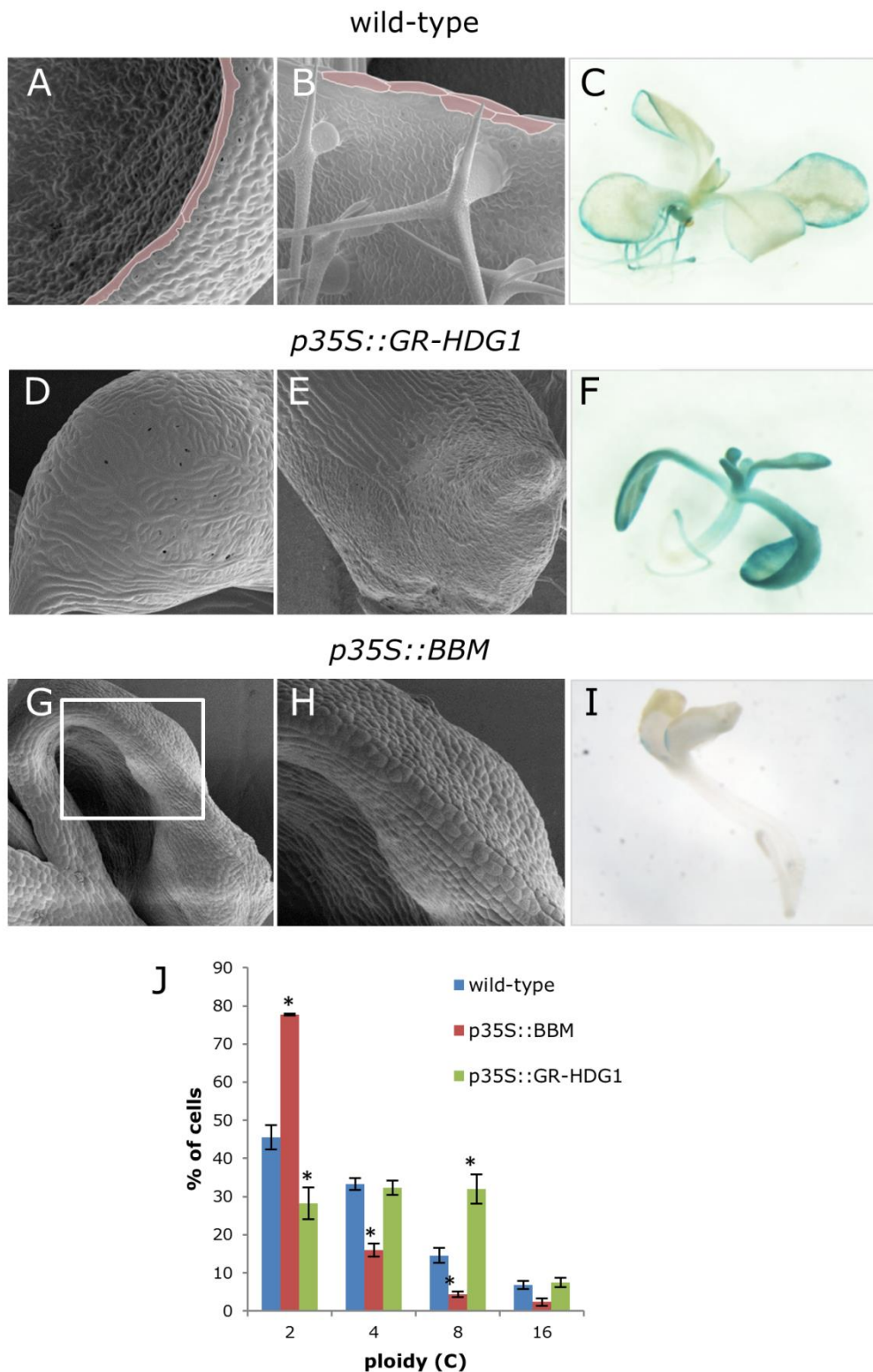


Fig. 5. *HDG1* overexpression promotes giant cell identity and differentiation

(A, B, D, E, G, H) SEM images of two-week-old seedlings.

(C, F, I) YJ158 expression in eight-day-old seedlings.

(A, B) Wild-type Col-0 cotyledon (A) and leaf (B) with elongated margin cells (false-coloured pink).

(D, E) Aberrantly-shaped leaves from DEX-treated *p35S::GR-HDG1* seedlings showing a large number of elongated cells.

(G, H) A *p35S::BBM* cotyledon (G), consisting of many small cells and lacking clear adaxial/abaxial cell types and (H) magnification of the cotyledon region indicated by the box in (G).

(C, F, I) YJ158 expression in a wild-type seedling (C), a DEX-treated *p35S::GR-HDG1* seedling with radialized leaves (F) and a *p35S::BBM* seedling (I).

(J) Ploidy analysis of wild-type (blue), *p35S::BBM* (red) and DEX-treated *p35S::GR-HDG1* 10 day-old seedlings (green). The error bars indicate standard errors and the asterisks significant differences ($p < 0.05$, Student's *t*-test) compared to wild-type seedlings.

In DEX-induced *p35S::GR-HDG1* seedlings with the narrow leaf phenotype (Fig. 3B), GUS staining was more intense, but still restricted to the margins, as in the control (data not shown), while *p35S::GR-HDG1* seedlings with gutter and/or pin-shaped leaves (Fig. 3G, Fig. 5D, E) showed GUS expression throughout the leaf surface, and also in the cotyledon blade and petiole (Fig. 5F). In contrast with *HDG1* overexpression lines, cotyledons of *p35S::BBM* seedlings consist of small, undifferentiated cells (Fig. 5G, H): the cotyledons lacked the jigsaw-shaped cells, stomata and the elongated margin cells of wild-type leaves (Fig. 5A, B). As expected based on cell morphology, YJ158 marker expression was weak (Fig. 5I) or completely absent in *p35S::BBM* seedlings.

Sepal giant cells are highly endoreduplicated, a differentiation process that occurs after the establishment of giant cell identity (Roeder et al., 2012). Endoreduplication is also considered a sign of advanced cell differentiation in leaf pavement cells (Melaragno et al., 1993). In accordance with our observations on changes in epidermal differentiation, we observed a shift toward higher ploidy levels in *p35S::GR-HDG1* seedlings and toward lower ploidy levels in *p35S::BBM* seedlings (Fig. 5J). Taken together, our results suggest that HDG1 and BBM promote and inhibit epidermal cell differentiation, respectively.

***hdg* mutants do not show obvious embryo or root meristem phenotypes**

To further investigate the roles of HDG proteins during development, we examined *hdg1*, *hdg11* and *hdg12* mutant phenotypes during embryogenesis and root development, the developmental stages in which BBM functions. However, none of the single, double or triple *hdg1*, *hdg11* or *hdg12* lines showed mutant phenotypes in these tissues (supplementary material Fig. S4). The lack of embryo and root phenotypes may reflect redundancy between members of the HDG family (Nakamura et al., 2006).

Cosuppression of *HDG* expression results in overproliferation

We observed that approximately 1% of *p35S::HDG* primary transformants formed ectopic shoots (Fig. 6A-B) or embryo-like tissue. None of the transformants with these phenotypes showed the overexpression phenotypes described above. Notably, these proliferating seedlings resemble *BBM* overexpression seedlings, which also show ectopic organ formation and somatic embryogenesis (SE) (Boutilier et al., 2002). Cell proliferation phenotypes were also observed at a similar frequency in *p35S::GR-HDG1* primary transformants that were grown on DEX-containing medium, but that failed to recover the wild-type phenotype after transfer to DEX-free medium (Fig. 6C-E). These data, together with the lack of similar phenotypes in *hdg1* T-DNA insertion mutants, suggest that cosuppression of multiple *HDG* genes underlies this cell proliferation phenotype.

We could detect *GR-HDG1* transgene expression in the *p35S::GR-HDG1* plants that recovered their wild-type phenotype after transfer to DEX-free medium, but could not detect *GR-HDG1* expression in the plants that failed to recover, i.e. continued to overproliferate, in the absence of DEX (supplementary material Fig. S5). Expression of seven selected *HDG* genes (*ATML1*, *PDF2*, *HDG2*, *HDG11*, *HDG12*, *ANL2*, *HDG3*) was also reduced in these proliferating lines compared to wild-type seedlings (supplementary material Fig. S5). However, endogenous *HDG1* expression was variably up- or down-regulated in these lines depending on the quantitative real-time RT-PCR (qPCR) primer set that was used (supplementary material Fig. S5). Similar variable qPCR results were previously reported for silenced genes and could be caused by incomplete degradation of mRNA fragments (Shepard et al., 2005; Holmes et al., 2010). Expression analysis in subsequent generations was further complicated by the limited survival and fertility of these seedlings, and by the reversion of the surviving lines to the characteristic DEX-dependent overexpression phenotypes in the subsequent generation. Although indirect, these results imply that the cell proliferation phenotypes observed in a subset of the *p35S::GR-HDG1* plants are not due to ectopic *HDG1* overexpression, but rather to *HDG* gene silencing.

To determine whether down-regulation of multiple *HDG* genes could cause these overproliferation phenotypes, we developed an artificial microRNA (amiRNA) construct (Schwab et al., 2006) that is predicted to target *HDG3*, *HDG7*, *HDG11*, *PDF2* and *ATML1*. We observed the same cell proliferation phenotypes in four primary transformants after transformation of this amiRNA construct to the *hdg1;anl2* double mutant (2/119 lines; Fig. 6F, G), and wild-type Col-0 (2/467 lines; Fig. 6H, I). These *HDG* amiRNA lines could not be propagated via seed. Together,

the *GR-HDG1* and amiRNA data provide support to the hypothesis that down-regulation of multiple *HDG* genes leads to ectopic cell proliferation.

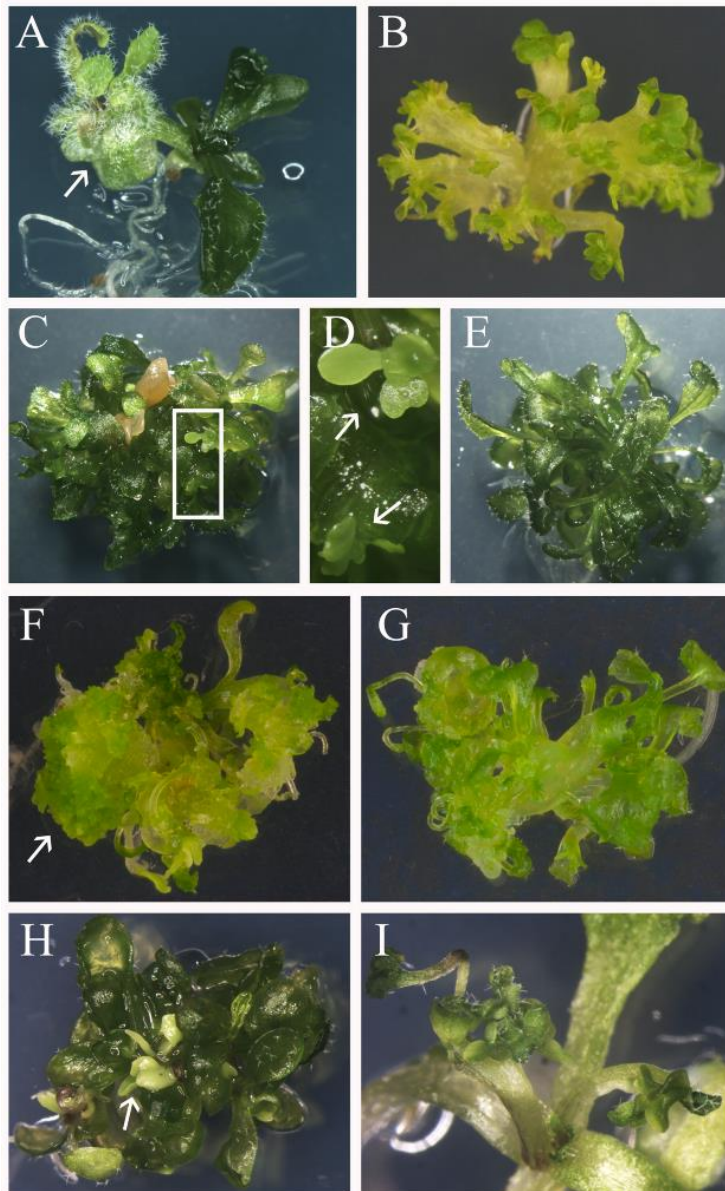


Fig. 6. *HDG* cosuppression results in ectopic meristem and embryo formation

(A) A *p35S::HDG1* seedling with ectopic shoots on the cotyledon (arrow).

(B) Part of a *p35S::HDG1* seedling that reiteratively formed ectopic shoots.

(C-E) *p35S::GR-HDG1* seedlings with proliferation phenotypes in the absence of DEX. (D) shows a magnification of the region indicated by a box in (C), containing light green, smooth and fleshy embryo-like tissue (arrows).

(F-I) Overproliferation phenotypes of seedlings expressing an amiRNA construct targeting multiple *HDG* genes in (F, G) the *hdg1;anl2* double mutant background and (H, I) a wild-type background. The arrows in (F) and (H) indicate somatic embryo-like tissue.

All seedlings were 4-5 weeks old, except for the seedling shown in (A), which was two weeks old.

Co-expression of HDG1 and BBM antagonizes both proteins' functions

If HDG1 and BBM function antagonistically in the regulation of cell proliferation, then co-overexpression of HDG1 and BBM should mitigate the other protein's overexpression phenotype. To test this hypothesis, we transformed a *p35S::BBM* construct to a characterized *p35S::GR-HDG1* line (>90% penetrance of the phenotype) and examined the phenotypes of the double transgenic seedlings before and after DEX-activation of the GR-HDG1 protein. The expression of both transgenes was verified using qPCR (supplementary material Fig. S6). The

effect of HDG1 overexpression on the BBM phenotype was dependent on the penetrance of the BBM somatic embryogenesis (SE) overexpression phenotype (Table 1). A line that showed a high penetrance of BBM-mediated SE was unaffected by co-overexpression of HDG1 (Table 1, line 1), while SE was reduced in double transgenic lines with a lower penetrance of this BBM overexpression phenotype (Table 1, lines 2-5). In addition, we observed that the HDG1 overexpression phenotype was also compromised in the co-overexpression lines compared to the phenotype of the parental *p35S::GR-HDG1* line, even in lines with a mild BBM overexpression phenotype (supplementary material Fig. S7). This suggests that BBM and HDG1 function antagonistically and that the balance between cell proliferation and differentiation depends on their relative concentrations.

Table 1. Phenotypes observed in *p35S::BBM/p35S::GR-HDG1* co-overexpression lines with and without DEX-activation of GR-HDG1.

Line	Percentage SE without DEX	Percentage SE with DEX
1	97% (<i>n</i> =141)	95% (<i>n</i> =229)
2	62% (<i>n</i> =290)	45% (<i>n</i> =359)
3	28% (<i>n</i> =460)	4% (<i>n</i> =565)
4	19% (<i>n</i> =102)	2% (<i>n</i> =139)
5	13% (<i>n</i> =77)	5% (<i>n</i> =78)
6	0% (<i>n</i> =99)	0% (<i>n</i> =173)
7	0% (<i>n</i> =67)	0% (<i>n</i> =114)

HDG1 represses transcription of meristem and cell proliferation genes

We performed microarray experiments to understand how HDG1 controls cell proliferation, as well as its functional relationship with BBM. Direct HDG targets were identified using 5-day-old *p35S::GR-HDG1* seedlings treated with DEX for 8 hours in the presence of the translational inhibitor cycloheximide (CHX). *HDG1* expression was significantly upregulated in CHX-treated *p35S::GR-HDG1* seedlings compared to CHX-treated wild-type seedlings, but was not increased in DEX+CHX-treated *p35S::GR-HDG1* seedlings compared to CHX-treated *p35S::GR-HDG1* seedlings (supplementary material Dataset S1 online; DEX+CHX experiment). These data indicate that *HDG1* is overexpressed, but that HDG1 does not regulate its own expression.

Statistical analysis identified 26 genes that were significantly differentially expressed at least two-fold in DEX+CHX-treated *p35S::GR-HDG1* seedlings compared to the controls (wild-type DEX+CHX-treated, and *p35S::GR-HDG1* CHX-treated; supplementary material Dataset S1 online;

DEX+CHX experiment). The expression level of the GR-HDG1 fusion protein did not change after CHX treatment compared to untreated samples (supplementary material Fig. S8), suggesting that the minimal change in gene expression is not due to a CHX-mediated reduction in HDG1. Since the dataset did not provide clear links with the observed *HDG1* overexpression phenotypes, we performed another microarray experiment in which *p35S::GR-HDG1* and wild-type seedlings were treated for 8 hours with only DEX. Using this approach, we identified 63 differentially expressed genes, including the *HDG1* transgene. In contrast to the DEX+CHX experiments, most of the genes that were differentially expressed in response to GR-HDG1 activation were down-regulated (79%; supplementary material Dataset S1 online; DEX experiment). The differential expression of a selection of these genes was validated by qPCR supplementary material Fig. S9).

HDG1 targeted a diverse group of genes, including those involved in transport (e.g. *ZIP1*, *SUC1*, *AAP4*) and hormone biosynthesis, transport or signalling (e.g. *GA3OX1*, *PIN5*, *ENP/MAB4*, *ARR16*), as well as genes involved in biosynthesis and transport of methionine-derived aliphatic glucosinolates (*MYB29*, *MAM1*, *CYP79F2*, *CYP83A1/REF1*, *IPMI1*, *IPMI2* and *BAT5*). Notably, HDG1 down-regulated the expression of five positive regulators of meristem development/cell proliferation: *CYCD3;1*, *CLE41*, *DAR2*, *RUL1* and *AIL5/PLT5*.

HDG and BBM transcriptional pathways intersect

The BBM and HDG1 transcription factors interact, suggesting that they might regulate a common set of target genes. We compared the list of HDG1 target genes with direct BBM target genes that were obtained by ChIP-seq analysis of BBM binding sites in somatic embryos (supplementary material Dataset S2 online, supplementary material Fig. S10). We observed BBM binding to 17 of the genes that showed differential expression after DEX-activation of the GR-HDG1 fusion protein, including *CLE41* and *AIL5/PLT5* (Table 2; supplementary material Fig. S11). We selected five genes that showed promoter binding close to the translational start site by BBM and that had a reasonable gene expression change upon GR-HDG1 activation in the microarray experiment. qPCR analysis of gene expression changes after BBM-GR or GR-HDG1 activation showed that *CLE41*, *RanBP2* and *TRM13* were antagonistically regulated by BBM (up) and HDG1 (down), while *AIL5* and *ATC* were downregulated by both BBM and HDG1 (Fig. 7). This suggests that HDG1 and BBM have common target genes that may be antagonistically regulated or co-regulated.

Table 2. Overlap between HDG1 and BBM target genes

HDG1 targets (microarray)		Score in BBM ChIP-seq					
Gene	Protein	2-log fold change	u3000	u2000	u1000	d0	d1000
AT3G24770	CLAVATA3/ESR-RELATED 41 (CLE41)	-1.83	3.67	6.5	0	0	0
AT2G27550	CENTRORADIALIS (ATC)	-2.02	0	7.83	0	2.07	2.35
AT5G57390	AINTEGUMENTA-like 5 (AIL5)	-1.41	0	0	16.3	1.59	1.93
AT3G15680	Ran BP2/NZF zinc finger-like superfamily protein	-2.69	7.06	1.46	3.9	0	4.41
AT1G15550	Gibberellin 3-oxidase 1 (GA3OX1)	-1.61	0	0	1.72	0	6.35
AT4G31820	ENHANCER OF PINOID (ENP)/MACCHI-BOU 4 (MAB4)	1.02	0	0	7.45	1.35	1.03
AT1G23090	Sulfate transporter 91 (AST91)	-1.40	1.02	3.26	0	5.81	0
AT5G03610	GDSL-like lipase	-1.18	1.02	1.18	4.56	0	1.22
AT1G15380	GLYOXYLASE I 4 (GLYI4) lactoylglutathione lyase	-1.13	1.17	1.11	4.12	1.05	1.08
AT3G15720	pectin lyase	-1.14	0	0	11.69	1.61	3.46
AT4G29920	Double Clp-N motif-containing P-loop nucleoside triphosphate hydrolases superfamily protein	-1.03	1.37	2.91	1.36	0	6.26
AT2G45900	Phosphatidylinositol N-acetylglucosaminyltransferase subunit P-related (TRM13)	-3.05	0	0	7.36	1.35	0
AT1G78370	glutathione S-transferase TAU 20	-1.69	0	1.13	0	0	4.59
AT1G07710	Ankyrin repeat family protein	1.62	9.45	0	0	1.22	0
AT5G22930	Protein of unknown function (DUF1635)	-1.40	0	6.47	4.52	1.09	1.98
AT4G36850	PQ-loop repeat family protein	-1.89	0	2.88	0	5.81	1.08
AT2G47560	RING/U-box superfamily protein	-1.21	0	4.94	0	2.56	0

The ChIP-seq score reflects the height of the binding peaks in the *pBBM::BBM-YFP* ChIP, with u3000 showing the maximum score value in the region 3kb-2kb upstream of the protein coding region, u2000 for the region 2kb-1kb upstream, u1000 for the region 1kb-0kb upstream, d0 for the coding region, and d1000 for the 0kb-1kb downstream of the coding region. Peaks with scores above 3.96 are considered statistically significant (FDR<0.05). The shaded rows indicate the genes that were selected for gene expression analysis.

Our previous observation that BBM transcriptionally activates the *HDG* gene *PDF2* and the epidermally-expressed *GASSHO1* (*GSO1*) (Supplemental Table 2 of (Passarinho et al., 2008)) prompted us to examine whether BBM binds other *HDG* and L1-expressed genes. BBM binding was observed at *HDG1*, *HDG5*, *HDG7*, *HDG8*, *HDG11*, *ANL2*, *PDF2*, *ATML1*, and a set of epidermis-expressed genes, including *GSO1*, *GSO2*, *CRINKLY4* (*ACR4*) and *WEREWOLF* (*WER*) (supplementary material Fig. S11). BBM binding was mostly observed in the promoters of these genes, however, in some cases introns were bound (supplementary material Fig. S12). Increased expression of *HDG12*, *PDF2*, *GSO1* and *GSO2* was observed when 5-day-old *p35S::BBM-GR*

seedlings were treated with DEX and CHX, showing that these *HDG* and *L1* genes are direct transcriptional targets of *BBM* (supplementary material Fig. S12).

The combined microarray and ChIP-seq data analysis suggest that *BBM* and *HDG1* regulate a common set of target genes, but this regulation appears to be complex, as both coordinately and oppositely-regulated transcription, as well as a transcriptional feedback loops between *AIL* and *HDG* genes were observed. Additionally, our results uncovered a role for *BBM* in the transcriptional control of additional epidermal regulatory genes.

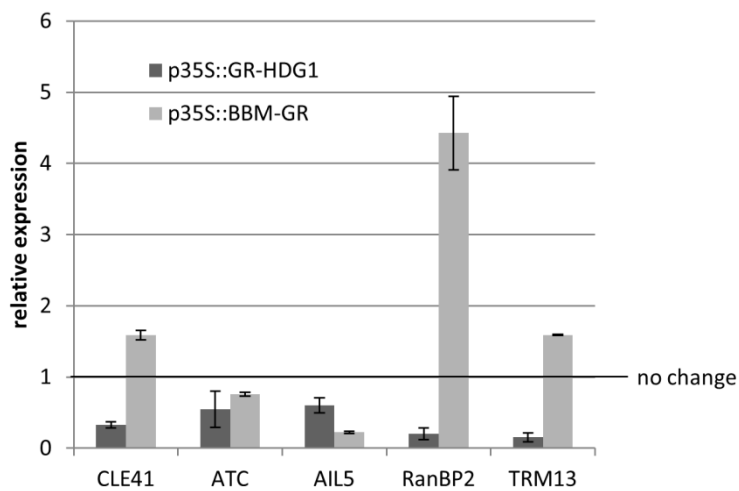


Fig. 7. *BBM* and *HDG1* regulate common target genes

qPCR analysis of overlapping *BBM/HDG1* target genes. The relative expression of five genes was determined in DEX+CHX-treated *p35S::BBM-GR* and DEX-treated *p35S::GR-HDG1* five day-old seedlings (two lines each) compared to DEX+CHX and DEX-treated *Col-0*, respectively. Error bars indicate standard errors of the two biological replicates.

DISCUSSION

Members of the *AIL* family have been well described with respect to their positive roles in stem cell maintenance. Here, we have shown that the *BBM AIL* protein interacts with and regulates the expression of *L1*-expressed *HDG* proteins. Analysis of gain- and loss-of-function *HDG* phenotypes suggests that *HDG* proteins function antagonistically to *AIL* proteins to keep cell proliferation processes in check.

***HDG* proteins stimulate cell differentiation**

We observed that *HDG1*, *HDG11* and *HDG12* overexpression seedlings were smaller compared to wild-type, and that they developed narrow leaves and accumulated anthocyanins. The most extreme overexpression phenotypes were observed in *HDG1* overexpression lines, which showed root and shoot meristem arrest. Another notable feature of *HDG1* overexpression lines was the increased formation of narrow, elongated cells on the leaf surface. In *Arabidopsis*

leaves, elongated cells are found on the abaxial leaf surface, on the petiole and along the leaf margin. The elongated cells formed in HDG1 overexpression lines resemble margin cells and showed increased expression of YJ158, which marks large/giant cells in the sepal, leaf margin and abaxial surface (Eshed et al., 2004; Roeder et al., 2012). Our data therefore suggest that HDG1 is able to specify the identity of margin cells. Recently, two other HDG proteins, ATML1 and HDG11, were shown to regulate sepal giant cell formation prior to endoreduplication (Roeder et al., 2012). The HDG1-mediated increase in cell ploidy in leaves may be a secondary consequence of the increase in margin cells. Alternatively, the increased ploidy levels might also reflect the increased differentiation/reduced meristematic growth that characterizes *HDG1* overexpression lines.

While overexpression of *HDG* genes induces meristem arrest and epidermal margin cell formation, knock-down of *HDG* expression by co-suppression or amiRNAs induced ectopic cell proliferation, including formation of shoots and embryo-like tissue. The formation of ectopic shoots was also observed occasionally in *pdf2;hdg3* or *atml1;hdg3* mutants (Nakamura et al., 2006). In addition, post-embryonic expression of *ATML1-SRDX* induced callus-like protrusions on cotyledons and leaves (Takada, 2013). Our *HDG* knock-down phenotypes were observed in a small proportion of transgenic lines, but were more severe than previously reported *HDG* loss-of-function phenotypes. Down-regulation of a larger number of *HDG* genes may have allowed us to overcome the high degree of functional redundancy within the HDG family (Nakamura et al., 2006). However, the low frequency of mutant phenotypes suggests that either *HDG* knock-down was inefficient or that it has negative impact on embryo viability (Abe et al., 2003; San-Bento et al., 2013). Taken together, these data suggest that *HDG* genes, besides their roles in the differentiation of specific epidermal structures, also have a general role in repressing cell proliferation in the epidermis. This role does not appear to be restricted to *HDG* genes, as loss-of-function or knock-down mutants in other epidermal-expressed genes that control epidermal differentiation also show cell over-proliferation phenotypes (Jin et al., 2000; Becraft et al., 2002; Ahn et al., 2004).

HDG1 target genes support its role in cell differentiation

We showed that HDG1 down-regulates the expression of genes involved in cell proliferation, including the D-type cyclin *CYCD3;1*. *CYCD3;1* overexpression leads to ectopic/increased cell divisions, reduced cell expansion and endoreduplication (Dewitte et al., 2003). Conversely, loss of *CYCD3* genes reduces leaf cell numbers and SAM size and stimulates

endoreduplication (Dewitte et al., 2007). HDG1 also inhibits the expression of *CLE41*, which encodes a B-type CLE signaling peptide. Overexpression of *CLE41* promotes the formation of axillary buds (Yaginuma et al., 2011), and co-overexpression of *CLE6* (A-type) and *CLE41* peptides induced ectopic divisions in root, leaf and the hypocotyl vasculature (Whitford et al., 2008), indicating a role for *CLE41* in cell proliferation. In addition, HDG1 represses the expression of *REDUCED IN LATERAL GROWTH1 (RUL1)*, which encodes a receptor-like kinase that positively regulates cambium activity (Agusti et al., 2011), and of *AIL5/PLT5*, which controls lateral root primordia initiation in a redundant fashion with *AIL6/PLT3* and *PLT7* (Hofhuis et al., 2013) and can induce increased organ size or SE when overexpressed (Nole-Wilson et al., 2005; Tsuwamoto et al., 2010). HDG1 could also indirectly down-regulate other *AIL* genes through *DAR2*, which was shown to act upstream of *PLT1/PLT2* in the control of root meristem size (Peng et al., 2013).

AILs and HDGs have antagonistic functions

We have shown that HDG and BBM/AIL proteins interact *in vitro* and *in planta*. The interaction between BBM and HDG proteins is limited mainly to embryo development, where the expression patterns of *BBM*, *HDG1*, *HDG11* and *HDG12* overlap extensively. However, as HDG proteins also interact with other AIL proteins, the expression patterns of the other *AIL* genes must be taken into account as well. For example, *PLT2* is expressed in all epidermal cells of the root meristem (Galinha et al., 2007), and overlaps with *HDG* expression in these cells. Although *BBM* is not expressed in the SAM, other *AILs* are expressed here, e.g *PLT7* (Mudunkothge and Krizek, 2012), and could interact with HDG proteins in the L1/L2 layers.

Interestingly, our *HDG1* overexpression phenotypes resemble *ail* loss-of-function phenotypes: *plt1;plt2* mutant roots terminate soon after initiation, whereas *plt1;plt2;ail6/plt3* mutants do not form any roots (Aida et al., 2004; Galinha et al., 2007), and *ant;ail6/plt3;plt7* mutants produce only a few leaves before the SAM terminates (Mudunkothge and Krizek, 2012). In addition to root and shoot meristem differentiation, we also observed other differentiation phenotypes in *HDG1* overexpression seedlings (ectopic formation of margin cells and higher ploidy levels), that were opposite to those observed in *BBM* overexpression seedlings (decreased cellular differentiation of cotyledons cells and reduced ploidy). Similarly, *AIL6* overexpression lines lack sepal giant cells (Krizek and Eaddy, 2012). In line with this antagonistic HDG-AIL model, we found that down-regulation of *HDG* expression by co-suppression or by an amiRNA leads to adventitious growth, similar to *BBM* or *PLT5/AIL5 (AIL)* overexpression (Boutilier et al., 2002; Tsuwamoto et al., 2010), and that co-overexpression of *BBM* and *HDG1* reduces the

overexpression phenotypes of both proteins. Taken together, our results suggest opposite roles for *AIL* and *HDG* genes, with *AILs* promoting meristem activity and *HDGs* stimulating meristem/cellular differentiation, and that these interactions are mediated at both transcriptional and protein-interaction level.

Molecular relationship between AIL and HDG proteins

We have shown that AIL and HDG proteins interact in yeast and *in planta* and that they have antagonistic functions. A mechanism for interacting proteins to exert antagonistic functions is through competitive inhibition. In this scenario, interaction of HDG and AIL proteins would inhibit their respective abilities to act as transcriptional regulators, with the balance between the amount of free HDG or AIL determining the developmental outcome. In support of this “titration model”, we found that BBM and HDG1 can suppress each other’s overexpression phenotypes in a dose-dependent manner. In this model, AIL overexpression would lead to opposite regulation of HDG target gene expression and *vice versa*. We have identified overlapping BBM and HDG1 gene targets, some of which are oppositely regulated upon BBM and HDG1 overexpression. Notably, some of the common target genes also appear to be co-regulated, suggesting that in addition to antagonistic functions, BBM and HDG also have common functions. This raises the possibility that they perform these functions in the same protein complex. Finally, we observed transcriptional cross-regulation between AILs and HDGs, suggesting an additional level of interaction.

Our results suggest that HDG and AIL act in concert to control cell proliferation and differentiation processes. Whether AIL-HDG function is cell-autonomous (Hacham et al., 2011; Knauer et al., 2013; Nobusawa et al., 2013) and how local BBM-HDG interactions regulate these processes at a molecular level are intriguing questions for further research.

MATERIALS AND METHODS

Plant material and growth conditions

The *pBBM::BBM-YFP* (Galinha et al., 2007), *p35S::BBM* (Boutilier et al., 2002), *p35S::BBM-GR* (Passarinho et al., 2008) lines, *HDG* T-DNA insertion alleles *hdg1-2* (SALK_062171), *hdg11-1* (SAIL_865 G09), *hdg12-2* (SALK_127261) and *anl2-t1* (SALK_000196) (Nakamura et al., 2006), and the giant cell marker line YJ158 (Eshed et al., 2004) have been previously described.

Plants were grown at 21 °C (16/8 hour light/dark regime) on rock wool plugs supplemented with 1 g/L Hyponex fertilizer, or in Petri dishes on medium containing half-strength Murashige and Skoog salts and vitamins (MS medium, pH 5.8), containing 0.8% agar and 1% sucrose (0.5MS-10). The medium was supplemented with 10 µM DEX when appropriate. To obtain a fully-penetrant *BBM-GR* overexpression phenotype it is necessary to sterilize seeds with liquid bleach, rather than bleach vapour.

Vector construction and transformation

All used primers are shown in supplementary material Table S3 (online).

For ectopic expression of *HDG1*, *HDG11* and *HDG12*, the open reading frames were amplified from Arabidopsis Col-0 cDNA and cloned into the Gateway (GW) overexpression vector pGD625 (Immink et al., 2002).

For inducible activation of *HDG1*, *HDG11* and *HDG12* the *HDG* coding regions were fused in-frame to the ligand binding domain of the rat glucocorticoid receptor (GR) coding region and then cloned into pGD625. For co-overexpression, a *p35S::GR-HDG1* line was transformed with a construct overexpressing the genomic Arabidopsis *BBM* fragment (*p35S::gAtBBM*) in pB7WG2.0 (Karimi et al., 2002).

The *BBM* and the *HDG* translational eGFP reporter constructs were made by cloning Arabidopsis Col-0 genomic DNA into the GW-compatible pGreenII vector AM884381 (NASC) (Zhong et al., 2008). The promoters of *BBM*, *HDG1*, *HDG11* and *HDG12* comprised respectively 4.2, 0.65, 2.7 and 1.2 kb upstream of the translational start site.

BBM and *HDG* cDNA entry clones were used to generate the *BBM-CFP* and *YFP-HDG* plasmids used for FRET-FLIM experiments, and cloned using recombination into GW-compatible sCFP3A and sYFP2 vectors (Karlova et al., 2011).

The *HDG* amiRNA construct was designed and generated according to WMD3 (Schwab et al., 2006) and cloned into pGD625.

For ChIP-seq experiments, a *pBBM::NLS-GFP* construct was generated in pGREEN using a 4.2 kb *pBBM* fragment. The *p35S::BBM-GFP* construct was using the *BBM* (At5g17430) Col-0 cDNA in pK7FWG2.0 (Karimi et al., 2002).

Arabidopsis Col-0 plants were transformed by the floral dip method (Clough and Bent, 1998) using *Agrobacterium tumefaciens* strain C58 carrying the pMP90 Ti plasmid.

FRET-FLIM

Protoplasts were transfected as described (Horstman et al., 2014b) and incubated overnight. All plasmid combinations were tested in three independent transfections (26-30 cells per combination), except for YFP-AP1, which was tested twice (26 cells). FRET-FLIM measurements were performed as previously described (Ruscinova et al., 2004). Photons were collected by a Hamamatsu HPM-100-40 Hybrid detector (Becker & Hickl; resolution 120 ps). 64 × 64 pixel images were acquired (60-150 seconds, count rate ca. 10⁴ photons per second) using the Becker and Hickl SPC 730 module. The data were analysed with SPCImage 3.10 software (Becker & Hickl) using a one- or two-component decay model for donor-alone and donor-plus-acceptor samples, respectively. The two-component analysis gives a slightly reduced fluorescence lifetime compared to the one-component analysis. Significant differences between samples were determined using a two-tailed Student's *t*-test.

GUS staining

β-glucuronidase (GUS) activity assays on seedlings were performed overnight at 37 °C as described (Soriano et al., 2014), using 2.5 mM potassium ferri- and ferrocyanide. Seedlings were cleared with 70% ethanol prior to imaging.

Flow cytometry

Ploidy measurements were performed (Iribov, The Netherlands) on whole 10 day-old seedlings (three to four replicates), using one seedling per replicate.

Microscopy

GFP was visualized with a Leica SPE DM5500 upright confocal laser scanning microscope using the LAS AF 1.8.2 software. Roots were counterstained with 10 µg/mL propidium iodide (PI), and embryos and shoot/flower meristems with 10 µg/mL FM4-64. GFP, YFP, PI and FM4-64 were

excited with a 488-nm solid-state laser and emissions were detected at band widths of 500–530, 510–560, 670–800 and 600–800 nm, respectively.

Cryo-SEM was performed as in (Fatouros et al., 2012), except that samples were sputter-coated with 10 nm tungsten and the analysis was performed with SE detection at 2 kV and 6.3 pA. Digital images were contrast-adjusted with Photoshop CS5.

ChIP-seq

ChIP-seq experiments were carried out using a GFP antibody on 1) 14- to 17-day-old 2,4-dichlorophenoxyacetic acid (2,4-D)-induced somatic embryo cultures (Mordhorst et al., 1998) (4.8 g, ectopic shoots and callus removed) carrying a *pBBM::BBM-YFP* construct (Galinha et al., 2007) in the *bbm-1* (SALK_097021; NASC) mutant background and on 2) embryogenic seedlings (1.75 g) derived from a *p35S::BBM-GFP* line. The *pBBM::BBM-YFP* construct complemented the embryo lethal phenotype of the *bbm;plt2* double mutant (not shown). 2,4-D somatic embryo cultures from a *pBBM::NLS-GFP* line and *p35S::BBM* seedlings served as negative controls for experiment 1) and 2), respectively. ChIP samples were prepared as described previously (Kaufmann et al., 2010; Smaczniak et al., 2012).

ChIP-seq libraries were sequenced on the Illumina HiSeq 2000 platform. Sequence reads that failed the CASAVA quality filter were eliminated. Sequence reads were mapped to the unmasked Arabidopsis genome (TAIR10; ftp://ftp.arabidopsis.org) using the SOAPaligner (v2) program (Li et al., 2009). A maximum of two mismatches and no gaps were allowed. Reads mapping in multiple genomic locations or to the chloroplast or mitochondrial genomes were discarded. ChIP-seq peaks were detected using CSAR (Muino et al., 2011) with default parameter values except for “backg”, which was set to 2. Enrichment was calculated as the ratio of normalized extended reads between the *pBBM::BBM-YFP* or *p35S::BBM-GFP* samples versus their corresponding controls. False discovery rate (FDR) thresholds were estimated by permutation of reads between IP and control sample using CSAR. ChIP-seq results were visualized using Integrated Genome Browser 8.1.2 (Nicol et al., 2009). The ChIP-seq data is made available via NCBI (GEO accession: GSE52400).

Microarray analysis

For each sample, approximately 40 five-day old seedlings were treated with 10 μ M DEX and/or 10 μ M CHX for 8 hours. All conditions were replicated in triplicate. Microarray analysis was performed by NASC using Affymetrix Arabidopsis Gene 1.0 ST Arrays. Arrays were

normalized using the RMA algorithm and differential expression was assessed with the LIMMA package and the Benjamini and Hochberg multiple testing correction (Benjamini and Hochberg, 1995; Carvalho and Irizarry, 2010). HDG1 targets were defined as showing a fold change >2 and FDR threshold of 0.05. The data is available via NCBI (GEO accession: GSE54312)

Quantitative real-time RT-PCR

For analysis of (*GR-*)*HDG1*, *BBM* and target gene expression, seedlings were grown and treated as described in the text (DEX and CHX, both at 10 μ M). DNase-treated RNA was used for cDNA synthesis. qPCR was either performed using the SYBR green mix from BioRad or the BioMark HD System (effect of BBM on *HDG/L1* gene expression) on a 96.96 dynamic array chip according to the manufacturer's instructions (Fluidigm) and analysed using the manufacturer's software. In all experiments, the relative expression levels were calculated according to the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001) using wild-type Col-0 as the calibrator and the *SAND* gene (Czechowski et al., 2005) as the reference. All primers are listed in supplementary material Table S3 (online).

SUPPLEMENTARY MATERIAL

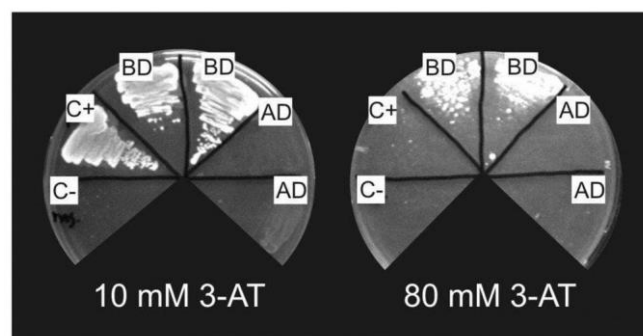


Fig. S1. Transcriptional auto-activation of BD-BnBBM in yeast.

The Gal4 DNA binding domain-BnBBM fusion protein (BD) induced *HIS3* marker gene expression in yeast in the absence of an interacting activation domain fusion protein when plated on synthetic dropout medium (SD) lacking histidine, in the presence of 10 mM 3-AT (left). AD-BnBBM (AD) did not induce *HIS3* marker expression. The transcriptional auto-activation activity of BD-BnBBM could not be suppressed by a higher 3-AT concentration (right).

Co-transformation of AD- and BD-fusions with petunia MADS box protein pairs were used as controls for positive (C+, BD-FLORAL BINDING PROTEIN2 (FBP2) and AD-FBP7) and negative (C-, BD-FBP2 and AD-FBP20) protein-protein interactions.

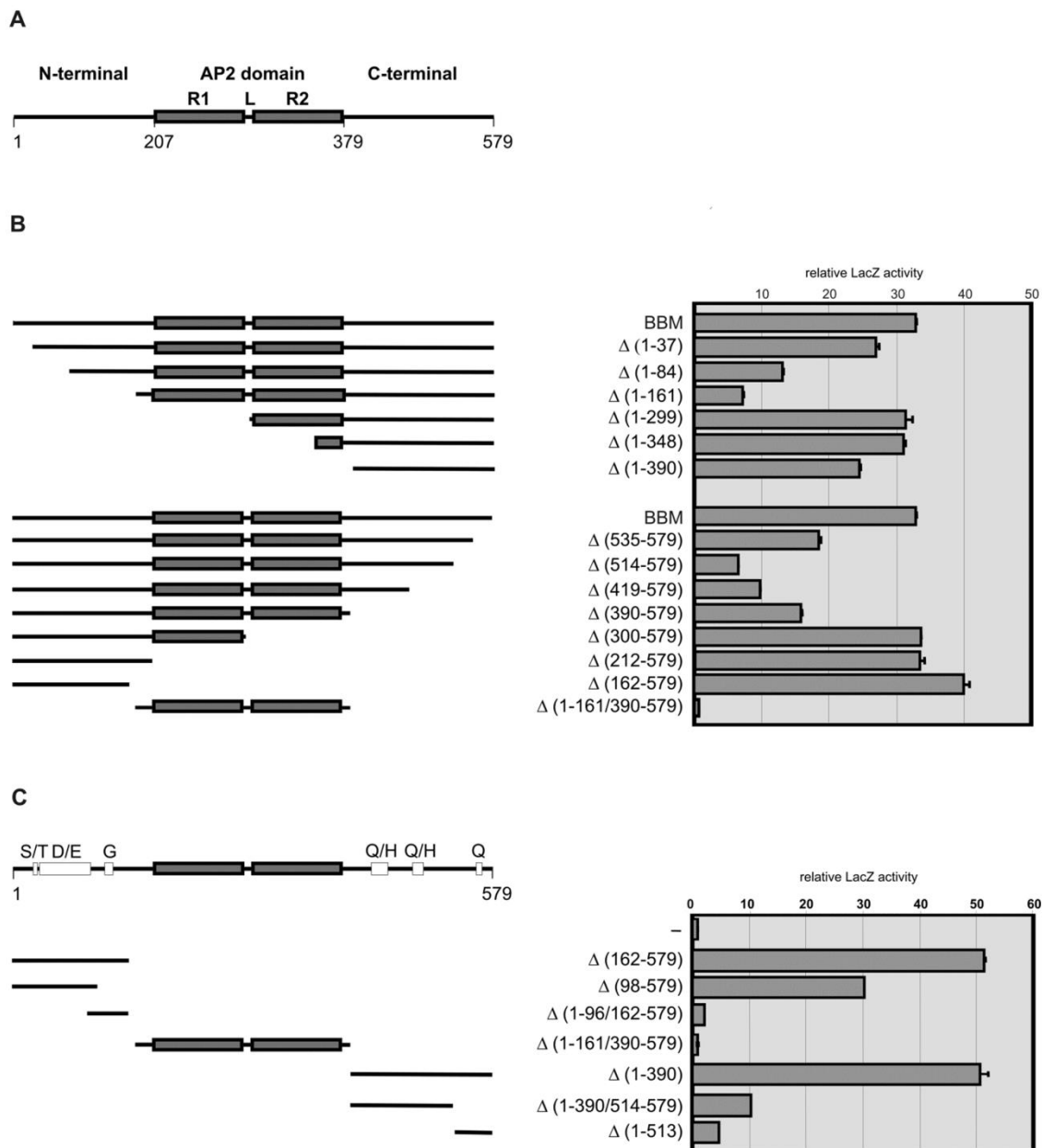


Fig. S2. The N- and C-terminal regions of BBM have transcriptional activation activity in yeast

(A) Schematic overview and amino acid positions of the defined regions of the BnBBM protein.

(B) Schematic representation of the Gal4 BD-BBM deletion fusion proteins (Gal4 BD- Δ BBM) used to localize the position of the BBM activation domains (left) and the relative LacZ activity obtained for each construct after transformation to yeast (right). The relative LacZ activity was calculated for each transformed construct or construct pair from the average of three replicates as described in Supplementary Materials and Methods. The standard error is indicated.

(C) Mapping of the BBM N- and C-terminal activation domains. Regions of the BBM protein enriched for specific amino acids are superimposed as white blocks along the BBM protein schematic, with the enriched amino acid(s) indicated above the blocks. The BBM deletion (Δ) derivatives assayed for transcriptional autoactivation and the relative LacZ activity obtained for each construct after transformation to yeast are shown below. The deleted amino acids are shown in brackets for each construct.

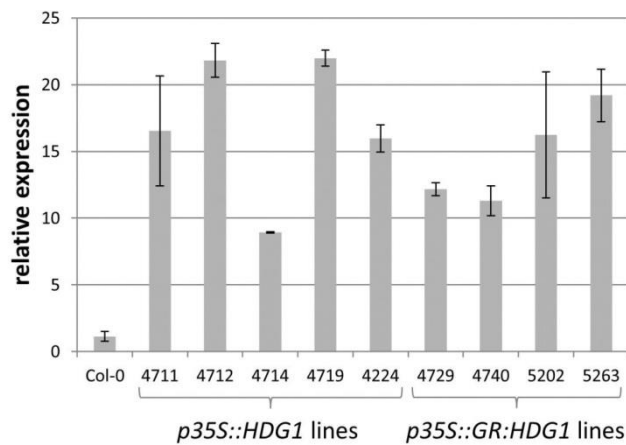
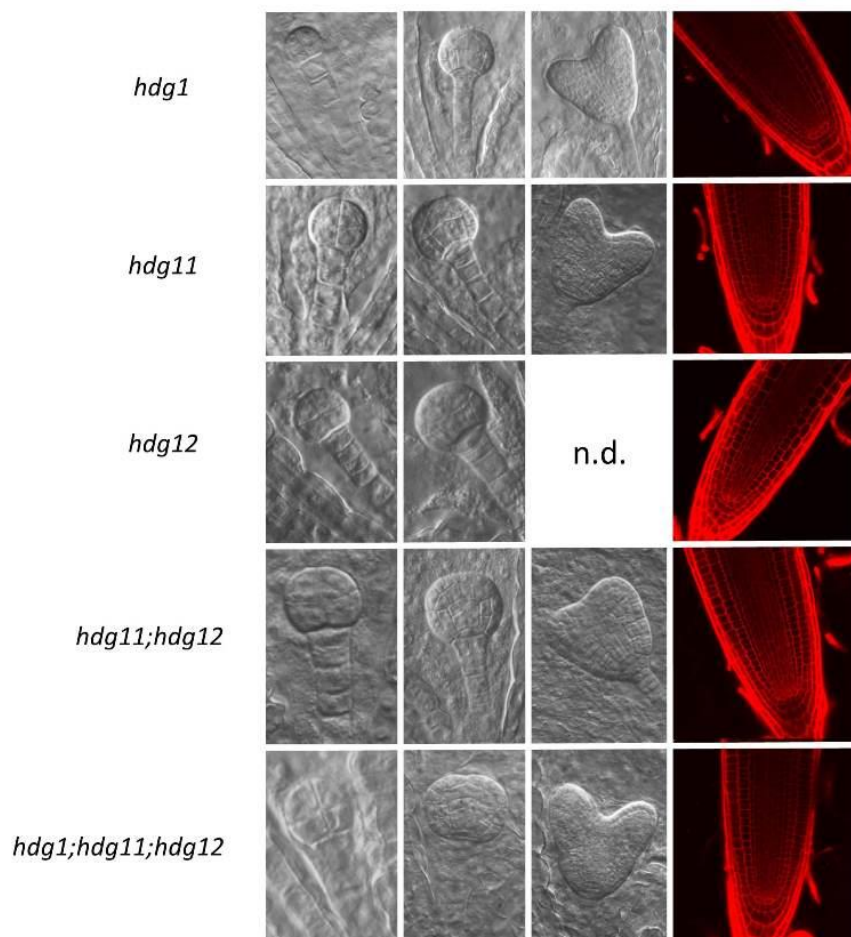


Fig. S3. HDG1 expression levels in *p35S::HDG1* and *p35S::GR-HDG1* lines

qPCR was performed on two pools of 20-40 11 day-old seedlings from five independent *p35S::HDG1* lines (4711, 4712, 4714, 4719 and 4224) and four independent *p35S::GR-HDG1* lines (non-DEX treated; 4729, 4740, 5202; 5263), and relative expression (compared to Col-0) was calculated as described in the Materials and Methods. Error bars indicate standard errors of the biological controls (two pools of seedlings).

Fig. S4. *hdg* mutants show normal root and embryo development

DIC microscopy images of cleared 8-to-32-cell (1st column), globular (2nd column) and heart (3rd column) stage *hdg* embryos, and confocal images of propidium iodide-stained *hdg* primary roots (4th column).



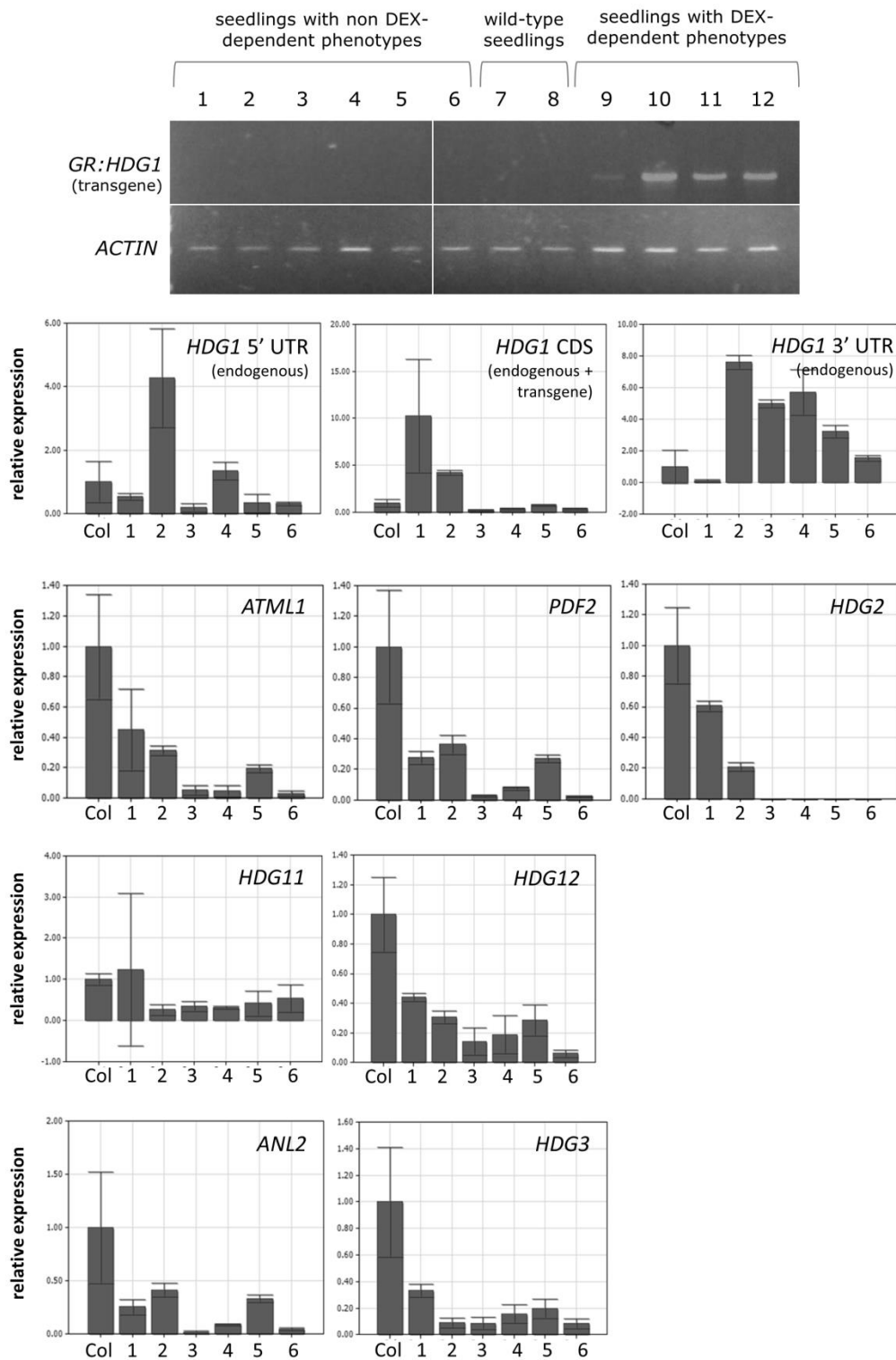


Fig. S5. Expression analysis of selected *HDG* genes in putative cosuppression lines

The expression level of the *GR-HDG1* transgene was examined by semi-quantitative RT-PCR amplification of the *GR-HDG1* transgene in individual seedlings with non DEX-dependent mutant phenotypes (gel images; lanes 1-6), wild-type seedlings (lanes 7-8) and seedlings with DEX-dependent mutant phenotypes (lanes 9-12). *ACTIN8* was used as a reference gene. The expression levels of several *HDG* genes in putative *HDG* cosuppression lines were examined by qPCR in individual seedlings. The relative expression of the target genes was calculated as described in the Material and Methods. Error bars indicate standard errors of the two technical replicates of each line. UTR, untranslated region; CDS, coding sequence.

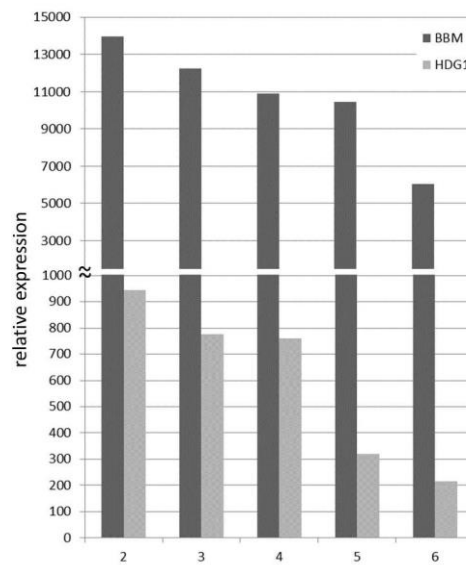


Fig. S6. *BBM* and *HDG1* expression levels in *p35S::BBM/p35S::GR-HDG1* lines

RNA was extracted from pools of 20-40 7 day-old seedlings from five co-overexpression lines (2-6, see Table 1) and wild-type Col-0 (all non-DEX treated). qPCR was performed as described in Materials and Methods using primers on the coding regions of *BBM* and *HDG1* (endogenous + transgene). The relative expression of *BBM* and *HDG1* was calculated as described in the Material and Methods. All lines show highly elevated *BBM* and *HDG1* expression levels compared to wild-type (set to 1). The variation in *HDG1* gene expression in the lines is likely due to differences in *p35S* promoter expression in lines with a different developmental status.

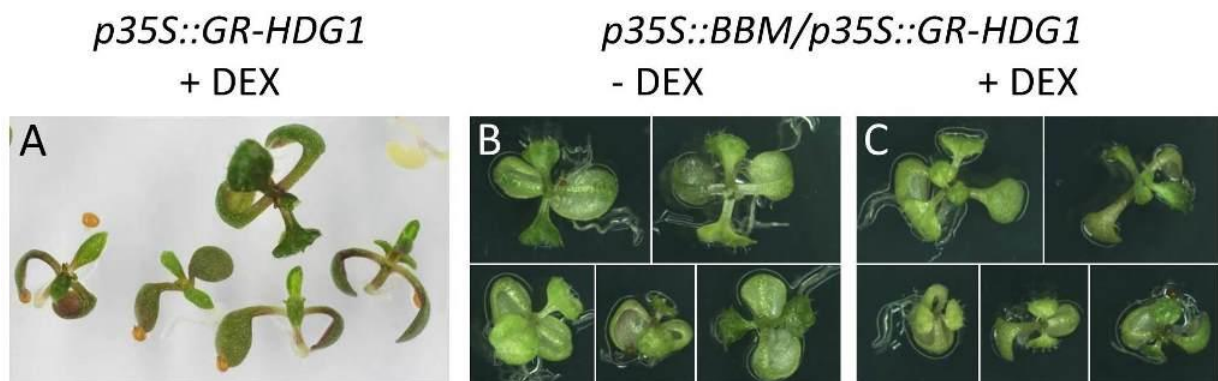


Fig. S7. *BBM* overexpression alleviates the *HDG1* overexpression phenotype

(A) 10 day-old DEX-induced seedlings of the *p35S::GR-HDG1* line that was used for transformation with *p35S::BBM*.

(B, C) 13 day-old seedlings of *p35S::BBM/p35S::GR-HDG1* co-overexpression line #6 grown on medium without DEX (B) and with DEX (C). In the absence of DEX, the co-overexpression line shows a mild *BBM* phenotype, characterized by lack of somatic embryos, reduced growth, round leaves, short petioles and epinastic cotyledons (B). DEX-induced *p35S::BBM/p35S::GR-HDG1* seedlings (C) showed phenotypes intermediate to those of the parent line (A) and of the uninduced *p35S::BBM/p35S::GR-HDG1* seedlings (C).

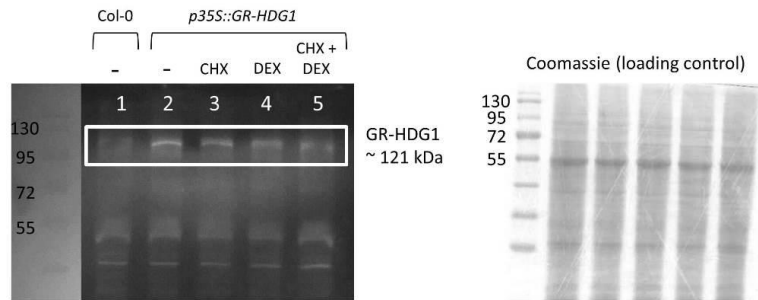


Fig. S8. Effect of dexamethasone and cycloheximide on GR-HDG1 stability

(A) Western blot of five day-old Col-0 and *p35S::GR-HDG1* seedlings treated for 8 hours with either ethanol (mock; '-'), cycloheximide (CHX), dexamethasone (DEX) or both CHX and DEX. The GR-HDG1 fusion protein (ca. 121 kDa) was detected using a primary GR antibody and with a secondary antibody (goat anti-rabbit IgG conjugated to peroxidase).

(B) The Coomassie-stained SDS-PAGE gel of the samples in (A) served as a loading control.

Western blotting and Coomassie staining were performed as described in Supplementary Materials and Methods.

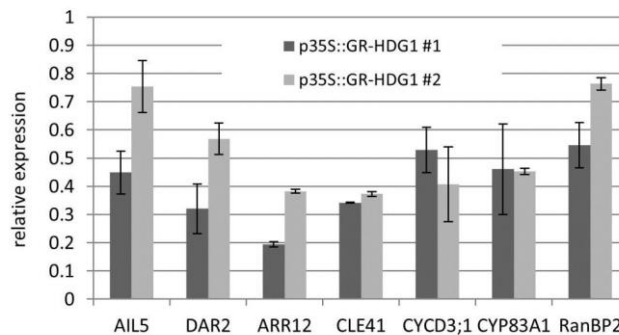


Fig. S9. Validation of HDG1 target genes

qPCR analysis of HDG1 target genes identified using microarray analysis. The relative expression of the *HDG1* target genes was calculated for DEX-treated *p35S::GR-HDG1* five day-old seedlings compared to DEX-treated Col-0 as described in Material and Methods. The error bars indicate the standard errors of the biological replicates (two pools of seedlings from each line).

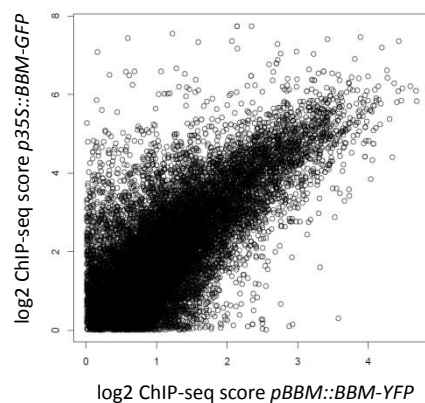


Fig. S10. Correlation between the *pBBM::BBM-YFP* and *p35S::BBM-GFP* ChIP-seq data

Scatterplot showing the correlation between the ChIP-seq experiments, where each circle represents the maximum log₂ ChIP-score value on the promoter region (1kb upstream of the start of the gene) of an Arabidopsis gene. The Pearson coefficient (R) = 0.73.

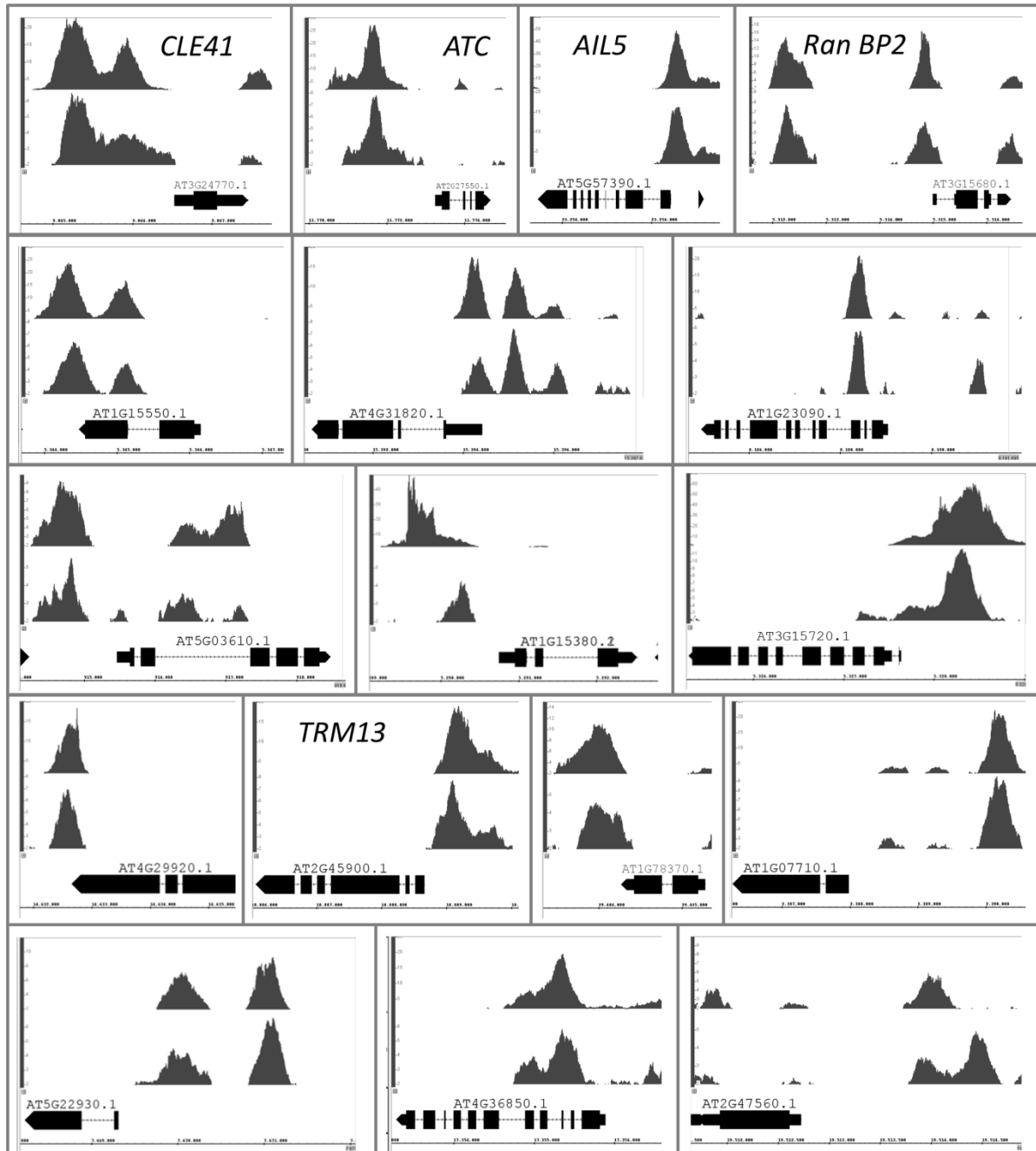


Fig. S11. Binding of BBM to common BBM-HDG1 target genes

ChIP-seq BBM binding profiles for common BBM/HDG1 target genes in somatic embryo tissue. The binding profiles from the *p35S::BBM-GFP* (upper panel) and *pBBM::BBM-YFP* (lower panel) ChIP-seq experiments are shown. The x-axis shows the nucleotide position of DNA binding in the selected genes using the TAIR 10 annotation. The y-axis shows the ChIP-seq score. Peaks with scores above 1.76 (for *p35S::BBM-GFP*) and 3.96 (for *pBBM::BBM-YFP*) are considered statistically significant (FDR<0.05).

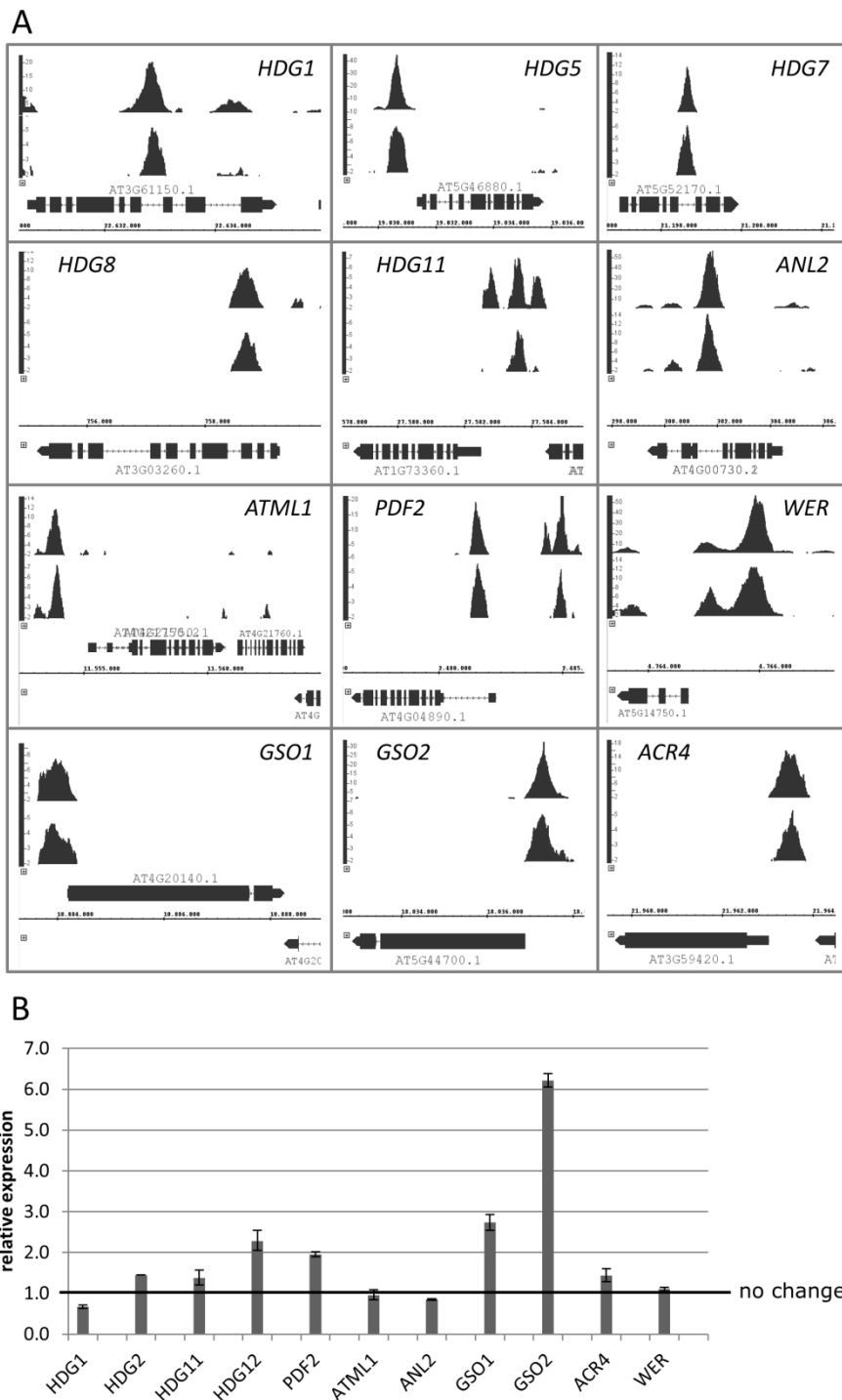


Fig. S12. BBM binds to and activates *HDG* and other L1 genes

A. ChIP-seq BBM binding profiles for *HDG* and L1 genes in somatic embryo tissue. The binding profiles from the *p35S::BBM-GFP* (upper panel) and *pBBM::BBM-YFP* (lower panel) ChIP-seq experiments are shown. The x-axis shows the nucleotide position of DNA binding in the selected genes using the TAIR 10 annotation. The y-axis shows the ChIP-seq score. Peaks with scores above 1.76 (for *p35S::BBM-GFP*) and 3.96 (for *pBBM::BBM-YFP*) are considered statistically significant (FDR<0.05).

B. qPCR analysis of BBM target genes identified using ChIP-seq. The relative expression of *HDG* and L1-expressed genes was calculated for DEX+CHX-treated *p35S::BBM-GR* five day-old seedlings compared to DEX+CHX-treated Col-0 as described in the Material and Methods. Error bars indicate standard errors of the biological replicates (two independent *p35S::BBM-GR* lines).

Table S1. The N- and C-terminal regions of BBM show transcriptional activation activity *in planta*

The two-component expression system used to detect transient activation of gene expression by BBM *in planta* is shown schematically above the table. In the activator construct, the *B. napus* 22A1 promoter (*P22a1*) was used to express a modified GAL4-BD (mBD)- Δ BBM fusion protein. In the reporter construct, four copies of the GAL4-BD binding site, or upstream activating sequence (4x UAS), were placed upstream of a minimal 35S CaMV promoter (*35S_{min}*) and the firefly *Luciferase* (*fLUC*) reporter gene. Individual activator constructs were co-bombarded with the reporter construct to *Brassica napus* microspore-derived embryos. Relative *fLUC* activity was calculated using a Renilla *luciferase* (*P22a1::RLUC*) construct as an internal standard as described in the Supplemental Materials and Methods section. **, proteins that significantly activated transcription of the reporter gene (Dunnet's test $p < 0.05$), as compared to the *P22a1::mBD* background control.



plasmid	relative activity
mBD	1.00 ± 0.61
mBD-BBM Δ (212-579)	725.90 ± 93.98**
mBD-BBM Δ (1-390)	15.14 ± 5.73**

Table S2. BBM-interacting proteins identified in a yeast two-hybrid screen

Clone	GenBank Accession	# clones	Closest Arabidopsis locus	NCBI accession	Identical amino acids	Conserved amino acids
BBP1A	DQ166821	5	At1g73360 (<i>HDG11</i>)	NP_177479	574/702 (82%)	610/702 (87%)
			At1g17920 (<i>HDG12</i>)	NP_564041	485/687 (71%)	553/687 (80%)
BBP1B	DQ182489	4	At1g73360 (<i>HDG11</i>)	NP_177479	571/702 (81%)	608/702 (87%)
			At1g17920 (<i>HDG12</i>)	NP_564041	487/689 (71%)	556/689 (81%)
BBP2	DQ182490	1	At3g61150 (<i>HDG1</i>)	NP_191674	570/715 (80%)	589/715 (82%)
			At4g00730 (<i>ANL2</i>)	NP_567183	434/683 (64%)	498/683 (73%)

SUPPLEMENTARY MATERIALS AND METHODS

Luciferase assays

Microspore-derived embryos at the late globular- to early heart-shaped stage (nine days after culture initiation) (Custers et al., 1994) were collected onto the surface of Millipore HATF0047 detergent-free membranes (Millipore, Amsterdam) using a filter holder apparatus. The membranes were then placed on NLN-13 medium solidified with 0.6% agarose and then bombarded with the DNA constructs. The particle bombardment procedure was performed according to (Fukuoka et al., 1998), except that each plate was bombarded twice. Each combination of an effector plasmid (promoter:BBM Δ -Gal4-BD), a reporter plasmid (UAS::firefly luciferase) and an internal standard plasmid (UAS::Renilla luciferase) were mixed respectively in a 2:2:1 molar ratio, adsorbed onto gold particles and used for bombardment. Experiments were performed in triplicate. Embryos were incubated *in situ* on the solid medium for 48 hours after bombardment and then collected for luciferase assays. The activity of the two reporter genes, firefly luciferase (*fLUC*) and *Renilla* luciferase (*rLUC*) were measured using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions. A *p35S::GUS* plasmid was bombarded as a positive control, and the empty pCmBDSN cassette as a control for background activity of the pCmBDSN-containing constructs.

Yeast two-hybrid experiments

The transcriptional autoactivation activity of the BBM deletion derivatives was quantified using the *LacZ* marker gene. *LacZ* activity was measured using o-nitrophenyl β -D galactopyranoside as a substrate. *LacZ* activity was normalized to OD600. The truncated BBM proteins were constructed in the BD Gal4 vector in frame with, and upstream of the Gal4 BD using restriction enzyme digests as follows: BBM Δ (1-37), *ScaI*; BBM Δ (1-84), *SphI*; BBM Δ (1-161), *StuI*; BBM Δ (1-299), *FspI*; BBM Δ (1-348), *PstI*; BBM Δ N(1-390), *HaeII*; BBM Δ (1-513), *HpaI*; BBM Δ (535-579), *CelII*; BBM Δ (514-579), *HpaI*; BBM Δ (419-579), *NheI*; BBM Δ (390-579), *HaeII*; BBM Δ (300-579), *FspI*; BBM Δ (212-579), *AosIII*; BBM Δ (162-579), *StuI*; BBM Δ (98-579), *HindIII*; BBM Δ (1-161/390-579), *StuI* /*HaeII*; BBM Δ (1-96/162-579), *HindIII* /*StuI*; BBM Δ (1-390/514-579), *HaeII* /*HpaI*.

A yeast two-hybrid screen for BBM-interacting proteins was performed based on the Sos-recruitment system (Aronheim et al., 1997) using CytoTrap vectors (Stratagene), and a yeast two-hybrid cDNA library constructed using mRNA from *B. napus* microspore-derived embryos at

globular and torpedo stage (pMyr vector, CytoTrap XR library kit, Stratagene). The *B. napus* *BBM1* (Bn*BBM*) cDNA was cloned in frame with the hSos protein to make *pSos:BBM*. Yeast transformation, library screening and interaction verification were performed according to the manufacturer's instructions (<http://www.sanfordresearch.org/ClassLibrary/Page/Images/files/Cytotrap%20manual.pdf>).

The BD and AD Gal4 plasmids (pDEST32 and pDEST22, Invitrogen) and the PJ69-4A and PJ69-4 α yeast strains (James et al., 1996) were used in the Gal4 yeast two-hybrid experiments. Yeast transformation and selection of transformed colonies was performed as in the Clontech Yeast Protocols Handbook (<http://www.clontech.com>; Protocol No. PT3024-1). The coding regions of Arabidopsis *HDG* and *AIL* cDNAs (supplementary material Table S3) were used to generate the *BD-HDG* and *AD-AIL* constructs, respectively, via GW cloning. The *HDG1* promoter fragment was placed downstream of the AD coding region by GW cloning to create a stop codon immediately after the GW linker.

Western blotting

Western blot analysis of GR-HDG1 was performed essentially as previously described (Lamb and Irish, 2003), except that the samples were run on a SDS/10% polyacrylamide gel, and the blot was blocked in 1X phosphate-buffered saline (PBS), 0.3% Tween-20 and 3% non-fat powdered milk for one and a half hours. The blot was incubated with the primary antibody (anti-GR at 1:1000, Pierce Antibodies; PA1-516) at room temperature for 3 hours, and with the secondary antibody (goat anti-rabbit IgG conjugated to peroxidase, Rockland 611-1302) for 1 hour. The blot was developed using Stable Peroxide and Luminol/Enhancer solutions (Thermo Scientific) and exposed to film.

The loading control gel was fixed for a few minutes using an acetic acid:ethanol:water (1:4:5) solution, stained overnight with Coomassie Brilliant Blue R 250 and de-stained by washing several times with 0.3% Tween-20.

Whole-mount imaging

hdg mutant seeds were incubated over-night in HCG solution (80 g chloral hydrate, 10 ml glycerol, 30 ml water) on microscopy slides, prior to observation using a Nikon Optiphot microscope equipped with differential interference contrast (DIC) optics.

ACKNOWLEDGEMENTS

We thank Taku Takahashi for providing the *hdg* mutants, Renze Heidstra for the *pBBM::BBM-YFP* line, John Bowman for the YJ158 line, Jan-Willem Borst for advice concerning FRET-FLIM, Adriaan van Aelst and Tiny Franssen-Verheijen for performing the cryo-SEM analyses, and Iris Heidmann and Mieke Weemen for assistance. This project was funded by a Technology Top Institute - Green Genetics grant to K.B. and a Science and Technology Agency of Japan fellowship to H. F.

AUTHOR CONTRIBUTIONS

A.H., K.B., G.A., H.F. and R.I. designed the research; A.H., H.F., P.P., C.G. performed research; G.S-P, L.N. and J.M.M. analyzed data and; A.H. and K.B. wrote the paper.

REFERENCES

- Abe, M., Katsumata, H., Komeda, Y., and Takahashi, T.** (2003). Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. *Development* **130**, 635-643.
- Agusti, J., Lichtenberger, R., Schwarz, M., Nehlin, L., and Greb, T.** (2011). Characterization of transcriptome remodeling during cambium formation identifies MOL1 and RUL1 as opposing regulators of secondary growth. *Plos Genet* **7**, e1001312.
- Ahn, J.W., Kim, M., Lim, J.H., Kim, G.T., and Pai, H.S.** (2004). Phytocalpain controls the proliferation and differentiation fates of cells in plant organ development. *The Plant journal : for cell and molecular biology* **38**, 969-981.
- Aida, M., Beis, D., Heidstra, R., Willemssen, 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.
- Aronheim, A.** (1997). Improved efficiency Sos recruitment system: expression of the mammalian GAP reduces isolation of Ras GTPase false positives. *Nucleic Acids Res* **25**, 3373-3374.
- Becraft, P.W., Li, K.J., Dey, N., and Asuncion-Crabb, Y.** (2002). The maize dek1 gene functions in embryonic pattern formation and cell fate specification. *Development* **129**, 5217-5225.
- Benjamini, Y., and Hochberg, Y.** (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300.
- Boutillier, K., Offringa, R., Sharma, V.K., Kieft, 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. *The 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.
- Carvalho, B.S., and Irizarry, R.A.** (2010). A framework for oligonucleotide microarray preprocessing. *Bioinformatics* **26**, 2363-2367.
- Clough, S.J., and Bent, A.F.** (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant journal : for cell and molecular biology* **16**, 735-743.
- Custers, J.M., Cordewener, J.G., Nöllen, Y., Dons, H.M., and Lockeren Campagne, M.** (1994). Temperature controls both gametophytic and sporophytic development in microspore cultures of Brassica napus **13**, 267-271.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R.** (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant physiology* **139**, 5-17.
- Dewitte, W., Riou-Khamlichi, C., Scofield, S., Healy, J.M., Jacquard, A., Kilby, N.J., and Murray, J.A.** (2003). Altered cell cycle distribution, hyperplasia, and inhibited differentiation in Arabidopsis caused by the D-type cyclin CYCD3. *The Plant cell* **15**, 79-92.
- Dewitte, W., Scofield, S., Alcasabas, A.A., Maughan, S.C., Menges, M., Braun, N., Collins, C., Nieuwland, J., Prinsen, E., Sundaresan, V., and Murray, J.A.** (2007). Arabidopsis CYCD3 D-type cyclins link cell proliferation and endocycles and are rate-limiting for cytokinin responses. *Proc Natl Acad Sci U S A* **104**, 14537-14542.
- Eshed, Y., Izhaki, A., Baum, S.F., Floyd, S.K., and Bowman, J.L.** (2004). Asymmetric leaf development and blade expansion in Arabidopsis are mediated by KANADI and YABBY activities. *Development* **131**, 2997-3006.

- Fatouros, N.E., Lucas-Barbosa, D., Weldegergis, B.T., Pashalidou, F.G., van Loon, J.J., Dicke, M., Harvey, J.A., Gols, R., and Huigens, M.E.** (2012). Plant volatiles induced by herbivore egg deposition affect insects of different trophic levels. *Plos One* **7**, e43607.
- Fukuoka, H., Ogawa, T., Matsuoka, M., Ohkawa, Y., and Yano, H.** (1998). Direct gene delivery into isolated microspores of rapeseed (*Brassica napus* L) and the production of fertile transgenic plants. *Plant cell reports* **17**, 323-328.
- 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.
- Hacham, Y., Holland, N., Butterfield, C., Ubeda-Tomas, S., Bennett, M.J., Chory, J., and Savaldi-Goldstein, S.** (2011). Brassinosteroid perception in the epidermis controls root meristem size. *Development* **138**, 839-848.
- Hofhuis, H., Laskowski, M., Du, Y., Prasad, K., Grigg, S., Pinon, V., and Scheres, B.** (2013). Phyllotaxis and rhizotaxis in Arabidopsis are modified by three PLETHORA transcription factors. *Current biology : CB* **23**, 956-962.
- Holmes, K., Williams, C.M., Chapman, E.A., and Cross, M.J.** (2010). Detection of siRNA induced mRNA silencing by RT-qPCR: considerations for experimental design. *BMC research notes* **3**, 53.
- Horstman, A., Willemsen, V., Boutilier, K., and Heidstra, R.** (2014a). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends in plant science* **19**, 146-157.
- Horstman, A., Tonaco, I.A., Boutilier, K., and Immink, R.G.** (2014b). A cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies. *International journal of molecular sciences* **15**, 9628-9643.
- Immink, R.G., Gadella, T.W., Jr., Ferrario, S., Busscher, M., and Angenent, G.C.** (2002). Analysis of MADS box protein-protein interactions in living plant cells. *Proc Natl Acad Sci U S A* **99**, 2416-2421.
- James, P., Halladay, J., and Craig, E.A.** (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436.
- Jin, P., Guo, T., and Becraft, P.W.** (2000). The maize CR4 receptor-like kinase mediates a growth factor-like differentiation response. *Genesis* **27**, 104-116.
- Karimi, M., Inze, D., and Depicker, A.** (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends in plant science* **7**, 193-195.
- Karlova, R., Rosin, F.M., Busscher-Lange, J., Parapunova, V., Do, P.T., Fernie, A.R., Fraser, P.D., Baxter, C., Angenent, G.C., and de Maagd, R.A.** (2011). Transcriptome and metabolite profiling show that APETALA2a is a major regulator of tomato fruit ripening. *The Plant cell* **23**, 923-941.
- Kaufmann, K., Muino, J.M., Osteras, M., Farinelli, L., Krajewski, P., and Angenent, G.C.** (2010). Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nature protocols* **5**, 457-472.
- Knauer, S., Holt, A.L., Rubio-Somoza, I., Tucker, E.J., Hinze, A., Pisch, M., Javelle, M., Timmermans, M.C., Tucker, M.R., and Laux, T.** (2013). A protodermal miR394 signal defines a region of stem cell competence in the Arabidopsis shoot meristem. *Dev Cell* **24**, 125-132.
- Krizek, B.A.** (1999). Ectopic expression of AINTEGUMENTA in Arabidopsis plants results in increased growth of floral organs. *Developmental genetics* **25**, 224-236.
- Krizek, B.A., and Eaddy, M.** (2012). AINTEGUMENTA-LIKE6 regulates cellular differentiation in flowers. *Plant molecular biology* **78**, 199-209.

- Lamb, R.S., and Irish, V.F.** (2003). Functional divergence within the APETALA3/PISTILLATA floral homeotic gene lineages. *Proc Natl Acad Sci U S A* **100**, 6558-6563.
- Li, R., Yu, C., Li, Y., Lam, T.W., Yiu, S.M., Kristiansen, K., and Wang, J.** (2009). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966-1967.
- Livak, K.J., and Schmittgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.
- Melaragno, J.E., Mehrotra, B., and Coleman, A.W.** (1993). Relationship between Endopolyploidy and Cell Size in Epidermal Tissue of Arabidopsis. *The Plant cell* **5**, 1661-1668.
- Mordhorst, A.P., Voerman, K.J., Hartog, M.V., Meijer, E.A., van Went, J., Koornneef, M., and de Vries, S.C.** (1998). Somatic embryogenesis in Arabidopsis thaliana is facilitated by mutations in genes repressing meristematic cell divisions. *Genetics* **149**, 549-563.
- Mudunkothge, J.S., and Krizek, B.A.** (2012). Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function. *The Plant journal : for cell and molecular biology* **71**, 108-121.
- Muino, J.M., Kaufmann, K., van Ham, R.C., Angenent, G.C., and Krajewski, P.** (2011). ChIP-seq Analysis in R (CSAR): An R package for the statistical detection of protein-bound genomic regions. *Plant methods* **7**, 11.
- Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K.T., and Takahashi, T.** (2006). Characterization of the class IV homeodomain-leucine zipper gene family in Arabidopsis. *Plant physiology* **141**, 1363-1375.
- Nicol, J.W., Helt, G.A., Blanchard, S.G., Jr., Raja, A., and Loraine, A.E.** (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**, 2730-2731.
- Nobusawa, T., Okushima, Y., Nagata, N., Kojima, M., Sakakibara, H., and Umeda, M.** (2013). Restriction of cell proliferation in internal tissues via the synthesis of very-long-chain fatty acids in the epidermis. *Plant signaling & behavior* **8**.
- Nole-Wilson, S., Tranby, T.L., and Krizek, B.A.** (2005). AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant molecular biology* **57**, 613-628.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M.W., 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 molecular biology* **68**, 225-237.
- Peng, Y., Chen, L., Lu, Y., Ma, W., Tong, Y., and Li, Y.** (2013). DAR2 acts as an important node connecting cytokinin, auxin, SHY2 and PLT1/2 in root meristem size control. *Plant signaling & behavior* **8**.
- Peterson, K.M., Shyu, C., Burr, C.A., Horst, R.J., Kanaoka, M.M., Omae, M., Sato, Y., and Torii, K.U.** (2013). Arabidopsis homeodomain-leucine zipper IV proteins promote stomatal development and ectopically induce stomata beyond the epidermis. *Development* **140**, 1924-1935.
- Roeder, A.H.K., Cunha, A., Ohno, C.K., and Meyerowitz, E.M.** (2012). Cell cycle regulates cell type in the Arabidopsis sepal. *Development* **139**, 4416-4427.
- Russinova, E., Borst, J.W., Kwaaitaal, M., Cano-Delgado, A., Yin, Y., Chory, J., and de Vries, S.C.** (2004). Heterodimerization and endocytosis of Arabidopsis brassinosteroid receptors BRI1 and AtSERK3 (BAK1). *The Plant cell* **16**, 3216-3229.
- San-Bento, R., Farcot, E., Galletti, R., Creff, A., and Ingram, G.** (2013). Epidermal identity is maintained by cell-cell communication via a universally active feedback loop in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*.

- Schwab, R., Ossowski, S., Riester, M., Warthmann, N., and Weigel, D.** (2006). Highly specific gene silencing by artificial microRNAs in Arabidopsis. *The Plant cell* **18**, 1121-1133.
- Shepard, A.R., Jacobson, N., and Clark, A.F.** (2005). Importance of quantitative PCR primer location for short interfering RNA efficacy determination. *Analytical biochemistry* **344**, 287-288.
- Smaczniak, C., Immink, R.G.H., Muino, J.M., Blanvillain, R., Busscher, M., Busscher-Lange, J., Dinh, Q.D., Liu, S.J., Westphal, A.H., Boeren, S., Parcy, F., Xu, L., Carles, C.C., Angenent, G.C., and Kaufmann, K.** (2012). Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. *P Natl Acad Sci USA* **109**, 1560-1565.
- Soriano, M., Li, H., Jacquard, C., Angenent, G.C., Krochko, J., Offringa, R., and Boutilier, K.** (2014). Plasticity in Cell Division Patterns and Auxin Transport Dependency during in Vitro Embryogenesis in Brassica napus. *The Plant cell* **26**, 2568-2581.
- Takada, S.** (2013). Post-Embryonic Induction of ATML1-SRDX Alters the Morphology of Seedlings. *Plos One* **8**.
- Takada, S., Takada, N., and Yoshida, A.** (2013). ATML1 promotes epidermal cell differentiation in Arabidopsis shoots. *Development* **140**, 1919-1923.
- 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 molecular biology* **73**, 481-492.
- Whitford, R., Fernandez, A., De Groot, R., Ortega, E., and Hilson, P.** (2008). Plant CLE peptides from two distinct functional classes synergistically induce division of vascular cells. *Proc Natl Acad Sci U S A* **105**, 18625-18630.
- Yaginuma, H., Hirakawa, Y., Kondo, Y., Ohashi-Ito, K., and Fukuda, H.** (2011). A novel function of TDIF-related peptides: promotion of axillary bud formation. *Plant & cell physiology* **52**, 1354-1364.
- Zhong, S., Lin, Z., Fray, R.G., and Grierson, D.** (2008). Improved plant transformation vectors for fluorescent protein tagging. *Transgenic research* **17**, 985-989.

Chapter 4

BABY BOOM and PLETHORA2 induce somatic embryogenesis in a dose- and context-dependent manner via the LAFL pathway

Anneke Horstman¹, Iris Heidmann², Mieke Weemen¹,
Gerco Angenent^{1,3} and Kim Boutilier¹

¹ Wageningen University and Research Centre, Bioscience, Droevendaalsesteeg 1, 6708 PB
Wageningen, The Netherlands

² Enza Zaden Research and Development B.V, Haling 1-E, 1602 DB Enkhuizen, The Netherlands

³ Laboratory of Molecular Biology, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

ABSTRACT

Somatic embryogenesis (SE) is an example of cellular totipotency, where embryos develop from vegetative cells rather than from gamete fusion. The *AINTEGUMENTA-LIKE* (*AIL*) transcription factor family comprises eight genes, which redundantly regulate meristem identity and growth. Ectopic expression of the *AIL* genes *BABY BOOM* (*BBM*) or *PLETHORA5/AIL5*, is sufficient to induce SE in *Arabidopsis thaliana* seedlings, but the roles of the other *AIL* genes in this process, as well as the signalling pathways underlying *AIL*-mediated SE, are not known. Here, we show that overexpression of all *AIL* genes, except for the phylogenetically-distinct *AIL1* and *AINTEGUMENTA*, induces SE, suggesting extensive overlap in *AIL* function. Using *BBM* and *PLT2* as representatives of *AIL* function, we show that *AIL*-mediated SE is dose-dependent, where a relatively high dose induces SE and a relatively low dose induces shoot (*BBM*) or root (*PLT2*) organogenesis. *AIL*-induced SE is also context-dependent, as early expression of *BBM* or *PLT2* induces SE directly from seedling tissues, whereas late expression induces SE indirectly from callus. Analysis of *BBM* regulatory pathways shows that *BBM* binds to and regulates genes with roles in maintaining embryo identity and/or somatic embryo induction including the *LAF1* genes, *LEC1*, *LEC2*, *FUS3* and *ABI3*, as well as *AGL15*. Mutant analysis identified these genes as positive regulators of *BBM*-mediated SE and their chromatin mediated repressors *PKL* and *VAL1* as negative regulators. Our results demonstrate that *AIL* proteins regulate overlapping pathways in a context- and dose-dependent manner to modulate plant development and place *BBM* and *PLT2* upstream of other known inducers of SE.

INTRODUCTION

AINTEGUMENTA-LIKE (AIL) genes form a small clade of eight members within the AP2 group of APETALA2/ethylene-responsive element-binding factor (AP2/ERF) transcription factors (Kim et al., 2006), and comprise *AINTEGUMENTA (ANT)*, *AIL1*, *PLETHORA1 (PLT1)*, *PLT2*, *PLT3/AIL6*, *CHOTTO1 (CHO1)/EMBRYOMAKER (EMK)/PLT5/AIL5* (hereafter named *PLT5/AIL5*), *PLT7* and *BABY BOOM (BBM)*. *AIL* genes are expressed in dividing tissues, including root, shoot and floral meristems (Nole-Wilson et al., 2005), where they act in a redundant manner to maintain a meristematic state (reviewed in (Horstman et al., 2014)). Single knock-out mutants of *AIL* genes show only minor defects, but double or triple mutants have stronger phenotypes related to reduced cell proliferation or altered cell identity. For example, the *ant* single mutant has smaller floral organs with partial loss of identity, a phenotype that is enhanced in the *ant;plt3/ail6* double mutant (Klucher et al., 1996; Krizek, 2009; Sharma et al., 2013). Combinations of *plt1*, *plt2*, *plt3/ail6* and *bbm* mutants are embryo lethal (*plt2;bbm*), rootless (*plt1;plt2;plt3/ail6*) or have a short root (*plt1;plt2*) (Aida et al., 2004; Galinha et al., 2007), and the *ant;plt3/ail6;plt7* triple mutant is impaired in shoot meristem maintenance (Mudunkothge and Krizek, 2012a).

In line with their loss-of-function phenotypes, overexpression of *AIL* transcription factors induces cell overproliferation phenotypes. Ectopic overexpression of *PLT5/AIL5* promotes somatic embryo and ectopic organ formation on seedlings (Boutilier et al., 2002; Tsuwamoto et al., 2010), while overexpression of *PLT1* and *PLT2* leads to ectopic development of hypocotyls, roots and quiescent centre cells (Aida et al., 2004). Besides promoting enhanced pluripotency and totipotency, *AIL* overexpression can also lead to an enlarged root meristem (*PLT2*) (Galinha et al., 2007) and to increased floral organ size due to increased cell number, as shown for *ANT*, *PLT5/AIL5* and *PLT3/AIL6* overexpression (Krizek, 1999; Nole-Wilson et al., 2005; Krizek and Eaddy, 2012). In contrast, sepals of seedlings expressing higher levels of *PLT3/AIL6* are small and undifferentiated, suggesting that high *PLT3/AIL6* levels inhibit cell differentiation (Krizek and Eaddy, 2012).

Genetic analysis shows both specific and overlapping roles for *AIL* genes, and that *AIL* proteins can partially or fully complement phenotypes of other *ail* mutants (Galinha et al., 2007), but it has been difficult to assign specific *AIL* functions based on the overexpression studies. *AIL* genes that show redundancy in loss-of-function studies, such as *BBM* and *PLT2*, do not show the same overexpression phenotypes (Boutilier et al., 2002; Aida et al., 2004), while overexpression

of the same gene e.g. *PLT5/AIL5*, can result in different overexpression phenotypes (Nole-Wilson et al., 2005; Yano et al., 2009; Tsumamoto et al., 2010). Whether these different phenotypes are due to differences in the expression level of the transgene or due to the screening approach is not clear. AIL target genes have only been identified for BBM (Passarinho et al., 2008), thus it is not known whether AIL proteins have the same or partially overlapping target genes.

Here, we focus on the role of *AIL* genes in somatic embryo induction. Besides AIL proteins, a number of other transcription factors have been identified that can induce or enhance somatic embryogenesis (SE) when ectopically expressed (Feher, 2014). These include two *LEAFY COTYLEDON 1 (LEC1)/LEC1-LIKE*; *ABSCISIC ACID (ABA)-INSENSITIVE3 (ABI3)*; *FUSCA3 (FUS3)*; *LEC2 (LAFL)* seed maturation genes (Jia et al., 2013), *LEC1* and *LEC2* (Lotan et al., 1998; Stone et al., 2001), and the MADS-domain transcription factor *AGAMOUS-LIKE15 (AGL15)* (Harding et al., 2003; Zheng et al., 2009). The developmental programs regulated by *AGL15* and *LEC2* have been well characterized and their pathways are interconnected, as *LEC2* and *AGL15* positively regulate each other's function (Braybrook et al., 2006; Zheng et al., 2009). Similar embryogenic phenotypes are observed in loss-of-function mutants of epigenetic regulators, including the CHD3 protein *PICKLE (PKL)* (Ogas et al., 1999), the B3-domain proteins *VP1/ABI3-LIKE1 (VAL1)* and *VAL2* (Suzuki et al., 2007), and the Polycomb Group proteins *CURLY LEAF (CLF)*, *SWINGER (SWN)*, *EMBRYONIC FLOWER2 (EMF2)*, *VERNALIZATION2 (VRN2)*, and *FERTILIZATION INDEPENDENT ENDOSPERM (FIE)* (Chanvivattana et al., 2004; Bouyer et al., 2011), which function to repress *LAFL* gene expression during the transition to post-embryonic growth.

Here, we show that the BBM clade of AIL proteins are potent inducers of SE that this function is dose- and context-dependent. In addition, we show that that AIL-induced SE is mediated in part by direct activation of *LAFL* genes and indirect activation of other components of the *LAFL* network.

MATERIALS AND METHODS

Plant material and growth conditions

The *lec2-1* (CS3868), *lec1-2* (CS3867), *fus3-3* (CS8014), *agl15-3* (CS16479), *fie* (SALK_042962), and *pkl-1* (CS3840) mutants were obtained from Nottingham Arabidopsis Stock Centre. The *val1-2* (*hsi2-5*), *val1-2;val2-1*, *abi3-8*, *abi3-9*, *abi3-10* and *abi5-7* mutants have been previously described (Nambara et al., 2002; Suzuki et al., 2007; Sharma et al., 2013). The *LEC1::LEC1-GFP* (Li et al., 2014) marker and the *35S::BBM* and *35S::BBM-GR* constructs were described previously (Boutilier et al., 2002; Passarinho et al., 2008). The *35S::BBM-GR* construct was introduced into the mutant lines by floral dip transformation (Clough and Bent, 1998).

Seeds were sterilized with liquid bleach (1 minute in 70% ethanol, followed by 20 minutes in commercial bleach (4%) containing 0.03% Tween-20, and then washed 4-5 times with sterile MilliQ water) before plating on solid medium (½MS-10: half-strength Murashige and Skoog salts and vitamins, pH 5.8, with 0.8% agar and 1% sucrose). Embryo rescue of the *lec1-2* mutant was performed by culturing ovules from sterilized siliques on solid ½MS-10 medium. For some experiments, sterilized seeds were dispensed in 190 ml containers (Greiner) with 30 ml liquid ½MS-10 medium. DEX and CHX (both Sigma) were added to the medium as described in the text. Solid and liquid (rotary shaker, 60 rpm/min) cultures were kept at 21 °C and 25 °C, respectively (16 hour light/8 hour dark regime). Plants were grown for seed collection at 21 °C (16h light/8h dark regime) on rockwool plugs (Grodan) supplemented with 1 g/L Hyponex fertilizer.

Vector construction and transformation

The *ANT*, *PLT3/AIL6*, *PLT7* and *PLT1* protein coding regions were amplified from Arabidopsis Col-0 genomic DNA and the *PLT2* protein coding region from cDNA, using the primers listed in Supplemental Table 1. The DNA fragments were cloned into the Gateway (GW) binary vector pGD625, which contains a double-enhanced cauliflower mosaic virus 35S promoter and an AMV translational enhancer (Immink et al., 2002). *BBM-GFP* was amplified from the *BBM::BBM-GFP* plasmid described in Chapter 3. The GW-compatible destination vector pARC146 (Danisman et al., 2012) was used for inducible ectopic activity of *PLT2* and *BBM-GFP*. This vector contains a double-enhanced cauliflower mosaic virus 35S promoter and an AMV translational enhancer, as well as the coding region of the ligand binding domain of the rat glucocorticoid receptor (GR) downstream of the GW cassette.

Leaf imaging and quantification of stomatal development

The first leaf pairs of nine day-old untreated or 0.1 μ M DEX-treated *35S::BBM-GR* seedlings were placed overnight in 70% ethanol at 4 °C, then transferred to 85% ethanol for 6 hours, and subsequently to 3% bleach overnight or until imaging. Leaves were mounted in HCG solution (80 g chloral hydrate, 10 ml glycerol, 30 ml water) prior to imaging with a Nikon Optiphot microscope.

The stomatal, meristemoid and stomatal lineage indices (SI, MI and SLI) were calculated as previously described (Peterson et al., 2013): $SI = (\text{number of stomata} / (\text{total number of stomata} + \text{non-stomatal epidermal cells})) \times 100$. For the SI, only mature stomata with a pore were counted. $MI = (\text{number of meristemoids} / (\text{total number of stomata} + \text{non-stomatal epidermal cells})) \times 100$. $SLI = (\text{number of stomata and stomata precursors} / (\text{total number of stomata} + \text{non-stomatal epidermal cells})) \times 100$.

Tissue sectioning

35S::BBM-GR and *35S::PLT2-GR* seedlings were fixed overnight in 3:1 ethanol (100%):acetic acid and dehydrated stepwise from 70 to 100% ethanol. The samples were then infiltrated in Technovit 7100 (including hardener 1) in three steps (Heraeus Kulzer, Germany), followed by Technovit 7100 plus hardeners 1 and 2 (Heraeus Kulzer, Germany). Four micron-thick sections were prepared using a rotary microtome (Zeiss HM340E) and Technovit blades (Adamas, The Netherlands). Sections were stained with 0.05% Toluidine Blue (Merck, Germany) for three minutes, and then rinsed well with water and air-dried. The sections were mounted in Euparal (Roth, Germany) and images were taken using an IX70 microscope (Olympus) with a DP70 camera and CellSens software (Olympus). Seven to ten seedlings per line per treatment were observed.

Confocal laser scanning microscopy

LEC1::LEC1-GFP seedlings were fixed for one week at 4 °C in 1x microtubule stabilizing buffer (MTSB: 50mM PIPES, 5mM MgSO₄, 5 mM EGTA, pH7.4) containing 4% paraformaldehyde. Fixed seedlings were washed three times with 0.2x MTSB and mounted in the same buffer containing 1% glycerol prior to imaging. Roots were counterstained with 10 μ g/mL propidium iodide (PI). Confocal laser scanning microscopy was performed with a Leica SPE DM5500 upright microscope using the LAS AF 1.8.2 software. GFP was excited with a 488-nm solid-state laser and its emission was detected at a band width of 500–530 nm. PI (roots) and red autofluorescence

(cotyledons) were used as a background signals (excited with a 532 nm laser and detected at 600-800 nm).

ChIP-seq

ChIP-seq experiments and data analysis were carried out as described in Chapter 3. Somatic embryo material generated from either 2,4-dichlorophenoxyacetic acid (2,4-D)-induced cultures or from a *BBM* overexpression line were used for ChIP. Somatic embryos from a *BBM::NLS-GFP* line, or embryogenic *35S::BBM* seedlings served as negative controls for the *BBM::BBM-YFP* and *35S::BBM-GFP* ChIPs, respectively. ChIP-seq results were visualized using Integrated Genome Browser (IGB) 8.1.11 (Nicol et al., 2009). The ChIP-seq data is available via NCBI (GEO accession: GSE52400).

Expression analysis of *BBM/PLT2* target genes

One- and five-day-old Col-0, *35S::BBM-GR* and *35S::PLT2-GR* seedlings (3 biological replicates of each) were treated for 3 hours with 10 μ M DEX plus 10 μ M CHX. RNA was extracted using the NucleoSpin RNA kit (Machery-Nagel) kit in combination with Plant RNA Isolation Aid (Ambion), treated with *DNA-free* (Ambion) and then used for cDNA synthesis with M-MLV Reverse Transcriptase (Life Technologies). Quantitative real-time RT-PCR (qPCR) analysis of *BBM/PLT2* target genes was performed using the BioMark HD System (Fluidigm) as described in Chapter 3. The data were normalized against the *SAND* gene (Czechowski et al., 2005) and relative gene expression was calculated according to Livak and Schmittgen (Livak and Schmittgen, 2001) by comparison with DEX + CHX-treated wild-type Col-0. The DNA primers are shown in Supplemental Table 1.

RESULTS

All BBM and PLT proteins induce SE

BBM and *PLT2* have redundant roles in embryogenesis and root meristem maintenance (Galinha et al., 2007), but show different overexpression phenotypes (Boutillier et al., 2002; Galinha et al., 2007; El Ouakfaoui et al., 2010). This observation, together with reported differences in the overexpression phenotypes described for the same *AIL* gene (*PLT5/AIL5*) (Nole-Wilson et al., 2005; Yano et al., 2009; Tsuwamoto et al., 2010), prompted us to investigate the overexpression phenotypes of the *AIL* family members using the same overexpression vector and under the same growth conditions. We generated *Arabidopsis* *35S::AIL* overexpression lines for the six *AIL* genes that have not been reported to induce SE when overexpressed, namely *ANT*, *AIL1*, *PLT1*, *PLT2*, *PLT3/AIL6* and *PLT7*, and found that overexpression of all these genes except the phylogenetically-distinct *ANT1* and *AIL1* (Kim et al., 2006) induced somatic embryogenesis in 7-26% of the primary transformants (Supplemental Figure 1, Supplemental Table 2). These numbers are in line with the percentage of embryogenic seedlings obtained after transformation with *35S::BBM* (Supplemental Table 2). We observed the large flower phenotype that has been reported previously for *35S::ANT* (Krizek, 1999; Mizukami and Fischer, 2000), demonstrating that the protein is expressed, but did not observe the previously reported conversion of the shoot apical meristem (SAM) into root identity in *PLT1* or *PLT2* overexpression lines (Aida et al., 2004; Galinha et al., 2007), neither in the primary transformants nor in subsequent generations. No mutant phenotypes were observed upon *AIL1* overexpression. These results show that all *AIL* proteins, except for *ANT* and *AIL1*, have the capacity to induce SE, and suggest that all *BBM*-clade proteins are functionally interchangeable with respect to somatic embryo induction.

BBM and *PLT2* have dose-dependent overexpression phenotypes

PLT2 functions in a dose-dependent manner in the root, with different levels of *PLT2* protein instructing different cellular outputs (Galinha et al., 2007). We employed fusions between two representative *AIL* proteins, *BBM* and *PLT2*, and the glucocorticoid receptor ligand-binding domain (GR, *35S::AIL-GR*) to investigate the dose-dependency of *AIL* overexpression phenotypes. The amount of nuclearly localized *BBM*-GFP-GR protein could be controlled by the DEX concentration. In the absence of DEX, GFP was localized to the cytoplasm, but became increasingly nuclear-localized with higher DEX concentrations, such that cytoplasmic GFP could no longer be detected in the presence of 1 μ M DEX (Supplemental Figure 2). These experiments

demonstrated that the proportion of a nuclear-localized GR fusion protein, and by extension AIL-GR protein, can be controlled by exposing plant tissue to different amounts of DEX.

We used the same DEX concentration range to regulate AIL-GR activity in *35S::AIL-GR* seedlings (Figure 1). Control seedlings (wild-type seedlings + DEX) did not show aberrant phenotypes when grown on DEX, whereas *35S::BBM-GR* and *35S::PLT2-GR* seedlings showed dose-dependent mutant phenotypes. The DEX concentration required to induce a specific phenotype was dependent on the strength of the transgenic line.

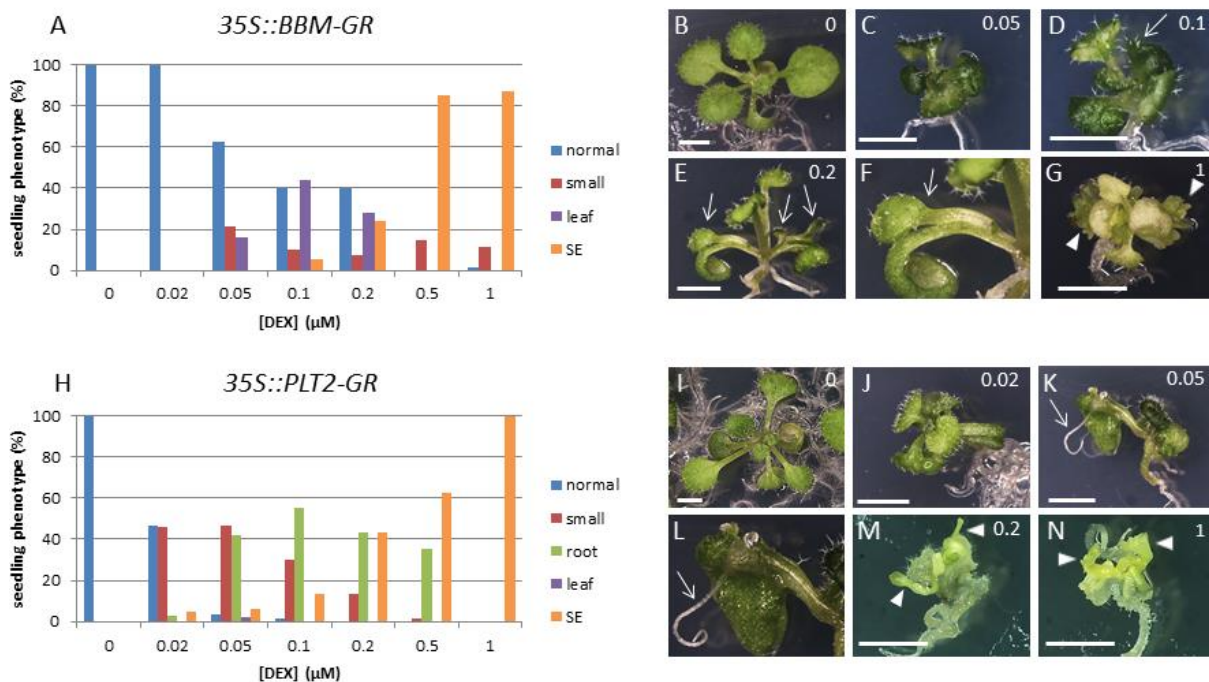


Figure 1. BBM and PLT2 have dose-dependent overexpression phenotypes

35S::BBM-GR and *35S::PLT2-GR* seedlings were grown for two weeks on medium containing different DEX concentrations. Frequency of *35S::BBM-GR* (A) and *35S::PLT2-GR* (H) phenotypes ($n=62$ to 196 seedlings). No additional phenotypes were observed in treatments above 1 μM DEX. Leaf, ectopic leaves; root, ectopic root; SE, somatic embryogenesis.

(B-G) Representative phenotypes of *35S::BBM-GR* seedlings grown on the DEX concentration (μM) indicated in each picture. (B) A normal looking seedling grown without DEX. (C) A small seedling showing epinastic growth of leaves and cotyledons. (D) A small, epinastic seedling with a trichome-bearing ectopic leaf (arrow) on the cotyledon petiole. (E) A seedling with ectopic leaves on the petioles of both cotyledons (arrows). (F) A magnified view of the ectopic leaf (arrow) in (E). (G) A seedling with somatic embryos on the cotyledon margins (arrowheads).

(I-N) Representative phenotypes of *35S::PLT2-GR* seedlings grown on the DEX concentration (μM) indicated in each picture. (I) A wild-type seedling grown in the absence of DEX. (J) A small seedling showing epinastic growth of leaves and cotyledons. (K) A small epinastic seedling with ectopic root formation on the cotyledon (arrow). (L) A magnified view of the ectopic root (arrow) shown in (K). (M, N) Seedlings with somatic embryos on the cotyledons (arrowheads). Scale bars represent 2.5 mm.

The dose-dependent phenotypes of strong AIL-GR lines (i.e. lines that show highly penetrant SE at a high DEX dose) are shown in Figure 1. At the lowest effective DEX concentrations *35S::BBM-GR* seedlings were stunted, with epinastic leaves (Figure 1A). Analysis of the first leaf pair and stomatal development suggested that a low BBM/PLT2 dose inhibits cell differentiation (Supplemental Figure 3). At intermediate DEX concentrations the seedlings were still small, but now formed leaf-like structures from their cotyledon petioles, which ranged from trichome-bearing protrusions (Figure 1D) to ectopic leaves (Figure 1E, F). At the highest effective DEX concentration, *35S::BBM-GR* seedlings also developed somatic embryos on their cotyledons (Figure 1A, G) (Passarinho et al., 2008). *35S::PLT2-GR* seedlings also showed stunted growth and somatic embryo formation at the lowest and highest effective DEX concentrations tested, respectively, but ectopic roots were more prevalent than shoots at intermediate DEX concentrations (Figure 1H, K, L). Phenotypically weaker *35S::BBM-GR* and *35S::PLT2-GR* transgenic lines showed a similar dose-dependent response, but the penetrance and severity of the phenotypes was lower (Supplemental Figure 4). For example, although the number of SE-forming seedlings was high in these weaker lines, they only produced a few somatic embryos at the tip of the cotyledon.

Our data suggest that BBM and PLT2 overexpression phenotypes are dose-dependent, with similar phenotypes at relatively low (stunted) and high doses (embryogenesis) and divergent phenotypes at an intermediate dose (shoot or root organogenesis).

BBM and PLT2 promote context-specific embryogenesis

Previously, we showed that there is an optimal developmental window for BBM-mediated SE; a significant drop in the number of seedlings that form somatic embryos is observed when DEX is added four days after seed germination (Passarinho et al., 2008). We examined this developmental competence in more detail by activating BBM-GR and PLT2-GR at different time points before and after germination. Germination is defined as the emergence of the radicle through the surrounding structures (Bewley, 1997) and is a two-step process in *Arabidopsis*, comprising testa rupture (d1) followed by radicle protrusion through the endosperm (endosperm rupture, d2).

When *35S::BBM-GR* seeds were placed directly in DEX-containing medium prior to or at endosperm rupture (d0-d2), 100% of the seedlings formed somatic embryos directly on their cotyledons after circa one week (Figure 2A; Figure 3A). In contrast, post-germination DEX treatment (d3-d4) induced callus formation on the adaxial side of the cotyledons, from which

visible somatic embryos developed approximately 14 days after BBM activation (ca. 40%; Figure 2A; Figure 3A). *35S::PLT2-GR* seedlings treated with DEX at the same time points, showed similar phenotypes (Figure 2B; Figure 3B) with two exceptions. Firstly, when *35S::PLT2-GR* seedlings were DEX-treated at endosperm rupture (d2), they did not form somatic embryos directly from the cotyledon as for BBM, but rather formed a whitish protrusion at the SAM that contained leaf-like tissue on its distal end (Figure 2B; Figure 3B), which developed somatic embryos 12 days after PLT2 activation (Figure 2B; Figure 3B).

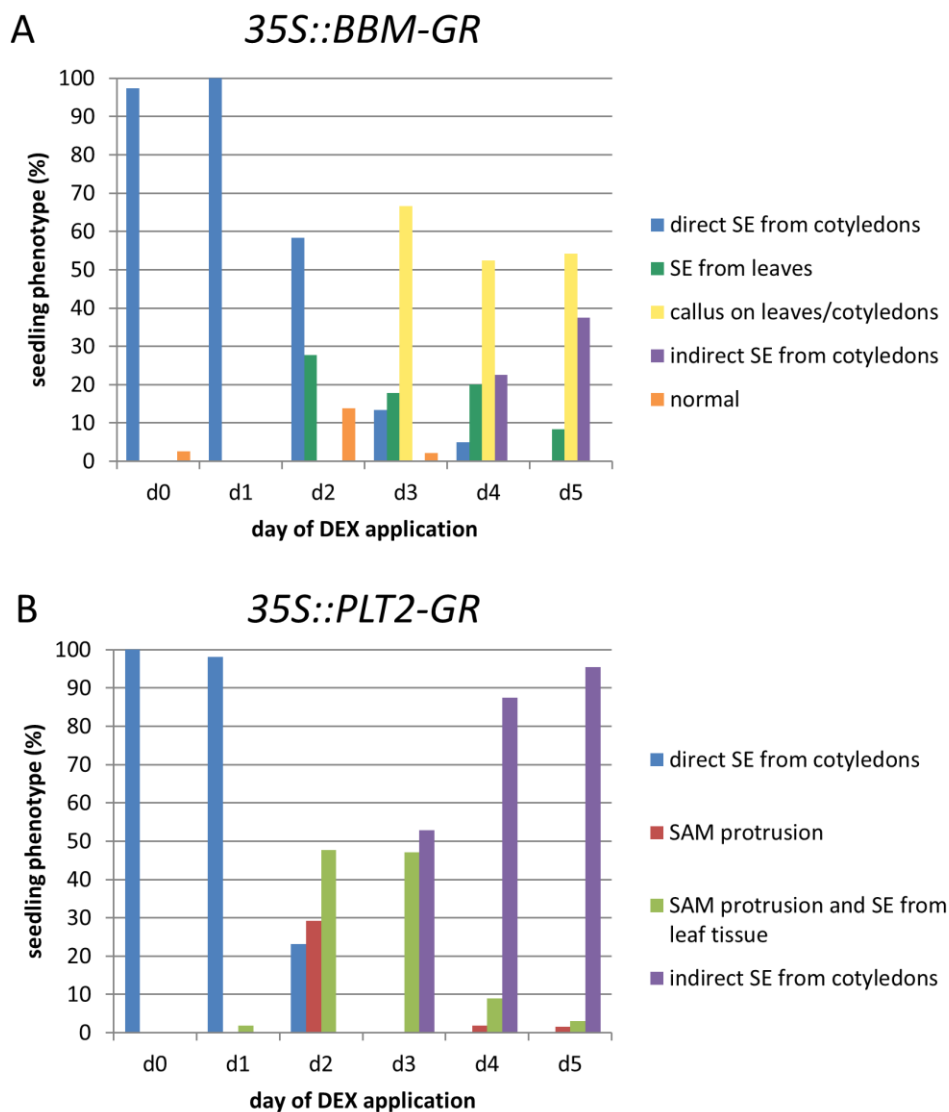


Figure 2. BBM and PLT2 induce stage-specific phenotypes

The effect of applying a relatively high BBM (A) or PLT2 dose (B) at different time points after sowing (d0-d5). The seedling phenotypes were scored two weeks after the DEX application. For each time point, between 31 and 70 seedlings were analysed. The quantification is shown for single *35S::BBM-GR* and *35S::PLT2-GR* lines, but similar results were obtained with other independent lines (Supplemental Figure 4). SE, somatic embryogenesis, SAM, shoot apical meristem.

Secondly, post-germination (d3-d4) DEX treatment of *35S::PLT2-GR* plants induced callus and somatic embryo formation on both the petioles and the cotyledons (Figure 2B; Figure 3B). These results suggest that the response to BBM and PLT2 ectopic expression depends on the developmental context in which the proteins are expressed.

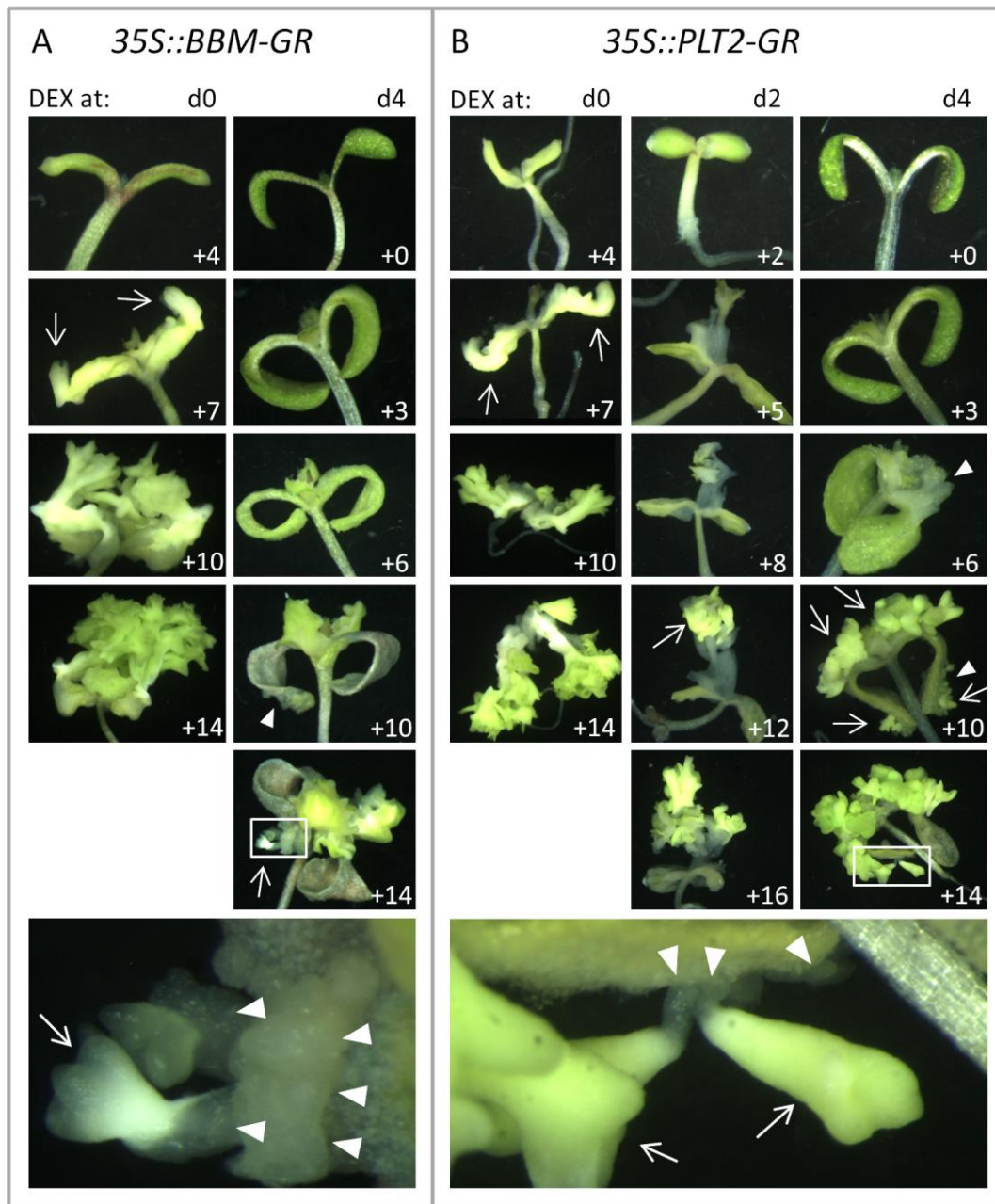


Figure 3. BBM and PLT2 promote context-specific embryogenesis

35S::BBM-GR (A) and *35S::PLT2-GR* (B) lines were treated with 10 μ M DEX at different time points (d0, d2, and d4), and the development of the seedlings was followed in time. The culture time after DEX application is indicated on the bottom right of each picture. The images at different individuals are from different individuals. The arrowheads and arrows indicate callus and somatic embryos/embryogenic tissue, respectively. The lower-most images in (A) and (B) are magnifications of the boxed regions in the respective '+14' images, and show the indirect development of somatic embryos from callus.

The timing and origin of somatic embryo formation in *35S::BBM-GR* and *35S::PLT2-GR* seedlings was examined in more detail using tissue sections. *35S::BBM-GR* seedlings that were DEX-induced at d0 showed anti- and periclinal cell division in the sub-epidermal layers on the adaxial side of the cotyledon, resulting in the formation of small cells at the position where elongated palisade cells are found in wild-type seedlings (Figure 4A).

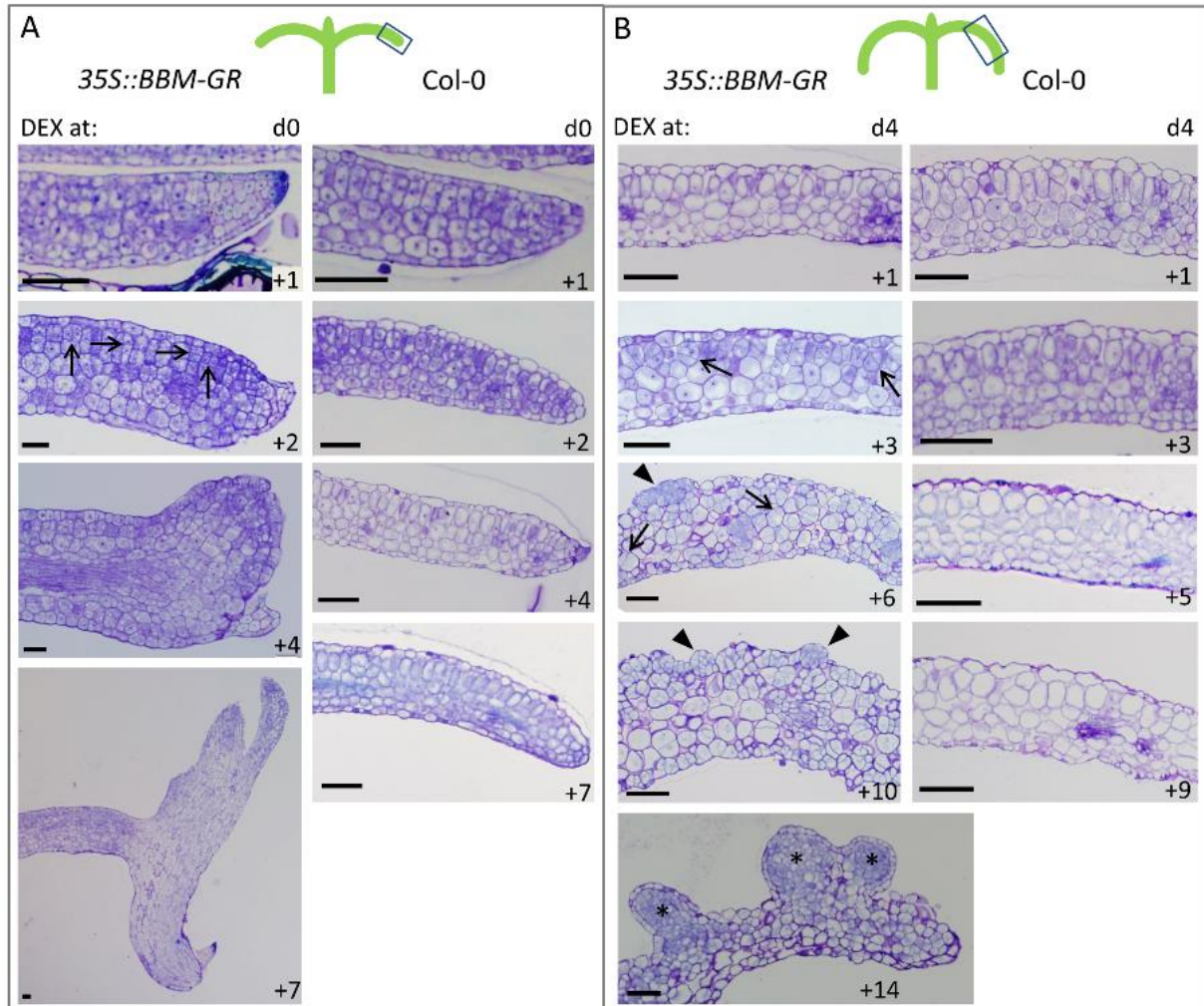


Figure 4. BBM induces direct and indirect SE

Sections of *35S::BBM-GR* and wild-type *Col-0* seedlings that were DEX-induced at d0 (**A**) or d4 (**B**), and cultured for the additional period indicated at the bottom right of each image (+1 to +14 days). The schematic illustrations depict the cotyledon regions (blue boxes) that were sectioned in the images below. *BBM-GR* activation at both d0 and d4 induces anticlinal, periclinal and oblique cell divisions, indicated by the horizontal, vertical and oblique arrows, respectively, on the adaxial side of the cotyledon. *BBM-GR* activation at d0 (**A**) induces cell divisions (+1, +2), and thickening of the cotyledon tip (+4), followed by the direct development of a somatic embryo from this area (+7). By contrast, *BBM-GR* activation at d4 (**B**) induces oblique and less compact cell divisions (+1, +3) and the formation of more compact cell masses (arrowheads) from which globular somatic embryos with a distinct epidermis (asterisks) develop. Scale bars, 100 μm .

After about four days of BBM-GR activation, a bump formed on the tip of the cotyledon that developed into a bipolar somatic embryo a few days later (Figure 4A). Later, somatic embryos also developed on more proximal parts of the cotyledon and secondary embryos formed on the primary somatic embryo on the cotyledon tip (Figure 3A +10 and +14). PLT2-GR activation at d0 induced a similar developmental change (Figure 3B). BBM-GR activation at d4 predominantly induced oblique cell divisions in the subepidermal cell layers on the adaxial side of the cotyledon and did not induce cell division at the cotyledon tips (Figure 4B). Moreover, in contrast to early BBM induction, larger, irregularly-shaped, vacuolate cells were formed proximal to the tip, resulting in a rough cotyledon surface (Figure 4B). Small clusters of small, cytoplasm-rich cells were observed on the cotyledon surface around seven days after BBM activation (Figure 4B). Ten days after BBM-GR activation, we observed larger globular-shaped structures enclosed by a smooth epidermis, which were set off from the underlying tissue by a thicker cell wall. These structures are reminiscent of globular-stage somatic embryos (Figure 4B). We observed the same phenotype after post-germination PLT2-GR activation, although somatic embryos developed faster (Figure 5A). Notably, BBM-GR and especially PLT2-GR activation induced proliferation of the cotyledon vasculature (Figure 5B). Somatic embryos always formed above this tissue, but we did not observe a direct connection between the proliferating vascular tissue and the somatic embryos.

We conclude that AIL-mediated SE is induced in two ways depending on the developmental stage of the explant: directly from cotyledons in a narrow window surrounding germination, and indirectly via a callus phase after germination. The data imply that the developmental competence for SE relies on context-specific co-factors.

BBM activates embryogenesis regulators

To understand the regulatory networks underlying AIL-mediated SE, we identified genes that were directly bound by BBM during somatic embryo development. Chromatin immunoprecipitation coupled to next-generation sequencing (ChIP-seq) (Chapters 3 and 5) showed that BBM bound to the promoter regions of transcription factor genes that have roles in promoting zygotic and/or somatic embryo development, including the *LAFL* seed maturation genes, *LEC1*, *LEC2*, *ABI3* and *FUS3* (but not *LEC1-LIKE*), and the MADS box transcription factor *AGL15* (Figure 6A).

We examined whether BBM binding regulates the expression of these genes during direct and indirect SE by inducing one day-old (early, direct) and five day-old (late, indirect)

35S::BBM-GR and *35S::PLT2-GR* seedlings with DEX in the presence of the translational inhibitor cycloheximide (CHX) (Gorte et al., 2011) and examining target gene expression using quantitative RT-PCR (qPCR). Early activation of BBM/PLT2-GR was characterized by upregulation of *LEC1*, *LEC2*, *FUS3* and *ABI3* gene expression (Figure 6B). In contrast, expression of *LEC1*, *LEC2*, *FUS3* and *ABI3* was not detected in five day-old induced Col-0 seedlings, nor was it detected in DEX-induced *35S::BBM-GR* and *35S::PLT2-GR* seedlings (Figure 6B). *AGL15* expression was not much affected by BBM/PLT2-GR activation at either of the two time points (Figure 6B). It might be that *LEC* genes are in an epigenetically silent state in five day-old seedlings and only become accessible after re-differentiation of the cells into callus.

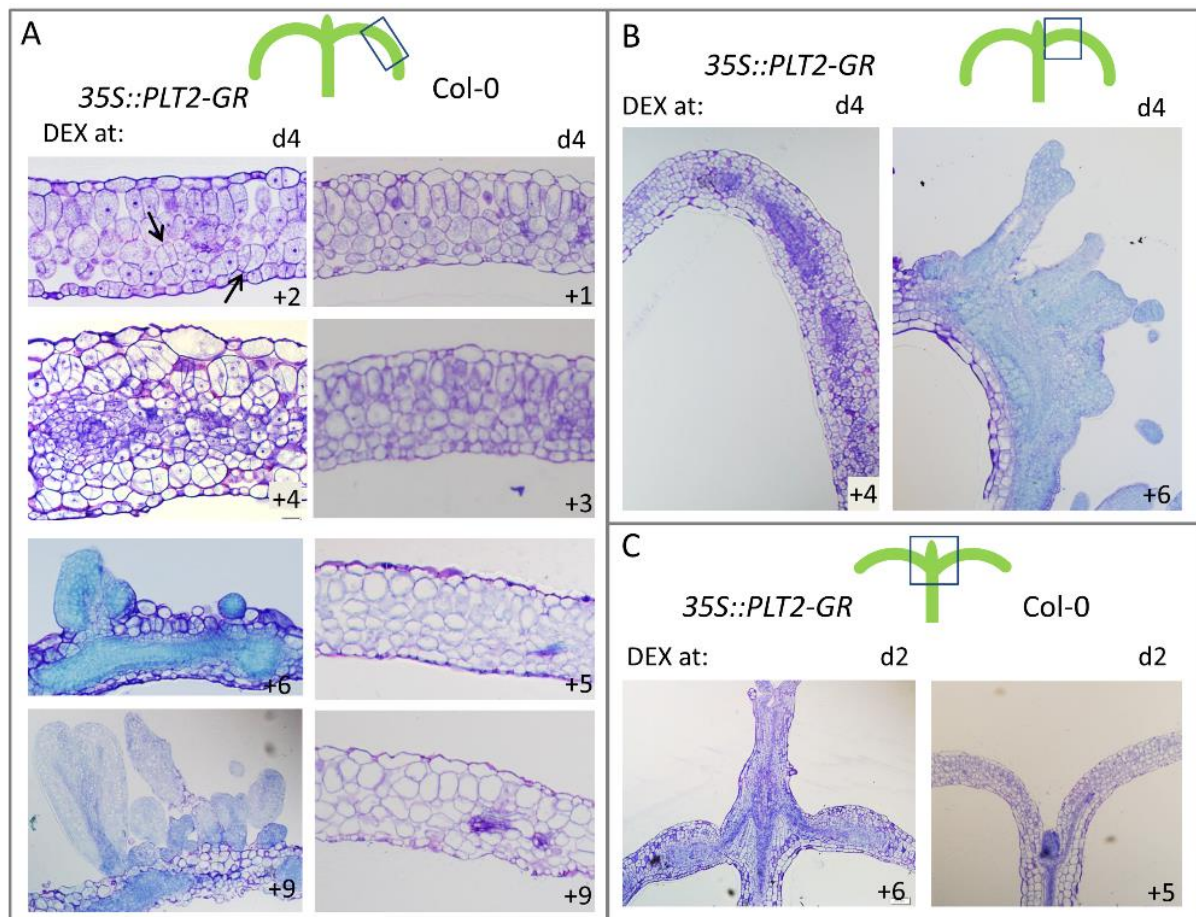


Figure 5. Post-germination activation of PLT2.

Sections of *35S::PLT2-GR* and wild-type Col-0 seedlings grown in medium supplemented with 10 μ M DEX at d4 (A, B) or d2 (C), and cultured for the additional period indicated on the bottom right of each image (+2 to +9 days). The schematic illustrations depict the cotyledon regions (blue boxes) that were sectioned in the images below.

(A, B) PLT2-GR activation at d4 induces cell divisions (+2, arrows) in/around the cotyledon vasculature (+4) in both the distal (A) and proximal (B) parts of the cotyledon. Extensive callus production is observed after 6 days, from which somatic embryos arise later (+9).

(C) PLT2-GR activation at d2 induces growth of the region below the SAM and swelling of the cotyledons.

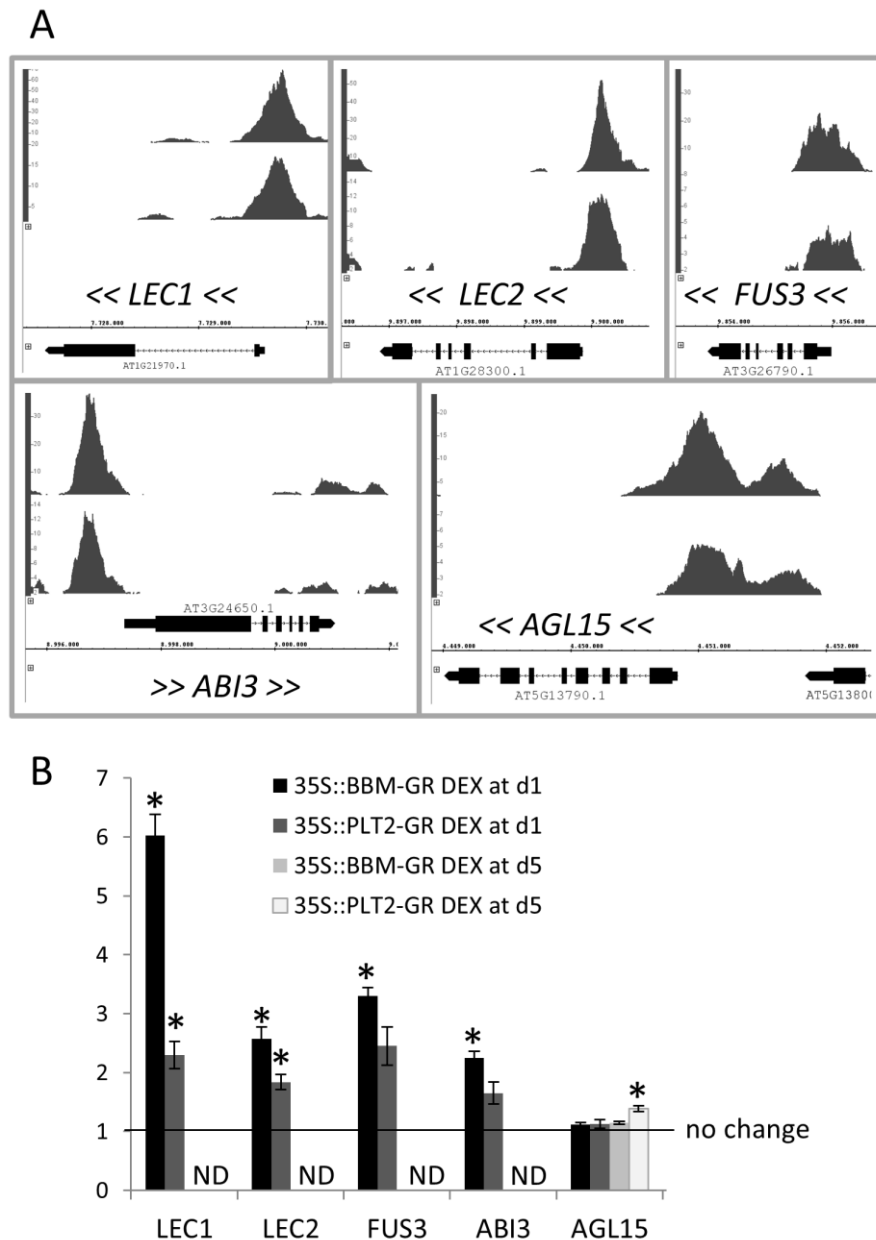


Figure 6. BBM binds and activates embryo-specific genes

(A) ChIP-seq BBM binding profiles for embryo-expressed genes in somatic embryo tissue. The binding profiles from the *35S::BBM-GFP* (upper profile) and *BBM::BBM-YFP* (lower profile) ChIP-seq experiments are shown. The x-axis shows the nucleotide position of DNA binding in the selected genes (TAIR 10 annotation), the y-axis shows the ChIP-seq score, and the brackets indicate the direction of gene transcription. Peaks with scores above 1.76 (for *35S::BBM-GFP*) and 3.96 (for *pBBM::BBM-YFP*) are considered statistically significant (FDR<0.05).

(B) The relative expression of embryo-specific genes was determined by quantitative real-time RT-PCR for DEX+CHX treated *35S::BBM-GR* and *35S::PLT2-GR* seedlings at d1 and d5 using DEX+CHX treated Col-0 as the calibrator and the *SAND* gene (Czechowski et al., 2005) as the reference. Error bars indicate standard errors of the three biological replicates. Statistically significant differences (*) between *35S::BBM-GR/35S::PLT2-GR* and Col-0 were determined using a Student's *t*-test ($p < 0.01$). ND, not detected.

Next, we used the *LEC1::LEC1-GFP* reporter (Li et al., 2014) to chart the dynamics of *LEC1* expression during BBM-induced SE. When DEX is added before germination (d1) *35S::BBM-GR* seedlings form somatic embryos directly on the cotyledon tip. Under these conditions, *LEC1-GFP* was observed one day after BBM-GR activation, in small patches of cells on the abaxial side of the cotyledon (Figure 7C, d1+1).

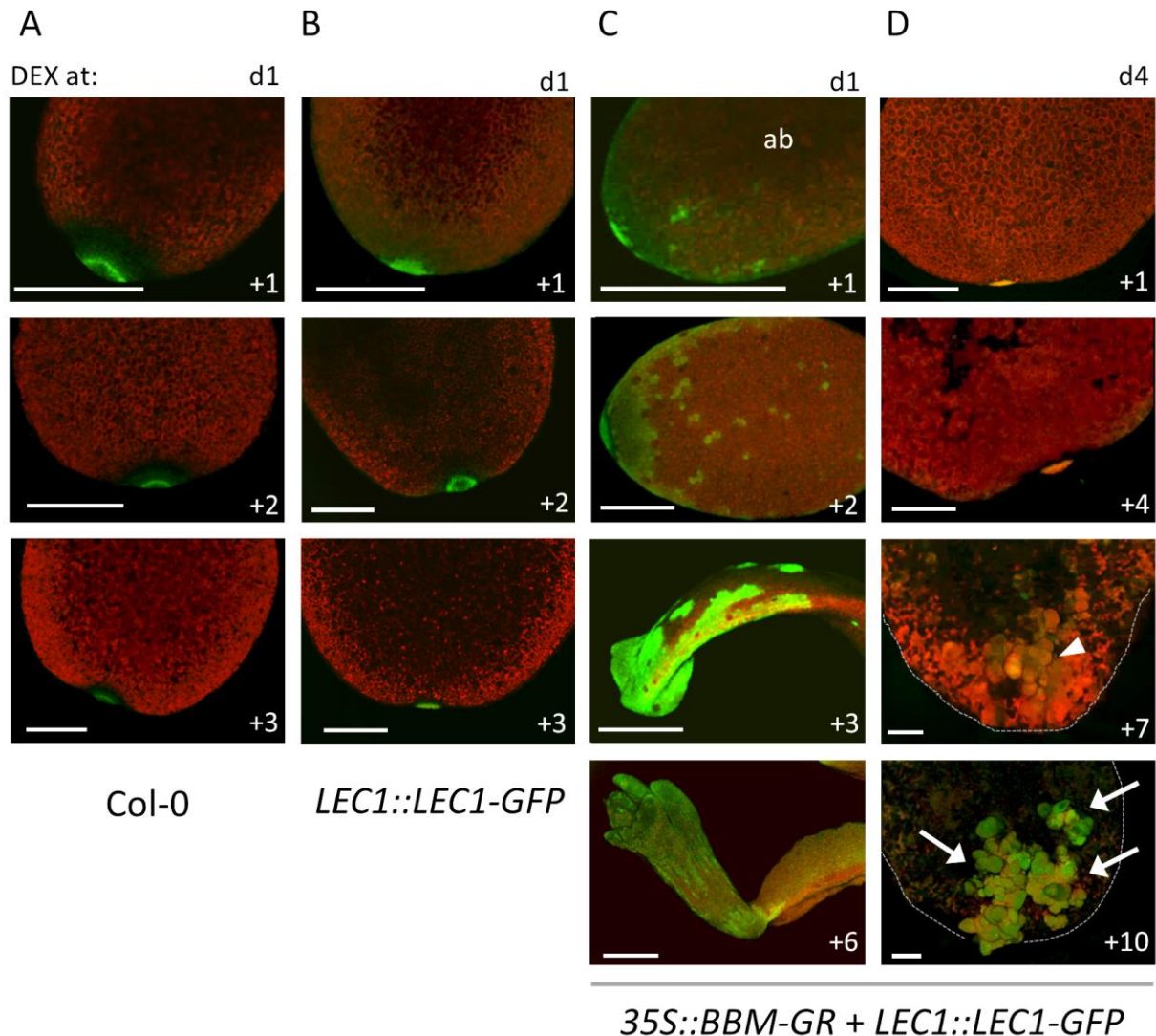


Figure 7: BBM-GR activates *LEC1* expression in a developmentally specific manner

Wild-type Col-0 (A), *LEC1::LEC1-GFP* (B) and *LEC1::LEC1-GFP + 35S::BBM-GR* (C, D) seedlings were treated with 10 μ M DEX at d1 or d4 and the GFP signal was observed from one to 10 days later (indicated on the bottom right of each picture). The images show the adaxial sides of cotyledons, unless indicated otherwise (ab, abaxial side). The green signal in Col-0 (A) and *LEC1::LEC1-GFP* (B) cotyledon tips is autofluorescence. Seedlings that were treated with DEX before germination show the first patches of ectopic *LEC1* expression one day after BBM activation (C). Seedlings that were treated with DEX after seed germination (D) show *LEC1* expression around 10 days after BBM-GR activation (d4+10), when embryogenic clusters are visible (arrows). The arrowhead in (d4+7) indicates the callus that is formed on the distal end of the cotyledon blade. The outline of the cotyledon margins in (D) is shown with dashed lines. Red autofluorescence was used to delineate the tissue. Scale bars, 250 μ m.

LEC1 expression expanded to the cotyledon tip and in patches of cells on the adaxial cotyledon blade (Figure 7C, d1+2), and then became stronger in the cotyledon and extended to the first leaves at the time when the cotyledon tip began to swell (Figure 7C, d1+3). Later, *LEC1* expression was observed in the outer layer of the somatic embryos, but not in the underlying seedling cotyledon (Figure 7C, d1+6). When DEX is added after germination (d4), *35S::BBM-GR* seedlings form callus on the cotyledon blade from which somatic embryos develop. *LEC1-GFP* could only be detected 10 days after DEX-induction (Figure 7D, 4+7, 4+10), where it was localized to the large globular-like embryo structures (Figure 5B). These results reinforce our qPCR-based expression analysis in which we observed rapid *LEC* expression when BBM was activated before germination, but no *LEC* expression when BBM is activated after germination. The observation that *LEC1-GFP* is initially absent from the callus that forms after post-germination *BBM-GR* activation, suggests that somatic embryo identity is established much later in this indirect pathway.

***LAFL* genes and *AGL15* are important for *BBM*-mediated direct SE**

We investigated the genetic relationship between *BBM* and its direct gene targets. Both *LEC1* and *LEC2* overexpression induces spontaneous SE in seedlings, while the *LEC2* target *AGL15* enhances the embryogenic potential in 2,4-D induced SE tissue culture when overexpressed (Lotan et al., 1998; Stone et al., 2001; Harding et al., 2003; Zheng et al., 2009). The other two *LAFL* proteins *FUS3* and *ABI3* do not induce SE when overexpressed, but *FUS3* overexpression confers cotyledon identity to leaves (Gazzarrini et al., 2004), and *ABI3* overexpression increases the expression of seed storage protein genes in leaves in response to ABA (Parcy et al., 1994; Parcy and Giraudat, 1997). Since *BBM* overexpression lines cannot be outcrossed without loss of the *BBM* phenotype, we introduced the *35S::BBM-GR* construct into the *lec1-2^{+/-}*, *lec2-1*, *fus3-3^{+/-}*, *agl15-3* and *abi3* (three alleles) mutant backgrounds via transformation. These mutants, except *agl15-3*, display defects during the later stages of embryogenesis with regard to storage protein accumulation, the acquisition of desiccation tolerance and dormancy (Meinke et al., 1994; Nambara et al., 2002). The *lec1-2* and *fus3-3* seeds are desiccation intolerant (Meinke et al., 1994), therefore heterozygous mutants were used for transformation.

In wild-type *Arabidopsis*, 6-7% of the primary (T1) *35S::BBM-GR* transformants was embryogenic when grown on DEX (Figure 8A). Transformation of the *lec1-2^{+/-}*, *lec2-1*, *fus3-3^{+/-}* and *agl15-3* mutants, resulted in a reduced percentage of *35S::BBM-GR* seedlings that formed embryogenic tissue (Figure 8A). *35S::BBM-GR* also severely inhibited growth and caused swelling

of the cotyledons in the *lec1-2*, *fus3-3* and *lec2-1* backgrounds (15-20%; Figure 8B), a phenotype which was not observed in DEX-activated *35S::BBM-GR* lines. Growth inhibition was also in the *agl15-3* mutants, but not cotyledon swelling (Figure 8B), a phenotype that was also observed in the wild-type background and that resembles *35S::BBM-GR* seedlings treated with low DEX concentrations (Figure 1C). Of the few embryogenic seedlings that were found in the *lec1-2^{+/-}* and *fus3-3^{+/-}* segregating populations none contained the *fus3-3* mutant allele, and only one contained the *lec1-2* mutant allele in the heterozygous state (Figure 8C). Immature embryos from this *lec1-2^{+/-}/35S::BBM-GR* plant were rescued to bypass the *lec1-2* desiccation intolerance. The embryos were separated phenotypically into *lec1-2* homozygous mutant and *lec1-2* heterozygous mutant/wild-type classes and placed on DEX-containing selective medium.

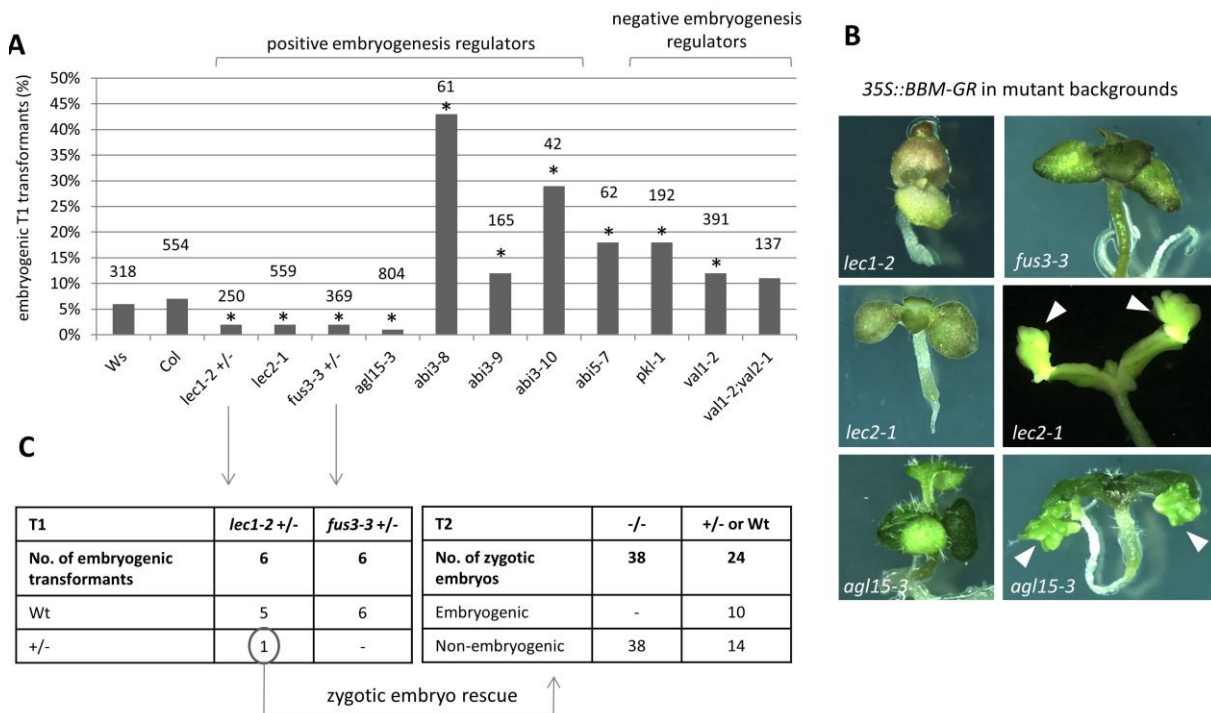


Figure 8. The efficiency of BBM-induced SE in embryogenesis mutants.

(A) Percentage of primary embryogenic transformants obtained after transformation of the *35S::BBM-GR* construct to wt (Ws/Col) or the indicated mutants. Statistically significant differences (*) between the mutant and the corresponding wt line were determined using a Pearson's chi-squared test ($p < 0.05$). The total number of transformants per line is indicated above each bar.

(B) Phenotypes of embryogenesis mutants that contain the *35S::BBM-GR* construct. In the *lec1-2* and *fus3-3* mutants, BBM-GR activation leads to severe growth inhibition, the *lec1-2* mutant was obtained via embryo rescue (C). Severe growth inhibition was also observed in the *lec2-1* mutant (left), but also embryogenic seedlings could be obtained (right). In the *agl15-3* mutant, BBM-GR activation leads to milder growth inhibition (left) and SE (right). Arrowheads indicate the somatic embryos formed on the cotyledon tips.

(C) Phenotype and genotype of the (progeny of) the embryogenic transformants obtained in the *lec1-2^{+/-}* and *fus3-3^{+/-}* segregating populations. The numbers of rescued embryos do not reflect the *lec1-2* phenotype segregation ratio. Phosphinothricin-resistance was used to select the *35S::BBM-GR* transgene.

Somatic embryos formed in wild-type/heterozygous *lec1-2* seedlings, but not in the homozygous *lec1-2* mutant seedlings (Figure 8B, C). Instead, growth was severely inhibited in the *lec1-2/35S::BBM-GR* mutants (Figure 8C). We could also obtain a few homozygous *fus3-3/35S::BBM-GR* lines, showing that the *fus3-3* seed maturation phenotype is not fully penetrant. However, no SE was observed in these lines (Figure 8B). These results suggest that *LEC1*, *LEC2*, *FUS3* are positive regulators of BBM-mediated SE, and that *LEC1* and *FUS3* are absolutely required for this process. Surprisingly, we found that *AGL15* also is a positive regulator of BBM-induced SE even though it is not transcriptionally regulated by BBM overexpression at the two time points examined; *AGL15* might be regulated by BBM at a later time point or function downstream of the *LAFL* genes (Braybrook et al., 2006) in BBM-induced SE.

In contrast to the results obtained with the *fus3*, *lec* and *agl15* mutants, transformation of the *35S::BBM-GR* construct to three different *abi3* mutants led to an enhanced SE response (Figure 8A). Notably, *abi3* is the only *LAFL* mutant that is insensitive to ABA and overexpression of *ABI3* does not lead to somatic embryogenesis (Parcy et al., 1994; Parcy and Giraudat, 1997). Of the three examined *abi3* alleles, *abi3-9* had the mildest effect on BBM-induced SE (Figure 8A). Interestingly, the *abi3-9* mutant was also found to be sensitive to ABA in the presence of glucose, in contrast to *abi3-8* and *abi3-10*, which were ABA-insensitive under these conditions (Nambara et al., 2002). In order to separate the effects of ABA-insensitivity and other embryo defects of *abi3* mutants on the BBM phenotype, we tested another ABA-insensitive mutant, *abi5-7*, which does not have any other reported embryo defects (Nambara et al., 2002). We also observed an enhanced BBM phenotype in the *abi5-7* mutant compared to wild-type (Figure 8A). These data suggest that BBM-mediated SE is suppressed by ABA signalling and that the enhanced BBM response in the *abi3* mutants is due to ABA-insensitivity, rather than to other defects in the *abi3* mutants.

Finally, we tested whether transcriptional repressors of the *LAFL* genes, PKL and VAL proteins, have an effect on the BBM phenotype. We observed that *pkl-1* and *val1-2* (*hsi2-5*) mutants enhanced the efficiency of BBM-mediated SE, as measured by a higher percentage of embryogenic primary transformants (Figure 8A). In the *val1-2;val2-1* double mutant, no significant change in SE-induction could be observed, which may be due to the lower number of transformants obtained in this mutant.

Together, the data show that members of the *LAFL* network, as well their upstream and downstream regulators are important components of the BBM signalling pathway during somatic embryo induction (Figure 9).

DISCUSSION

AIL transcription factors play key roles throughout plant development, where they regulate processes such as meristem identity and maintenance, cell proliferation, organ size and organ development (Horstman et al., 2014). Functional redundancy among AILs has been demonstrated using loss-of-function mutants, but these shared functions have been difficult to reconcile with the range of phenotypes observed in AIL overexpression studies. Using the same overexpression and growth conditions, we have shown that AIL proteins have overlapping functions that are expressed in a dose-dependent manner. Our data suggest that the variety of overexpression phenotypes observed in different studies can be explained in part by differences in transgene expression levels.

Dose-dependent AIL function

We have shown that relatively high expression of the six BBM clade of AIL proteins induced somatic embryogenesis. By contrast, overexpression of the remaining two AIL proteins, ANT and AIL1, was not sufficient to induce SE. The *ANT* and *AIL1* genes comprise the basalANT lineage within the *AIL* family, while the remaining proteins belong to the euANT lineage (Kim et al., 2006). The expression pattern of *ANT* also differs from that of other *AIL* genes; *ANT* is expressed at the meristem periphery in the shoot apical and flower meristems, while the other *AIL/PLT* genes are expressed throughout these meristems (Prasad et al., 2011; Mudunkothge and Krizek, 2012b). This suggests that the two groups of AIL proteins regulate distinct processes.

We showed that a high BBM/PLT2 dose induces SE, a lower dose induces organogenesis and the lowest dose inhibits differentiation. Although we did not examine the dose-dependency of the other AILs, it is likely that they have similar dose-response phenotypes. It was suggested that PLT2 and, by extension, other AIL proteins act as morphogens, regulating root meristem size and maintenance in a dose-dependent manner through a protein concentration gradient, with a high AIL dose instructing stem cell fate, an intermediate AIL dose leading to cell division, and a low AIL dose causing differentiation (Galinha et al., 2007). Our results on seedling cotyledons and leaves also support a dose-dependent AIL output in these tissues, but suggest that a low AIL dose prevents differentiation rather than promoting differentiation. In analogy, a low AIL dose in the root might not actively instruct cellular differentiation, rather, it might simply be ineffective, thereby allowing cellular differentiation. We showed that a high AIL dose induces SE in cotyledons, but it is not known whether this proceeds through a stem cell pathway as instructed

by a high AIL dose in the root. Likewise, it is not known whether higher AIL concentrations than are found in the stem cell niche are required for organogenesis and embryogenesis under normal growth conditions/*in planta*. Measurement of cellular AIL protein levels would help to relate the endogenous protein expression levels to those in overexpression lines.

It is currently unclear how different AIL concentrations instruct separate cellular outputs. The AIL dose-dependent phenotypes could result from different expression levels of the same target genes and/or from dose-dependent activation of specific target genes. A transcription factor gradient can regulate different sets of target genes through differences in binding site number and affinity (Rogers and Schier, 2011). In this model, target genes with many or high-affinity binding sites are activated by low levels of the transcription factor, whereas genes with few or low-affinity binding sites are only activated at high transcription factor levels. For example, the transcription factor Bicoid regulates anterior-posterior axis patterning in *Drosophila* embryos through a protein gradient, and Bicoid target genes with high-affinity binding sites were expressed at lower Bicoid levels in contrast to targets with low-affinity binding sites (Driever et al., 1989). Genome-wide AIL-DNA binding studies using different AIL dosages could reveal whether such high- and low-affinity AIL binding sites exist.

We observed some differences in the dose- and context- dependent overexpression phenotypes of BBM and PLT2. For example, intermediate doses of BBM and PLT2 mainly (though not exclusively) induce ectopic shoot and root formation, respectively. It is not clear how AIL specificity is determined. The *in vitro* DNA binding sites of ANT and PLT5/AIL5 appear to be very similar (Nole-Wilson and Krizek, 2000; Yano et al., 2009), but they need to be better defined for each AIL protein. We have previously shown that multiple AILs interact with HDGs, however the individual AIL-HDG interactions differed (Chapter 3). Defining the overlapping and unique target genes for each AIL transcription factor and the protein complexes in which they function may shed light on how specificity is achieved.

AILs trigger two distinct SE pathways

We observed that the developmental context in which AILs are expressed also affects the SE process. BBM and PLT2 can induce SE in two ways: either directly and quickly or indirectly and slowly. Direct SE was observed when BBM and PLT2 are activated before or during germination, and indirect SE when activated after germination. During direct SE, cells in the L1/L2 layers of the cotyledon divide, and somatic embryos develop from the cotyledon tips. The indirect SE pathway seems to take a different route: the upper layers become rough and irregular, the underlying

tissue proliferates and somatic embryos are formed on the cotyledon blade. Previously, it was shown that organogenesis from aerial tissues starts from pericycle-like cells around the vasculature and proceeds via a lateral root pathway (Che et al., 2007; Atta et al., 2009; Sugimoto et al., 2010). Embryogenic callus can also be derived from pericycle-like cells (Sticklen, 1991; Yang et al., 2010). Indirect BBM/PLT2-induced somatic embryogenesis does not appear to originate from vascular-derived callus, but rather from the ground tissue. However, future research should focus on whether this embryogenic callus originates from a similar lateral root pathway or a completely different developmental program.

In *Arabidopsis*, late zygotic embryo stages and dry seeds are the only stages that have been reported to undergo direct SE. All other tissues form callus and then somatic embryos, regardless of the inducing factor (2,4-D/transcription factor; discussed in Chapter 1). Our results reinforce the existence of such a developmental window of competence for direct SE, and the idea that tissues outside this window require more extensive reprogramming, callus formation, before the embryo program can be initiated.

BBM-mediated SE requires LEC and FUSCA gene expression

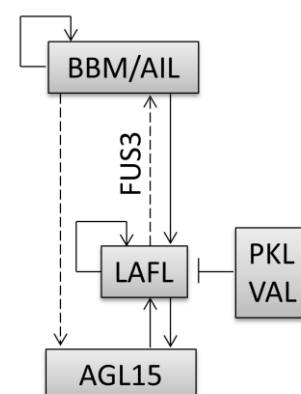
Besides AILs, the LEC1/LEC2 and AGL15 transcription factors can also induce or enhance SE respectively, when overexpressed (Lotan et al., 1998; Stone et al., 2001; Harding et al., 2003; Zheng et al., 2009). *LEC* genes are important regulators of seed maturation and it was previously suggested that LEC2 overexpression might promote SE through dehydration stress resulting from the ectopic activation of seed maturation processes in vegetative tissues (Stone, 2008). Overexpression of *FUS3* does not induce SE, but does induce cotyledon identity in leaves: they have a rounder shape, lack trichomes and accumulate seed storage proteins (Gazzarrini et al., 2004). Loss of function of the LAFL repressors PKL and VAL also induces spontaneous SE (Ogas et al., 1999; Suzuki et al., 2007). We showed that BBM acts upstream of these embryogenesis regulators and that their expression is important for SE (Figure 9). BBM overexpression in *lec1*, *lec2*, *fus3* and *agl15* mutants either eliminated or reduced SE. The importance of *FUS3* for BBM-mediated SE was unexpected as *FUS3* overexpression does not enhance or induce SE, however, the *fus3* mutant is also impaired in 2,4-D induced SE (Gaj et al., 2005). The reduction of BBM-mediated SE in the mutants could be explained in two ways: (1) the developmental defects in the mutants change the physiological state of the seed in such a way that it is no longer responsive for BBM-mediated SE, or (2) that BBM-induced SE relies on suitable transcriptional activation of these target genes, which is hampered in the mutants. Several lines of evidence

support the latter scenario. First, we observed a reduced responsiveness to BBM in segregating *lec1* and *fus3* populations, which contain wild-type and heterozygous plants. However, the few embryogenic transformants in these populations were mainly wild-types, suggesting that the *lec1* and *fus3* mutations already affect BBM-induced SE in the heterozygote state. Heterozygous *lec1* and *fus3* mutants do not show reported growth defects, suggesting that reduced *LEC1* or *FUS3* expression in the heterozygous mutants, rather than a change in the physiological state of the tissue, reduces the response to BBM overexpression. Secondly, the *abi3* mutant shows similar maturation defects as the other *LAFL* mutants, yet we observed no negative effect of *abi3* mutations on the BBM overexpression phenotype. Therefore, we hypothesize that the lack of elevated expression of the *LAFL* genes reduces BBM-induced SE in the mutants. This hypothesis is further strengthened by our observations that mutations in *LAFL* repressors (*PKL/VAL*; Figure 9) enhance BBM-mediated SE, probably by facilitating elevated *LAFL* gene expression. The enhanced BBM response in the *abi3* and *abi5* mutants is intriguing. Exogenous ABA application is reported to either inhibit or promote somatic embryo induction, depending on the experimental system (Rai et al., 2011). In Arabidopsis, the ABA-insensitive *abi3* and *abi5* mutants have a negative effect on 2,4-D induced direct SE from immature zygotic embryos, but so do ABA hypersensitive mutants (Gaj et al., 2006), making it difficult to assign a single role to ABA in this system.

It was previously shown that transcriptional feedback loops exist within the *LAFL* network between known regulators of SE (Figure 9) (To et al., 2006; Jia et al., 2013). Recently, it was also shown that the *Phaseolus vulgaris* ABI3-like factor (Pv-ALF), which binds to the promoter of Arabidopsis *PLT5/AIL5* *in vitro*, and that *PLT5/AIL5* is required for activation of seed storage genes by Pv-ALF (Sundaram et al., 2013). In addition, *FUS3* binds to the first exon/intron of *BBM* *in vivo*, although direct transcriptional regulation by *FUS3* was not investigated (Wang and Perry, 2013). Here, we uncovered another regulatory layer in which BBM stimulates the expression of *LAFL* genes during the induction of SE.

Figure 9: SE gene regulatory networks.

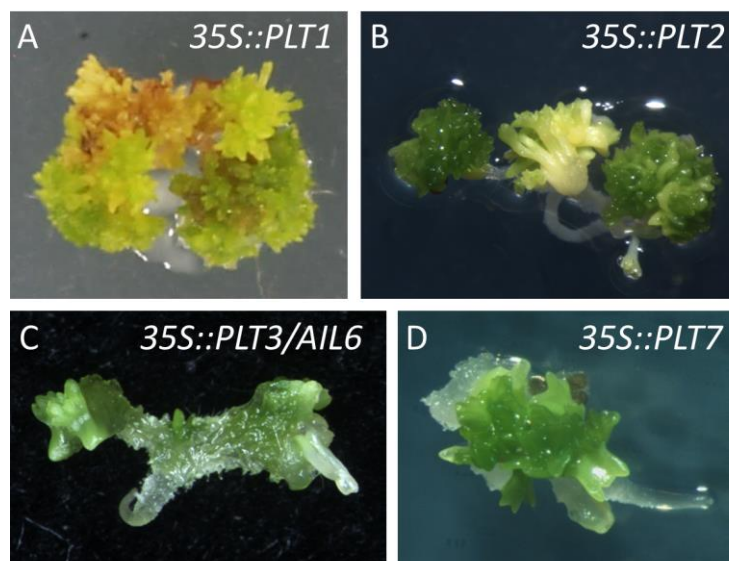
Schematic representation of the genetic interactions between genes involved in SE. The solid lines indicate DNA binding plus transcriptional activation or repression, while the dashed lines indicate DNA binding in the absence of transcriptional regulation.



ACKNOWLEDGEMENTS

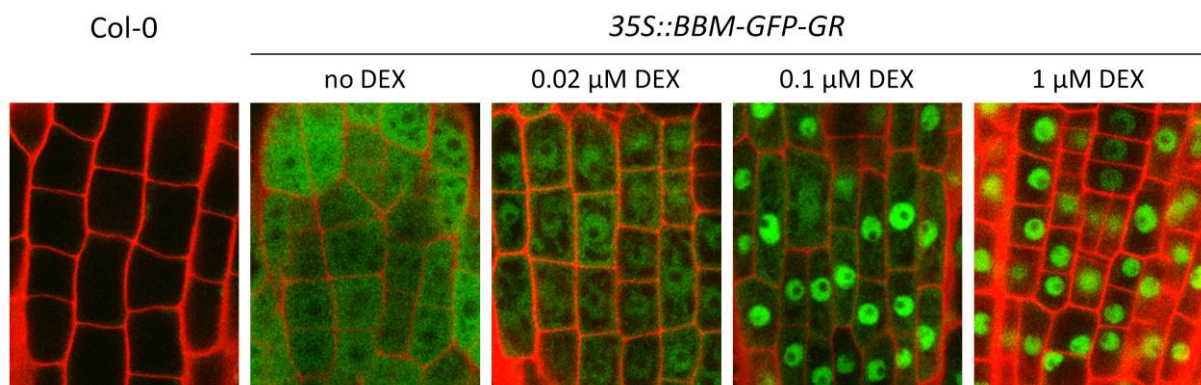
We thank Nirmala Sharma for the *val1-2 (hsi2-5)* mutant, Masaharu Suzuki for the *val1-2;val2-1* mutant and Bas Dekkers for the *abi3-8*, *abi3-9*, *abi3-10* and *abi5-7* mutants. This project was supported by a Technology Top Institute Green Genetics grant to K.B. and I.H.

SUPPLEMENTAL MATERIAL



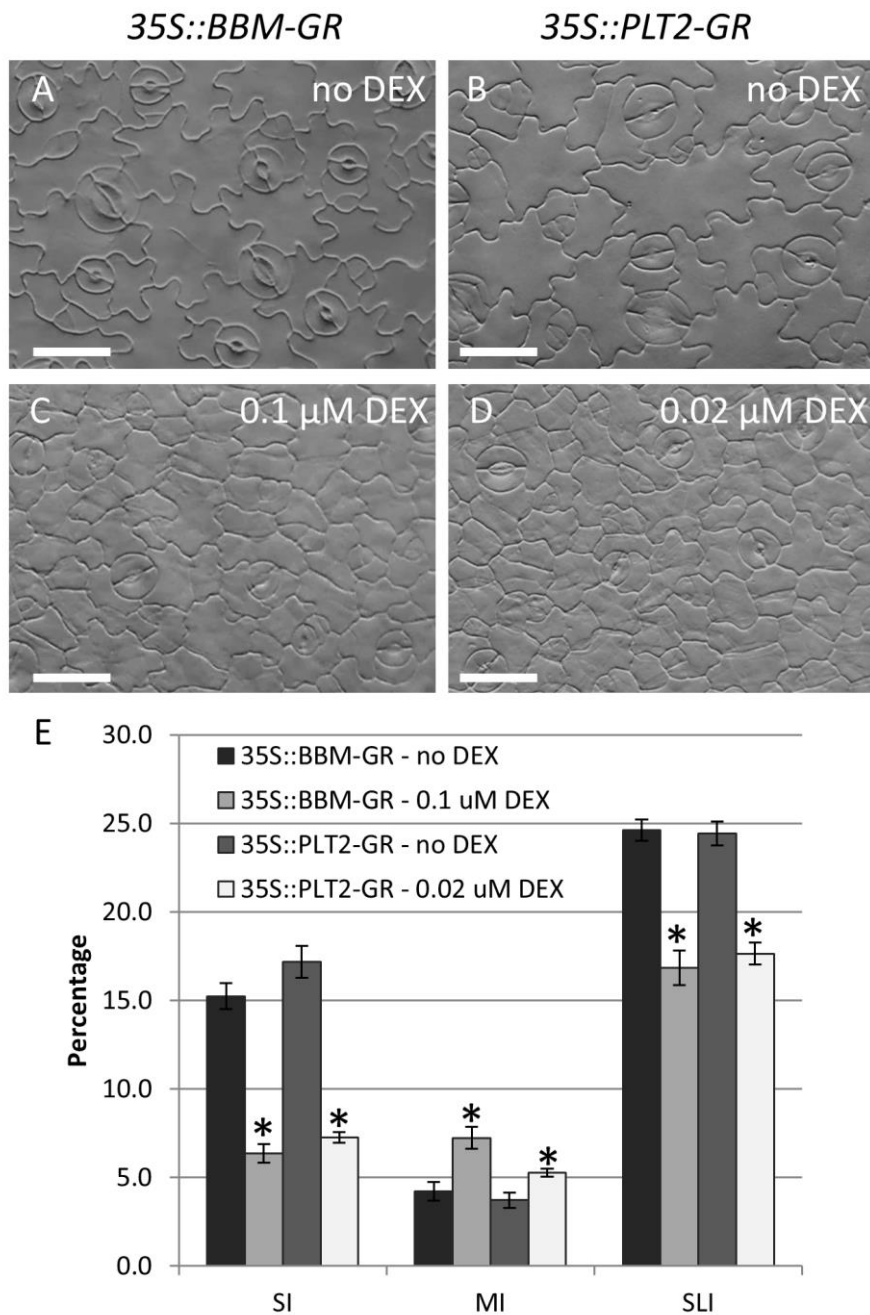
Supplemental Figure 1. Overexpression of AIL/PLT proteins induces somatic embryogenesis

Somatic embryo phenotypes of Arabidopsis primary transformants: *35S::PLT1* (A); *35S::PLT2* (B); *35S::PLT3/AIL6* (C); and *35S::PLT7* (D). Seedlings were grown on selection medium for 12 days (C), 3 weeks (D), 4 weeks (B) or 7 weeks (A).



Supplemental Figure 2. BBM-GFP-GR nuclear localization increases with increasing dexamethasone concentration

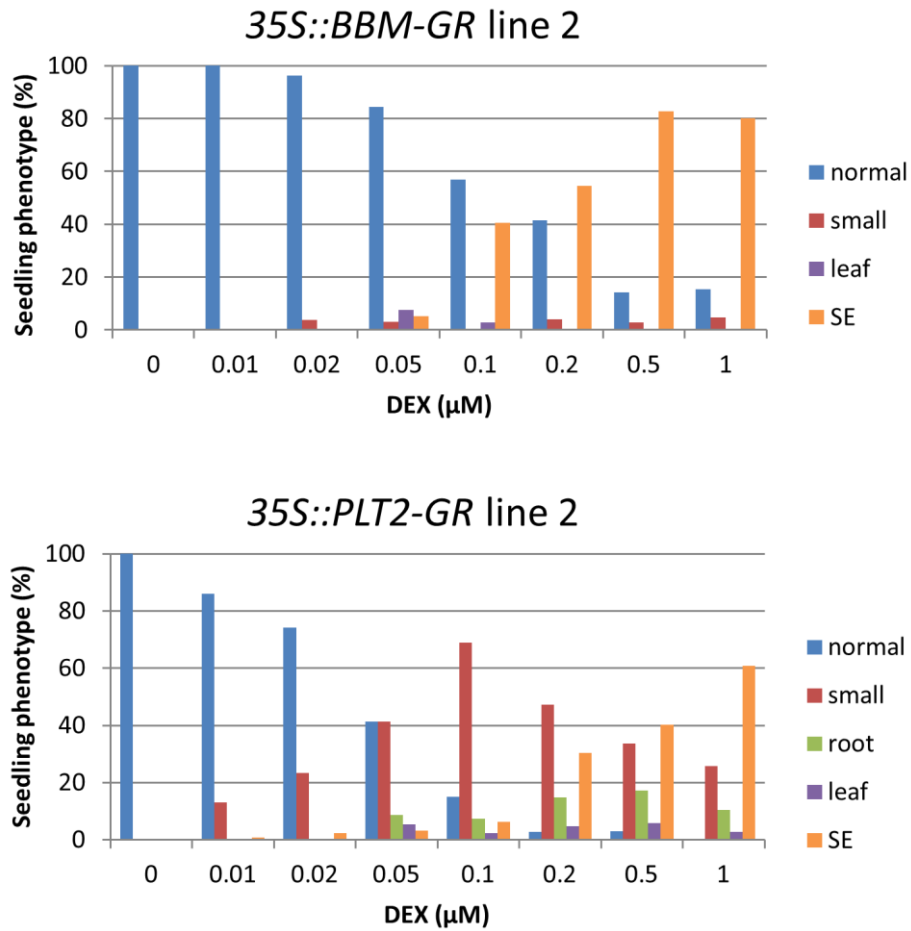
The effect of dexamethasone (DEX) on BBM localization in roots of *35S::BBM-GFP-GR* seedlings grown for seven days in medium containing the indicated DEX concentration. Non-DEX treated (Col-0) roots are shown as a GFP-negative control. Green, GFP. Red, propidium iodide.



Supplemental Figure 3. A low BBM/PLT2 dose inhibits differentiation of leaf epidermal cells

The abaxial sides of cleared first leaves of nine day-old *35S::BBM-GR* (A, C) and *35S::PLT2-GR* (B, D) seedlings grown on medium without DEX (A, B) or with 0.1 or 0.02 μM DEX (C, D). A relatively low BBM or PLT2 dose leads to the development of smaller and less-lobed leaf pavement cells compared to the control. Scale bars, 25 μm .

(E) Stomatal differentiation in DEX-treated *35S::BBM-GR* and *35S::PLT2-GR* seedlings is reduced compared to untreated seedlings. Fewer mature stomata were found in leaves of DEX-treated *35S::BBM-GR/PLT2-GR* seedlings, as reflected by a lower stomatal index (SI), while the number of stomatal meristemoids was increased (meristemoid index, MI). The stomatal lineage index (SLI), reflecting the total number of stomata and stomatal precursors, was lower in DEX-treated *35S::BBM-GR/PLT2-GR* leaves than in the control. For each index, eight images were analysed with total cell numbers between 125 and 350 per image. Error bars indicate standard errors. *, statistically significant difference compared to the control ($p < 0.05$ in Student's *t*-test).



Supplemental Figure 4. BBM/PLT2 dose-dependent overexpression phenotypes in independent transgenic lines

Effect of DEX dose on the development of additional, independent *35S::BBM-GR* and *35S::PLT2-GR* lines. The experimental conditions were the same as for the lines shown in Figure 1. No additional phenotypes were observed in treatments above 1 μM DEX. $n=200$ to 350 seedlings. Leaf, ectopic leaves; root, ectopic root; SE, somatic embryogenesis.

Supplemental Table 1. Primers used in this study

Cloning		
<i>ANT</i>	FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGAAGTCTTTTTGTGATAATGATGA
	RV	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAAGAATCAGCCCAAGCAG
<i>AIL1</i>	FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGAAGAAATGGTTGGGATTTT
	RV	GGGGACCACTTTGTACAAGAAAGCTGGGTATTAGTGGCCGGCGC
<i>PLT3/AIL6</i>	FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGATGGCTCCGATGACG
	RV	GGGGACCACTTTGTACAAGAAAGCTGGGTATTAGTAAGACTGATTAGGCCAGAGG
<i>PLT7</i>	FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGGCTCCTCCAATGACG
	RV	GGGGACCACTTTGTACAAGAAAGCTGGGTATTAGTAAGACTGGTTAGGCCACAA
<i>PLT1</i>	FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGAATTCTAACAACCTGGCTTGG
	RV	GGGGACCACTTTGTACAAGAAAGCTGGGTATTACTCATTCCACATAGTGAAAACAC
<i>PLT2</i>	FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGAATTCTAACAACCTGGCTCG
	RV+stop	GGGGACCACTTTGTACAAGAAAGCTGGGTATTATTTCATTCCACATCGTGAAAAC
	RV-stop	GGGGACCACTTTGTACAAGAAAGCTGGGTATTTCATTCCACATCGTGAAAAC
<i>BBM-GFP</i>	FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAACCTCGATGAATAACTGGTT
	RV-stop	GGGGACCACTTTGTACAAGAAAGCTGGGTACTTGTACAGCTCGTCCATGC
Gene expression analysis		
<i>SAND</i>	FW	AACTCTATGCAGCATTTGATCCACT
	RV	TGATTGCATATCTTTATCGCCATC
<i>LEC1</i>	FW	ACAAGAACAATGGTATCGTGGTCC
	RV	GAGATTTTGGCGTGAGACGGTAA
<i>LEC2</i>	FW	ATCGCTCGCACTTCAACACAG
	RV	AACAAGGATTACCAACCAGAGAACC
<i>FUS3</i>	FW	TCTTCTTCTTTAACCTTCTCTCTTTCC
	RV	ACCGTCCAATCTTCCATTCTTATAGG
<i>ABI3</i>	FW	GGCAGGGATGGAAACCAGAAAAGA
	RV	GGCAAAACGATCCTTCCGAGGTTA
<i>AGL15</i>	FW	GAACGATTGCTGACTAACCAACTTG
	RV	GCAAAGTTGTGCTGAATCGGTGTT

Supplemental Table 2. The efficiency of AIL-induced SE

Construct	No. of primary transformants	No. of transformants with SE	% SE
<i>35S::ANT</i>	89	0	-
<i>35S::AIL1</i>	228	0	-
<i>35S::AIL6</i>	171	45	26%
<i>35S::AIL7</i>	57	10	18%
<i>35S::PLT1</i>	136	9	7%
<i>35S::PLT2</i>	96	10	10%
<i>35S::BBM</i>	81	19	22%

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.
- Atta, R., Laurens, L., Boucheron-Dubuisson, E., Guivarc'h, A., Carnero, E., Giraudat-Pautot, V., Rech, P., and Chriqui, D. (2009). Pluripotency of Arabidopsis xylem pericycle underlies shoot regeneration from root and hypocotyl explants grown in vitro. *Plant J* **57**, 626-644.
- Bewley, J.D. (1997). Seed Germination and Dormancy. *The Plant cell* **9**, 1055-1066.
- Boutilier, K., Offringa, R., Sharma, V.K., Kieft, 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. *The 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 Genetics* **7**.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3468-3473.
- Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**, 5263-5276.
- Che, P., Lall, S., and Howell, S.H. (2007). Developmental steps in acquiring competence for shoot development in Arabidopsis tissue culture. *Planta* **226**, 1183-1194.
- Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant Journal* **16**, 735-743.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant physiology* **139**, 5-17.
- Danisman, S., van der Wal, F., Dhondt, S., Waites, R., de Folter, S., Bimbo, A., van Dijk, A.J., Muino, J.M., Cutri, L., Dornelas, M.C., Angenent, G.C., and Immink, R.G.H. (2012). Arabidopsis Class I and Class II TCP Transcription Factors Regulate Jasmonic Acid Metabolism and Leaf Development Antagonistically. *Plant physiology* **159**, 1511-1523.
- Driever, W., Thoma, G., and Nusslein-Volhard, C. (1989). Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. *Nature* **340**, 363-367.
- El Ouakfaoui, S., Schnell, J., Abdeen, A., Colville, A., Labbe, H., Han, S., Baum, B., Laberge, S., and Miki, B. (2010). Control of somatic embryogenesis and embryo development by AP2 transcription factors. *Plant molecular biology* **74**, 313-326.
- Feher, A. (2014). Somatic embryogenesis - Stress-induced remodeling of plant cell fate. *Biochim Biophys Acta*.
- Gaj, M.D., Zhang, S.B., Harada, J.J., and Lemaux, P.G. (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of Arabidopsis. *Planta* **222**, 977-988.
- Gaj, M.D., Trojanowska, A., Ujczak, A., Medrek, M., Koziol, A., and Garbaciak, B. (2006). Hormone-response mutants of Arabidopsis thaliana (L.) Heynh. impaired in somatic embryogenesis. *Plant Growth Regul* **49**, 183-197.

- 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.
- Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M., and McCourt, P.** (2004). The transcription factor FUSCA3 controls developmental timing in Arabidopsis through the hormones gibberellin and abscisic acid. *Developmental cell* **7**, 373-385.
- Gorte, M., Horstman, A., Page, R.B., Heidstra, R., Stromberg, A., and Boutilier, K.** (2011). Microarray-based identification of transcription factor target genes. *Methods in molecular biology* **754**, 119-141.
- 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 physiology* **133**, 653-663.
- 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.
- Immink, R.G.H., Gadella, T.W.J., Ferrario, S., Busscher, M., and Angenent, G.C.** (2002). Analysis of MADS box protein-protein interactions in living plant cells. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 2416-2421.
- Jia, H., McCarty, D.R., and Suzuki, M.** (2013). Distinct roles of LAFL network genes in promoting the embryonic seedling fate in the absence of VAL repression. *Plant physiology* **163**, 1293-1305.
- Kim, S., Soltis, P.S., Wall, K., and Soltis, D.E.** (2006). Phylogeny and domain evolution in the APETALA2-like gene family. *Mol Biol Evol* **23**, 107-120.
- Klucher, K.M., Chow, H., Reiser, L., and Fischer, R.L.** (1996). The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. *The Plant cell* **8**, 137-153.
- Krizek, B.** (2009). AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. *Plant physiology* **150**, 1916-1929.
- Krizek, B.A.** (1999). Ectopic expression of AINTEGUMENTA in Arabidopsis plants results in increased growth of floral organs. *Developmental genetics* **25**, 224-236.
- Krizek, B.A., and Eaddy, M.** (2012). AINTEGUMENTA-LIKE6 regulates cellular differentiation in flowers. *Plant molecular biology* **78**, 199-209.
- Li, H., Soriano, M., Cordewener, J., Muino, 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. *The Plant cell* **26**, 195-209.
- Livak, K.J., and Schmittgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**, 402-408.
- Lotan, T., Ohto, M., Yee, K.M., West, M.A.L., 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.
- Meinke, D.W., Franzmann, L.H., Nickle, T.C., and Yeung, E.C.** (1994). Leafy Cotyledon Mutants of Arabidopsis. *Plant Cell* **6**, 1049-1064.
- Mizukami, Y., and Fischer, R.L.** (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 942-947.
- Mudunkothge, J.S., and Krizek, B.A.** (2012a). Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function. *The Plant journal : for cell and molecular biology* **71**, 108-121.
- Mudunkothge, J.S., and Krizek, B.A.** (2012b). Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function. *Plant J* **71**, 108-121.

- Nambara, E., Suzuki, M., Abrams, S., McCarty, D.R., Kamiya, Y., and McCourt, P.** (2002). A screen for genes that function in abscisic acid signaling in *Arabidopsis thaliana*. *Genetics* **161**, 1247-1255.
- Nicol, J.W., Helt, G.A., Blanchard, S.G., Raja, A., and Loraine, A.E.** (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**, 2730-2731.
- Nole-Wilson, S., and Krizek, B.A.** (2000). DNA binding properties of the *Arabidopsis* floral development protein AINTEGUMENTA. *Nucleic acids research* **28**, 4076-4082.
- Nole-Wilson, S., Tranby, T.L., and Krizek, B.A.** (2005). AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant molecular biology* **57**, 613-628.
- 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*. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13839-13844.
- Parcy, F., and Giraudat, J.** (1997). Interactions between the ABI1 and the ectopically expressed ABI3 genes in controlling abscisic acid responses in *Arabidopsis* vegetative tissues. *Plant J* **11**, 693-702.
- Parcy, F., Valon, C., Raynal, M., Gaubier-Comella, P., Delseny, M., and Giraudat, J.** (1994). Regulation of gene expression programs during *Arabidopsis* seed development: roles of the ABI3 locus and of endogenous abscisic acid. *Plant Cell* **6**, 1567-1582.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M.W., 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 molecular biology* **68**, 225-237.
- Peterson, K.M., Shyu, C., Burr, C.A., Horst, R.J., Kanaoka, M.M., Omae, M., Sato, Y., and Torii, K.U.** (2013). *Arabidopsis* homeodomain-leucine zipper IV proteins promote stomatal development and ectopically induce stomata beyond the epidermis. *Development* **140**, 1924-1935.
- Prasad, K., Grigg, S.P., Barkoulas, M., Yadav, R.K., Sanchez-Perez, G.F., Pinon, V., Blilou, I., Hofhuis, H., Dhonukshe, P., Galinha, C., Mahonen, A.P., Muller, W.H., Raman, S., Verkleij, A.J., Snel, B., Reddy, G.V., Tsiantis, M., and Scheres, B.** (2011). *Arabidopsis* PLETHORA transcription factors control phyllotaxis. *Current biology : CB* **21**, 1123-1128.
- Rai, M.K., Shekhawat, N.S., Harish, Gupta, A.K., Phulwaria, M., Ram, K., and Jaiswal, U.** (2011). The role of abscisic acid in plant tissue culture: a review of recent progress. *Plant Cell Tiss Org* **106**, 179-190.
- Rogers, K.W., and Schier, A.F.** (2011). Morphogen gradients: from generation to interpretation. *Annu Rev Cell Dev Biol* **27**, 377-407.
- Sharma, N., Bender, Y., Boyle, K., and Fobert, P.R.** (2013). High-level expression of sugar inducible gene2 (HSI2) is a negative regulator of drought stress tolerance in *Arabidopsis*. *BMC plant biology* **13**, 170.
- Sticklen, M.B.** (1991). Direct Somatic Embryogenesis and fertile Plants from Rice Root Cultures. *Journal of plant physiology* **138**, 577-580.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11806-11811.
- Sugimoto, K., Jiao, Y.L., and Meyerowitz, E.M.** (2010). *Arabidopsis* Regeneration from Multiple Tissues Occurs via a Root Development Pathway. *Dev Cell* **18**, 463-471.

- Sundaram, S., Kertbundit, S., Shakirov, E.V., Iyer, L.M., Juricek, M., and Hall, T.C.** (2013). Gene networks and chromatin and transcriptional regulation of the phaseolin promoter in *Arabidopsis*. *Plant Cell* **25**, 2601-2617.
- Suzuki, M., Wang, H.H., and McCarty, D.R.** (2007). Repression of the LEAFY COTYLEDON 1/B3 regulatory network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. *Plant physiology* **143**, 902-911.
- To, A., Valon, C., Savino, G., Guilleminot, J., Devic, M., Giraudat, J., and Parcy, F.** (2006). A network of local and redundant gene regulation governs *Arabidopsis* seed maturation. *The Plant cell* **18**, 1642-1651.
- 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 molecular biology* **73**, 481-492.
- Wang, F., and Perry, S.E.** (2013). Identification of direct targets of FUSCA3, a key regulator of *Arabidopsis* seed development. *Plant physiology* **161**, 1251-1264.
- Yang, J.L., Seong, E.S., Kim, M.J., Ghimire, B.K., Kang, W.H., Yu, C., and Li, C.** (2010). Direct somatic embryogenesis from pericycle cells of broccoli (*Brassica oleracea* L. var. *italica*) root explants. *Plant Cell, Tissue and Organ Culture (PCTOC)* **100**, 49-58.
- Yano, R., Kanno, Y., Jikumaru, Y., Nakabayashi, K., Kamiya, Y., and Nambara, E.** (2009). CHOTTO1, a putative double APETALA2 repeat transcription factor, is involved in abscisic acid-mediated repression of gibberellin biosynthesis during seed germination in *Arabidopsis*. *Plant physiology* **151**, 641-654.
- 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. *The Plant cell* **21**, 2563-2577.

Chapter 5

Genome-wide BABY BOOM binding sites during somatic embryogenesis in *Arabidopsis*

Anneke Horstman¹, Jose M. Muino¹, Iris Heidmann², Omid Karami³, Remko
Offringa³, Gerco Angenent^{1,4}, Kim Boutilier¹

¹ Wageningen University and Research Centre, Bioscience, Droevendaalsesteeg 1, 6708 PB
Wageningen, The Netherlands

² Enza Zaden Research and Development B.V, Haling 1-E, 1602 DB Enkhuizen, The Netherlands

³ Molecular and Developmental Genetics, Sylviusweg 72, 2333 BE Leiden, The Netherlands

⁴ Laboratory of Molecular Biology, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

Abstract

AINTEGUMENTA-LIKE (AIL) transcription factors play important roles in plant development and overexpression of some AILs, including BABY BOOM (BBM), induces somatic embryogenesis. Many genetic interactions have been described for AIL proteins, but their direct target genes are largely unknown. We studied the genome-wide, *in vivo* DNA binding sites of BBM during somatic embryogenesis by chromatin immunoprecipitation followed by sequencing (ChIP-seq). We identified a BBM consensus DNA binding motif, GCAC(A/G)NNNN(C/T)CNAN(A/G), that resembles the reported ANT binding motif. In addition, we show that BBM binds and regulates auxin biosynthesis genes and *AT-HOOK-LIKE (AHL)* genes, and that both pathways are important components of BBM-induced SE.

Introduction

Plant growth relies on maintenance of the stem cell niches from which all post-embryonic organs arise. BABY BOOM (BBM) and other members of the AINTEGUMENTA-LIKE (AIL) clade of AP2 transcription factors (AINTEGUMENTA (ANT), AIL1, PLETHORA1 (PLT1), PLT2, AIL5/PLT5, AIL6/PLT3 and PLT7) play important roles in stem cell maintenance and organ growth (Horstman et al., 2014). For example, in *Arabidopsis thaliana* (Arabidopsis), development is compromised when a mutant *BBM* allele is combined with mutations in other *AIL* genes: the *plt2;bbm* double mutant arrests at the 2-celled stage of embryogenesis and combinations of *bbm*, *plt1* and *plt3* have short roots as a result of meristem differentiation (Galinha et al., 2007). By contrast, overexpression of *BBM* induces the formation of somatic embryos on cotyledons and leaves (Boutilier et al., 2002), supporting its role in regulating cell identity and proliferation.

Although the functions of *BBM* and other *AIL* genes have been well described, the molecular mode of action of these transcription factors has received less attention (reviewed in (Horstman et al., 2014)). AIL proteins contain two DNA-binding AP2 domains separated by a linker region (Riechmann and Meyerowitz, 1998). The crystal structure of AP2/ERF proteins has only been resolved for ERF1, which contains only a single AP2 domain. The AP2 domain of ERF1 forms a three-stranded anti-parallel β -sheet that binds DNA and an α -helix of which the function is unknown (Allen et al., 1998). Homology modelling of ANT on the ERF1 crystal structure suggests that, like ERF1, each AP2 repeat of ANT also forms an α -helix. However, unlike ERF1, the first AP2 repeat of ANT is predicted to form two instead of three β -sheets, while the second AP2 repeat is not predicted to form any β -sheets (Krizek, 2003).

AIL DNA binding studies have been performed using *in vitro* methods, with by far the majority of research focussed on the ANT protein. Both AP2 domains of ANT are required for DNA binding and each domain is thought to use different amino acids to contact the DNA (Nole-Wilson and Krizek, 2000; Krizek, 2003). A SELEX study based on reiterative *in vitro* selection of ANT binding sites showed that ANT binds to the DNA consensus sequence gCAC(A/G)N(A/T)TcCC(a/g)ANG(c/t), which shows similarity to the CCGA core binding site of other AP2/ERF proteins (Nole-Wilson and Krizek, 2000). It was proposed that the first and second AP2 domain of ANT bind to the 5' and 3' part of this sequence, respectively, with the ANT linker region serving as a bridge between these two domains (Nole-Wilson and Krizek, 2000). PLT5/AIL5 is also able to bind to the same DNA sequence (Yano et al., 2009), suggesting that different AILs have similar *in vitro* DNA binding properties. It is not known how DNA binding

specificity is achieved for the different AIL proteins, neither are there genome-wide *in vivo* DNA binding data available.

Genetic evidence suggests that *AIL* genes are hubs in a wide range of genetic networks, many of which are centred on the plant growth regulator auxin. AIL proteins interact with auxin pathways throughout plant development and at multiple levels, including via ARFs and through PIN feedback loops (Aida et al., 2004; Blilou et al., 2005; Krizek, 2009; Krizek, 2011b, a; Prasad et al., 2011; Pinon et al., 2013; Horstman et al., 2014). However, most of the known AIL genetic interactions are indirect, and with the exception of BBM, the direct transcriptional output of AIL transcription factors is not known (Horstman et al., 2014).

Our previous study provided the first insight into BBM molecular function by identifying BBM targets through a microarray-based approach (Passarinho et al., 2008), but only a few of these BBM targets have been functionally characterized, and only a few provided clear links with known regulators of cell proliferation or somatic embryogenesis. For example, besides stimulating its own expression, *BBM* was also shown to up-regulate expression of *ROOT GROWTH FACTOR8/GOLVEN6 (RGF8/GLV6)*, which encodes a homolog of the secreted peptide RGF1 that has a role in root meristem maintenance (Matsuzaki et al., 2010). *BBM* was also shown to regulate expression of *ACTIN DEPOLYMERIZING FACTOR9 (ADF9)*, which regulates actin filament turnover and is required for cell proliferation *in vitro* (Burgos-Rivera et al., 2008).

To obtain a better understanding of *BBM* function, we identified *in vivo* *BBM* target genes in somatic embryos using chromatin immunoprecipitation (ChIP) combined with massively-parallel DNA sequencing (ChIP-seq). We examined the transcriptional regulation of a subset of *BBM* target genes with known roles in auxin biosynthesis and somatic embryo induction. We also identified a *BBM* consensus binding motif, GCAC(A/G)NNNN(C/T)CNAN(A/G), which was enriched in the ChIP-seq peaks and resembles the reported ANT binding motif (Nole-Wilson and Krizek, 2000), as well as a second motif that is similar to the Basic Pentacystein (BPC) binding motif. Our results support the concept of AIL proteins as ‘hubs in a plethora of networks (Horstman et al., 2014), and provide a direct link to known AIL genetic pathways, as well as a novel downstream pathways.

Material and methods

Plant material and growth conditions

The T-DNA insertion lines *bbm-1* (SALK_097021) and *plt2* (SGT4287) were obtained from NASC. The *p35S::BBM-GR* and *DR5rev::GFP* lines were described previously (Boutilier et al., 2002; Benkova et al., 2003; Galinha et al., 2007; Passarinho et al., 2008).

Plants were grown at 21°C (16 h light/8 h dark regime) on rock wool plugs supplemented with 1 g/L Hyponex fertilizer, or in Petri dishes on medium containing half-strength Murashige and Skoog salts and vitamins (pH 5.8) supplemented with 0.8% agar and 1% sucrose (0.5MS-10).

ChIP-seq

ChIP-seq experiments and data analysis were carried out as described in Chapter 3. Somatic embryos obtained from either 2,4-D-induced cultures or from a BBM overexpression line were used for ChIP. Somatic embryos from a *pBBM::NLS-GFP* line, or embryogenic *p35S::BBM* seedlings served as negative controls for the *pBBM::BBM-YFP* and *p35S::BBM-GFP* ChIPs, respectively. Sequencing and mapping against the Arabidopsis genome yielded from 110 to 160 million uniquely mapped reads for the four immunoprecipitations. ChIP-seq results were visualized using Integrated Genome Browser (IGB) 8.0.14 (Nicol et al., 2009). The ChIP-seq data is available via NCBI (GEO accession: GSE52400).

DNA binding motif analysis

For motif identification, 201 bp DNA sequences (± 100 bp around the peak summits) were submitted to MEME version 4.0.0 (Bailey and Elkan, 1994), using the top 443 *pBBM* ChIP peaks (ChIP-seq score >9) and default MEME parameters.

Gene expression analysis

The effect of BBM overexpression on *AHL* gene expression was examined by inducing five day-old Col-0 and *p35S::BBM-GR* seedlings (four biological replicates of each) for three hours with 10 μ M dexamethasone (DEX) plus 10 μ M cycloheximide (CHX). RNA was isolated using the Invitex kit, treated with DNaseI (Invitrogen) and then used for cDNA synthesis with the Taqman cDNA synthesis kit (Applied Biosystems). Quantitative real-time RT-PCR (qPCR) was performed using the SYBR green mix from BioRad (3 minutes 95°C followed by 40 cycles of 15 seconds 95°C and 1 minute 60°C).

The effect of BBM overexpression on the expression of auxin biosynthesis genes was examined by inducing one and five day-old Col-0 and *p35S::BBM-GR* seedlings (three biological replicates of each) for three hours with 10 μ M DEX plus 10 μ M CHX. RNA was isolated using the NucleoSpin RNA kit (Machery-Nagel) in combination with Plant RNA Isolation Aid (Ambion), treated with DNA-free (Ambion) and then used for cDNA synthesis with M-MLV Reverse Transcriptase (Life Technologies). qPCR analysis was performed using the BioMark HD System (Fluidigm) as described in Chapter 3.

The relative expression level of the target genes was calculated according to the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001) using wild-type Col-0 as the calibrator and the *SAND* gene (At2g28390; (Czechowski et al., 2005)) as the reference. The PCR primers are listed in Supplemental Table 1.

Results

Experimental set-up

To date, all knowledge on DNA binding by AIL transcription factors has been obtained using *in vitro* methods. We therefore performed a genome-wide analysis of *in vivo* BBM DNA binding sites using ChIP-seq. Given the native and overexpression functions of BBM in embryogenesis we used both the pBBM and the p35S promoters to drive expression of a BBM-YFP or BBM-GFP protein, respectively, in somatic embryo tissue (Figure 1).

We confirmed the functionality of the BBM-YFP and BBM-GFP fusion proteins prior to ChIP-seq by crossing a *pBBM::BBM-YFP* line into the *bbm^{+/-};plt2* background and evaluating its ability to complement the embryo lethal phenotype of the *bbm;plt2* double mutant (Galinha et al., 2007). Siliques derived from a selfed *bbm^{+/-};plt2* mutant contained seeds that either lacked embryos or contained two-celled embryos (~30%), as well as seeds with older embryos (~70%), while wild-type siliques only contained embryos of similar stages (Figure 2). This suggests that the *bbm/plt2* double mutant arrests at the two-celled embryo stage or earlier. The progeny of *pBBM::BBM-YFP;bbm;plt2^{+/-}* plants showed a normal 1:2:1 segregation ratio of the *plt2* allele (16:31:19), indicating that the *pBBM::BBM-YFP* construct completely complements the *bbm;plt2* mutant (Figure 2). The functionality of the BBM-GFP fusion was also confirmed by the ability of *35S::BBM-GFP* seedlings to produce somatic embryos.

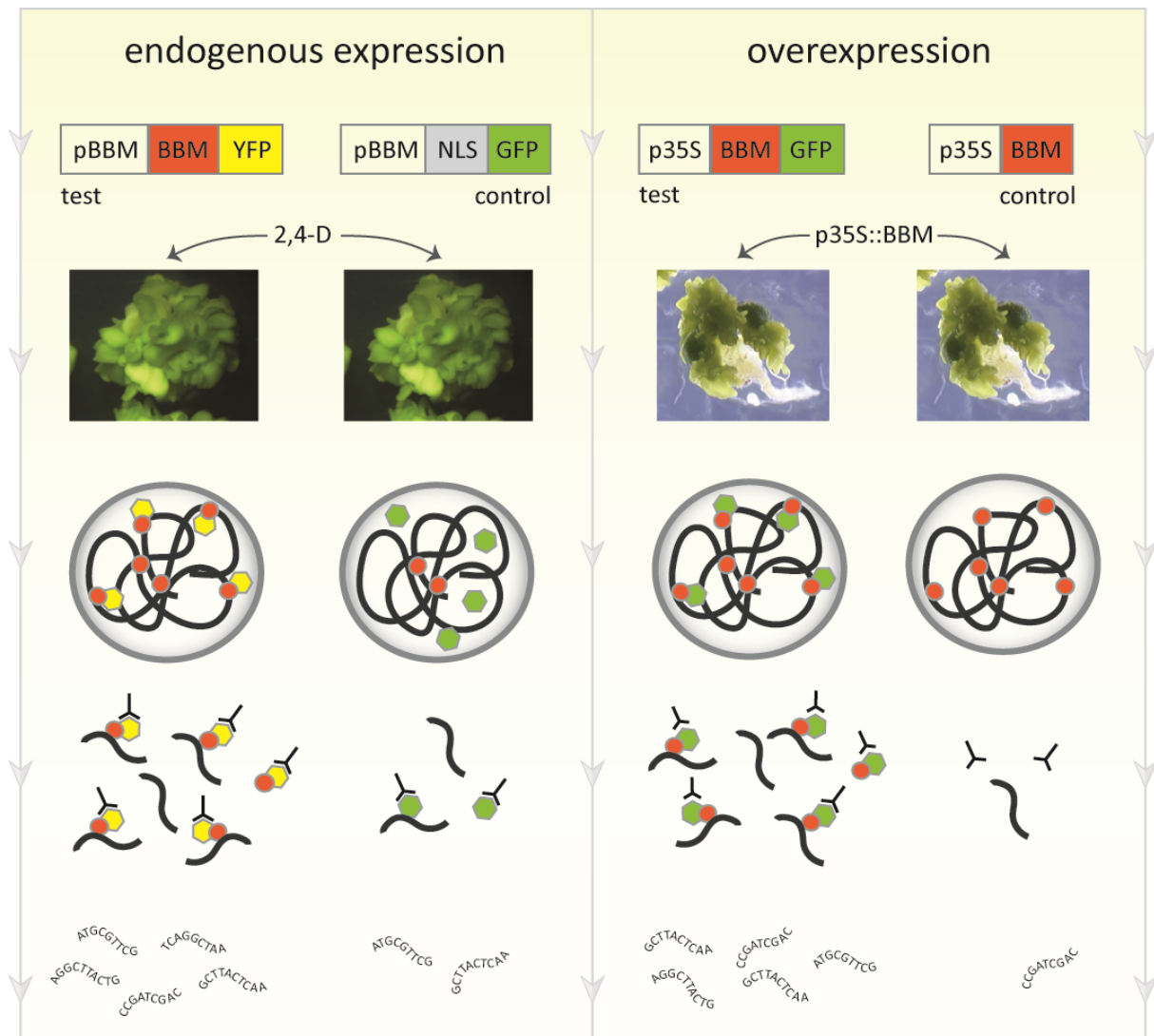


Figure 1. The experimental setup of the ChIP-seq experiments

Overview showing the tissue, DNA constructs, negative controls and experimental procedure used in the two BBM ChIP-seq experiments. SE was either induced by auxin (2,4-D) in the case of endogenous expression or by overexpression of BBM under the control of the 35S promoter. Red circles represent the BBM protein, yellow and green hexagons represent YFP and GFP, respectively, and Δ represents the GFP antibody. The bound and unbound DNA fragments in the control samples represent the background noise that is also present in the test samples.

Genome-wide identification of BBM binding sites

To identify potential BBM targets, we searched for genes that contained one or more binding sites within 3 kb upstream of the 5' end and 1 kb downstream of the 3' end of the annotated gene. Each binding site was also assigned a ChIP-seq score that represents the height of the DNA binding peak. Using a cut-off of $FDR < 0.01$, we identified 1016 genes as putative BBM targets in the *pBBM::BBM-YFP* ChIP, and 21,421 genes in the *p35S::BBM-GFP* ChIP. The difference in target numbers might be caused by ChIP efficiency differences.

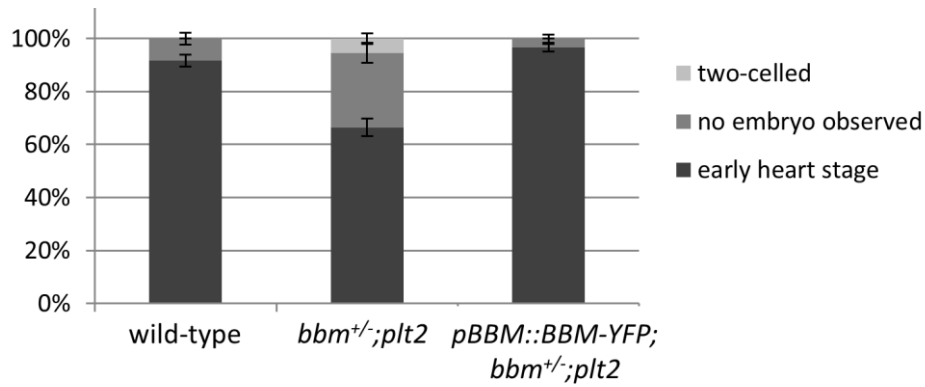


Figure 2. The *pBBM::BBM-YFP* construct rescues the *bbm;plt2* embryo phenotype.

Seeds from wild-type, *bbm^{+/-};plt2* and *pBBM::BBM-YFP;bbm^{+/-};plt2* silques were cleared and the percentage of different embryo types was quantified. Error bars represent the standard errors of five replicates (siliques).

This is consistent with the higher *p35S* ChIP-seq scores compared to *pBBM* ChIP-seq scores for the same genes (Figure 3), and could be due to a more efficient IP of the BBM-GFP protein compared to BBM-YFP, differences in fusion protein expression levels or tissue-specific differences. Alternatively, the control for the *p35S* ChIP (*p35S::BBM*) may not reflect non-specific binding as accurately as the control for the *pBBM* ChIP (*pBBM::NLS-GFP*) does. Because of these reasons we reduced the number of *p35S::BBM-GFP* candidate target genes by limiting our further analysis to the top 1000 genes with the highest ChIP-seq score. The DNA binding site with the highest ChIP-seq score in the *pBBM* experiment is located within the *BBM* gene itself (Table 1), which is in agreement with our earlier work showing that BBM upregulates its own expression (Passarinho et al., 2008). Although the set-up and material for both ChIP experiments were different (endogenous expression versus overexpression, 2,4-D induced somatic embryos versus BBM-induced seedlings and different negative controls), the correlation between the ChIP-seq data was high ($R = 0.72$; Figure 3).

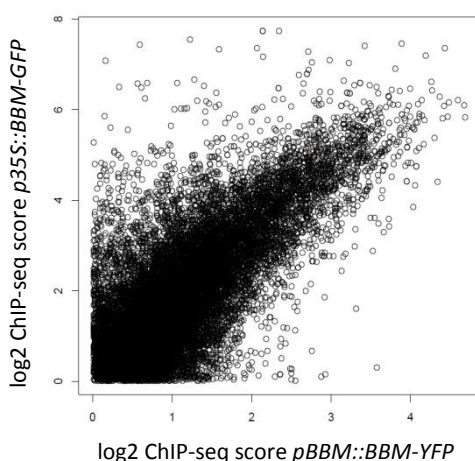


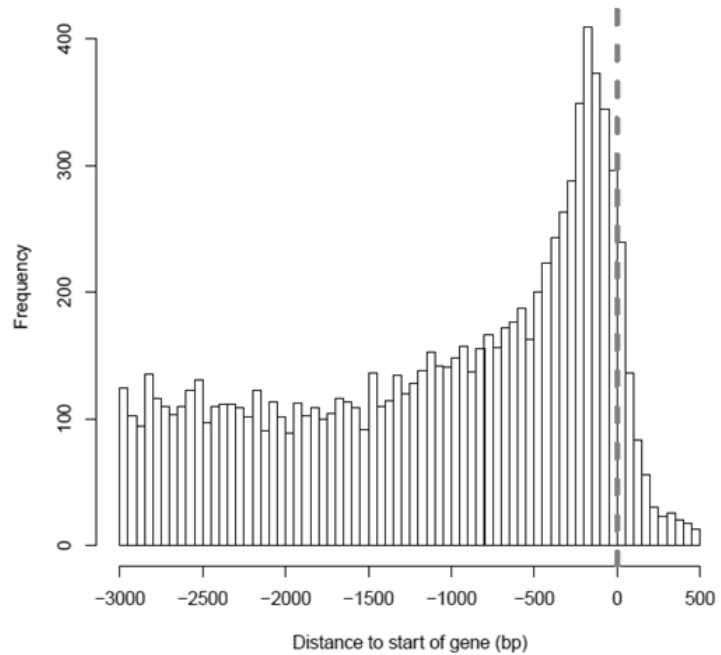
Figure 3. Correlation between the *pBBM::BBM-YFP* and *p35S::BBM-GFP* ChIP-seq data

Scatterplot showing the correlation between the ChIP-seq experiments, where each circle represents the maximum log₂ ChIP-score value on the promoter region (1kb upstream of the start of the gene) of an Arabidopsis gene. The Pearson coefficient (R) = 0.73.

Transcription factors regulate target gene expression by binding to *cis*-regulatory elements, which are DNA sequences that are mostly located in the proximal promoters of genes. We analysed the spatial distribution of BBM binding sites within 3kb upstream and 0.5kb downstream of the transcription start site of genes and found that BBM predominantly binds in a region spanning a few hundred base pairs upstream of the transcriptional start site of genes (Figure 4), which has also been shown for other Arabidopsis transcription factor ChIP-seq data sets (Heyndrickx et al., 2014).

Figure 4. BBM preferentially binds close to transcriptional start sites

Frequency of the pBBM ChIP-seq peak maxima positioned relative to the transcriptional start of genes (zero position), as annotated in TAIR 10.



We also examined the overlap in the *pBBM* and *p35S* ChIP experiments by comparing the top 1000 binding sites (syn. ChIP-seq scores) in each experiment. Approximately 45% of the binding sites were found in both experiments (Figure 5). The overlap became increasingly smaller with a more stringent list of top binding sites (Figure 5). These data indicate that there is substantial overlap between the DNA binding sites in both experiments, but their absolute ranking is not well conserved.

	pBBM		p35S
top 100	82	18	82
top 200	148	52	148
top 500	318	182	318
top 1000	547	453	547

Figure 5. Overlap between the *pBBM* and *p35S* ChIP target genes.

The number of unique (light grey) and overlapping (dark grey) targets in the top 100/200/500/1000 genes of both ChIPs are shown.

The BBM binding motif resembles that of ANT

We used MEME (Bailey and Elkan, 1994) to identify *cis*-regulatory motifs bound by BBM. This analysis yielded a TC-rich motif (Figure 6; motif 1; *E*-value = 2.9E-46), which resembles the GA-rich binding motif of BASIC PENTACYSTEIN (BPC) proteins (reverse complemented) (Meister et al., 2004; Kooiker et al., 2005). However, this motif is not enriched in the centre of the ChIP-seq peaks, indicating that this is likely not the BBM DNA binding motif.

A second sequence motif GCAC(A/G)NNNN(C/T)CNAN(A/G) was found enriched in the centre of the peaks (Figure 6; motif 2; *E*-value = 1.2E-36), suggesting that it represents a *bona fide* BBM binding motif. The BBM binding motif resembles the ANT binding motif, gCAC(A/G)N(A/T)TcCC(a/g)ANG(c/t) (Nole-Wilson and Krizek, 2000) (Figure 7), which was also shown to be bound by AIL5/PLT5 (Yano et al., 2009). The flanking regions of the ANT binding sequence, gCAC(A/G) and CC(a/g)ANG, are similar to the BBM binding motif, particularly the 5' flanking motif, while the core of the motif is not conserved (Figure 6). Our results provide evidence that different AIL transcription factors bind a similar, but not identical DNA sequence.

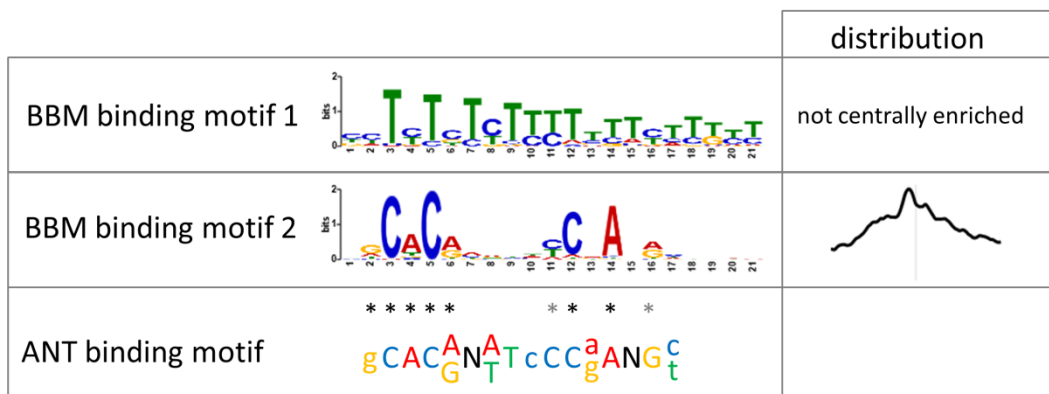


Figure 6. BBM binds to a DNA sequence similar to the ANT DNA binding motif

DNA motifs that were enriched in BBM immunoprecipitated regions and their distribution within these regions. The motifs are displayed as Position Weight Matrices that specify the probability of the occurrence of each possible nucleotide at each position in the motif. The height of the stack is the information content in bits (Schneider and Stephens, 1990). The similarity between BBM binding motif 2 and the ANT binding motif (Nole-Wilson and Krizek, 2000) is indicated with asterisks (black, highly similar; grey, similar).

Table 1. BBM targets involved in auxin, embryogenesis, root development and adaxial/abaxial polarity specification

Locus	Gene	Gene function	<i>pBBM</i> top X	<i>p35S</i> top X
Auxin				
AT1G04610	<i>YUC3</i>	auxin biosynthesis	100	200
AT4G28720	<i>YUC8</i>	auxin biosynthesis	1000	1000
AT1G70560	<i>TAA1</i>	auxin biosynthesis	1000	
AT3G51060	<i>STY1</i>	activator of <i>YUC</i> genes	200	
AT3G28860	<i>ABCB19</i>	auxin transport	100	200
AT1G73590	<i>PIN1</i>	auxin transport	500	200
AT2G01420	<i>PIN4</i>	auxin transport	100	1000
AT3G23050	<i>IAA7</i>	auxin response	100	1000
AT4G39403	<i>PLS</i>	auxin response	200	500
AT5G62000	<i>ARF2</i>	auxin response	200	1000
AT2G28350	<i>ARF10</i>	auxin response		500
AT5G25890	<i>IAA28</i>	auxin response	1000	1000
AT3G23030	<i>IAA2</i>	auxin response	1000	
AT1G30330	<i>ARF6</i>	auxin response	1000	500
Embryogenesis				
AT5G17430	<i>BBM</i>	SE inducer, embryo/root development	100	
AT1G51190	<i>PLT2</i>	SE inducer, embryo/root development	500	
AT5G57390	<i>AIL5/PLT5</i>	SE inducer, germination, phyllo-/rhizotaxy	200	1000
AT1G21970	<i>LEC1</i>	SE inducer, embryo development/maturation	100	500
AT1G28300	<i>LEC2</i>	SE inducer, embryo development/maturation	500	500
AT3G20910	<i>NF-YA9</i>	SE inducer, embryo development/maturation	500	
AT3G24650	<i>ABI3</i>	embryo development/maturation	500	
AT4G12080	<i>AHL1</i>			1000
AT1G63480	<i>AHL12</i>		1000	100
AT3G55560	<i>AHL15</i>	SE inducer	1000	1000
AT5G49700	<i>AHL17</i>			500
AT3G60870	<i>AHL18</i>		1000	
AT3G04570	<i>AHL19</i>		1000	
AT2G45430	<i>AHL22</i>		1000	
AT4G12050	<i>AHL26</i>			1000
Root development				
AT3G54220	<i>SCR</i>	root patterning	100	500
AT1G79580	<i>SMB</i>	root cap	200	
AT1G11130	<i>SUB</i>	root epidermis	500	
AT4G37650	<i>SHR</i>	root patterning	500	
AT5G60810	<i>RGF1</i>	maintenance of the root stem cell niche	500	
AT2G04025	<i>RGF3</i>	maintenance of the root stem cell niche	500	
AT3G30350	<i>RGF4</i>	maintenance of the root stem cell niche		1000
AT2G03830	<i>RGF8</i>	maintenance of the root stem cell niche	500	1000
AT5G14750	<i>WER</i>	root epidermis	500	500
AT5G03150	<i>JKD</i>	root patterning	500	100
AT4G00730	<i>ANL2</i>	root epidermis	500	500
AT2G27230	<i>LHW</i>	root vasculature	1000	
AT1G46264	<i>SCZ</i>	root patterning	1000	
AT2G46410	<i>CPC</i>	root epidermis	1000	
Abaxial/adaxial polarity				
AT5G16560	<i>KAN1</i>	abaxial identity	200	500
AT1G32240	<i>KAN2</i>	abaxial identity	1000	1000
AT4G00885	<i>miR165B</i>	abaxial identity (targets <i>PHV</i> , <i>PHB</i> , <i>REV</i>)		100
AT3G61897	<i>miR166B</i>	abaxial identity (targets <i>PHV</i> , <i>PHB</i> , <i>REV</i>)	500	500
AT2G34710	<i>PHB</i>	adaxial identity		100
AT5G60690	<i>REV</i>	adaxial identity		1000

Functional categories of the top 1000 BBM-YFP and BBM-GFP target genes (based on ChIP-seq score, which represents the height of a ChIP-seq peak), together with their ranking (top 100/200/500/1000) in their respective ChIP datasets. The categories are a combination of BiNGO- and author-assigned categories. SE, somatic embryogenesis.

BBM binds to genes with diverse functions

We performed a BiNGO analysis (Maere et al., 2005) on the top 1000 target genes in both CHIP experiments to identify overrepresented gene ontology (GO) categories (not shown). As expected from BBM's function in the root, categories related to root development (Table 1), meristem initiation and meristem maintenance were overrepresented. In addition, the top 1000 targets included several genes involved in adaxial/abaxial polarity specification (Table 1) and shoot development, including *ASYMMETRIC LEAVES 2-LIKE 1 (ASL1)*, *CUP-SHAPED COTYLEDON 2 (CUC2)*, *microRNA319a (miR319A/JAW)*, *ROTUNDIFOLIA 2/10/15/18*, and *TUMOROUS SHOOT DEVELOPMENT 2 (TSD2)*. BBM also bound to genes involved in (somatic) embryogenesis (SE) (Table 1), including the *LEAFY COTYLEDON (LEC)* genes (discussed in Chapter 4), *NUCLEAR FACTOR Y SUBUNIT A9 (NF-YA9)* (Mu et al., 2013) and the recently identified *AT-HOOK LIKE (AHL)* genes (Omid Karami, PhD thesis, Leiden University, unpublished), which are discussed below. Finally, many genes related to auxin biosynthesis, transport and signalling were also identified (Table 1). Most of these genes were among the top 1000 targets in both the *pBBM* and *p35S* CHIP experiments, but often with a different ranking (Table 1). One remarkable difference between the two CHIPs is that in the *pBBM* CHIP, BBM bound only to genes that promote abaxial identity, while in the *p35S* CHIP, BBM bound to both abaxial- and adaxial-specifying genes. Interestingly, the adaxial-specifying *PHABULOSA (PHB)* was shown to occasionally induce SE through direct upregulation of *LEC2* (Tang et al., 2012).

Previously, direct BBM target genes were identified that were differentially expressed in four day-old *p35S::BBM-GR* seedlings after an eight hour induction with dexamethasone (DEX) (Passarinho et al., 2008). In chapter 4, we showed that *p35S::BBM-GR* seedlings of this developmental stage mostly produce somatic embryos from callus (indirect SE), while younger seedlings form somatic embryos via direct SE. We observed that there is overlap between the CHIP and microarray dataset, but surprisingly, only few of these target genes were found in the top 1000 BBM targets identified by CHIP-seq (Table 2).

BBM activates *AHL* genes

We observed BBM binding to several members of the AT-hook containing, nuclear localized protein (*AHL*) gene family (Table 1). Interestingly, a few of these DNA-binding proteins were found to promote SE in Arabidopsis (Omid Karami, PhD thesis, Leiden University, unpublished) (Figure 7A, B). In contrast to *p35S::BBM* seedlings, *p35S::AHL15* seedlings also produced rooty callus (Figure 7A). *AHL15* overexpression also enhances 2,4-D-induced SE from immature zygotic embryos (Figure 7C), and in accordance, *ahl15* mutant explants show a reduced response in the same system (Figure 7D). This negative effect on somatic embryo formation was enhanced in a triple *ahl15;ahl19;amiRAHL20* mutant, in which the expression of two close homologs, *AHL19* and *AHL20*, was also reduced (Figure 7D).

BBM bound to the promoter regions of *AHL15*, *AHL19* and *AHL20*, close to the transcriptional start site (Figure 7E). To determine whether these genes are also transcriptionally regulated by BBM, we analysed gene expression changes in *p35S::BBM-GR* plants after DEX application, in combination with the translational inhibitor cycloheximide (CHX) to identify direct transcriptional effects. BBM stimulated the expression of *AHL15* and *AHL20*, but no statistically significant difference in *AHL19* expression was observed (Figure 7F).

Next, we investigated the requirement for *AHL* genes in BBM-induced SE by transforming the *p35S::BBM-GR* construct into the triple *ahl15;ahl19;amiRAHL20* mutant. In wild-type Col-0, this construct induces SE in 7% of the primary transformants (Chapter 4), but in the *ahl15;ahl19;amiRAHL20* mutant only 1.6% of the transformants was embryogenic (2/130). Our results suggest that BBM-induced SE is partly achieved through activation of *AHL* genes.

Table 2. Overlap between BBM targets obtained via microarray and ChIP-seq

AGI no.	Annotation	<i>pBBM</i> top X	<i>p35S</i> top X
Transcription			
At5g17430	BABY BOOM (BBM)	1000	2000
At5g39820	NAC domain protein (ANAC094)		
At1g16070	Tubby family protein (TLP8)	5000	
At1g65300	MADS-box protein (PHERES2)		
At5g46640	AT-hook protein (AHL8)	5000	5000
At3g60580	zinc finger protein C2H2-type		
At1g51140	Basic Helix-Loop-Helix protein (BHLH122)	2000	5000
At5g10960	CCR4-NOT transcription complex protein	5000	5000
Signaling			
At5g45780	leucine-rich receptor-like kinase kinase, LRRII group		5000
At2g34020	calcium-binding EF hand protein	1000	2000
At4g11320	cysteine proteinase		
At1g61610	S-locus lectin protein kinase		
At5g59100	subtilisin-like serine protease, S8 family		
Protein-protein interactions			
At5g48130	BTB-POZ domain protein, NPH3 family (NRL27)	2000	
At3g54780	RING H2 domain protein	2000	5000
At5g48510	BTB-POZ domain protein, speckle-type		
At4g38140	RING H2 domain protein		
At3g19380	U-box /armadillo domain protein	2000	2000
At4g35070	RING/U box domain protein		5000
At3g15680	zinc finger protein, RanBP2-type	2000	
Cell wall/cell membrane-localized			
At5g47440	PH domain containing protein (PH16)	1000	2000
At5g03260	laccase-like multicopper oxidase		
At4g03210	xyloglucan endotransglycosylase (XTH9)	2000	1000
At5g48900	pectate lyase (PLL21)	5000	5000
At5g01870	lipid-transfer protein (LTP6-like)		
At1g76790	O-methyltransferase	5000	5000
At4g02290	endo- β -1,4-glucanase	2000	1000
At4g27520	ENOD-like GPI-anchored arabinogalactan protein (AGP) /phytoeyanin	2000	2000
At5g48140	polygalacturonase	2000	
Other			
At3g26200	Cytochrome P450 (CYP71B22)		
At2g03830	Root growth factor8 (RGF8)	1000	1000
At4g34970	Actin-depolymerizing factor (ADF9)	5000	
At5g11890	EMBRYO DEFECTIVE 3135 (EMB3135)	1000	5000
At4g14690	EARLY LIGHT INDUCED PROTEIN2 (ELIP2)	5000	
At5g62490	AtHVA22b		
At1g64590	Short-chain dehydrogenase / reductase (SDR)		
At5g02550	Expressed protein	5000	1000
At4g02360	Expressed protein, DUF538		5000
At5g44560	SNF7 protein	5000	
At3g18800	Expressed protein	5000	5000
At2g41800	Unknown protein, DUF642	1000	5000
At3g02960	copper chaperone (ATX1)		
At3g60150	hypothetical protein, DUF598/498		

BBM target genes that were identified using a microarray-based approach (Passarinho et al., 2008) and their ranking (top 1000/2000/5000) in the *pBBM* and *p35S* ChIP datasets.

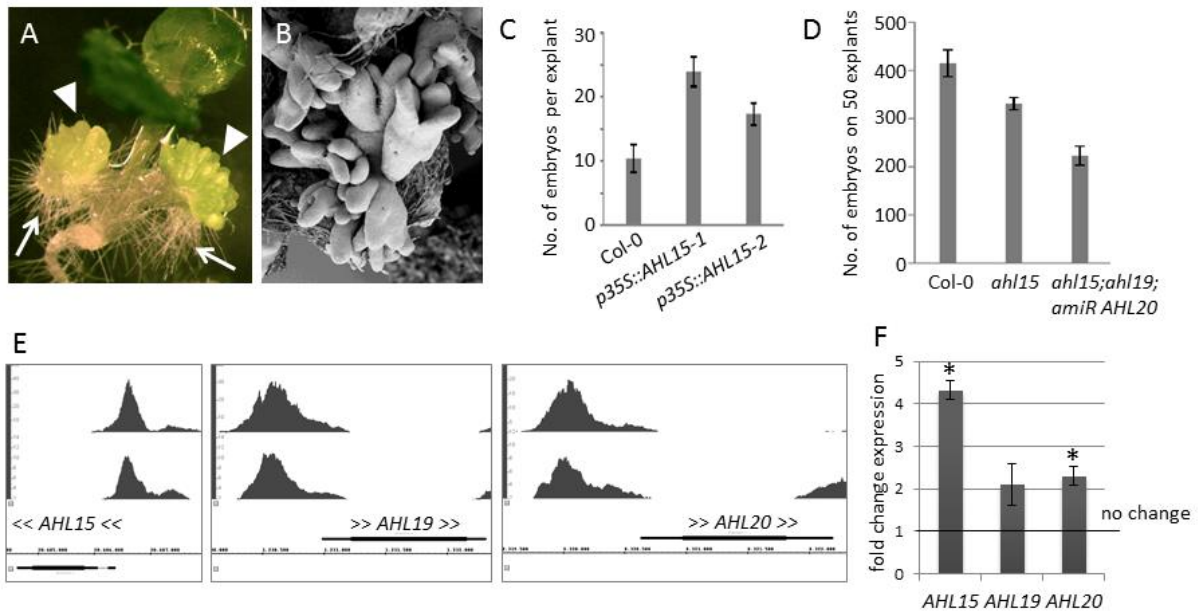


Figure 7. BBM activates SE-promoting AHL genes

(A) A two-week-old *p35S::AHL15* seedling with somatic embryos (arrow heads) and rooty callus (arrows) on the cotyledons.

(B) Scanning electron microscopy image showing somatic embryos on *35S::AHL15* cotyledons.

(C) Number of somatic embryos induced by 2,4-D in wild-type and *p35S::AHL15* immature zygotic embryos (Gaj, 2001). Error bars indicate minimum and maximum values of 50 explants per genotype.

(D) Effects of mutations in *AHL15* and its close homologs on SE-induction from immature zygotic embryos by 2,4-D. Error bars indicate the minimum and maximum values of 4 biological replicates, with 50 explants per replicate. The data shown in (A–D) were generated by Omid Karami (PhD thesis, Leiden University, unpublished).

(E) ChIP-seq BBM binding profiles for *AHL* genes. The binding profiles from the *p35S::BBM-GFP* (upper profile) and *pBBM::BBM-YFP* (lower profile) ChIP-seq experiments are shown. The x-axis shows the nucleotide position of DNA binding in the selected genes (TAIR 10 annotation), the y-axis shows the ChIP-seq score, and the brackets indicate the direction of gene transcription.

(F) The relative expression of *AHL* genes was determined by qPCR for DEX+CHX treated *35S::BBM-GR* seedlings as described in the Materials and methods. Error bars indicate standard errors of four biological replicates. Statistically significant differences (*) between *35S::BBM-GR* and Col-0 were determined using a Student's *t*-test ($p < 0.05$).

BBM-induced auxin biosynthesis is required for SE

In Arabidopsis, indole acetic acid (IAA) is the major active auxin and is mainly synthesized from tryptophan by the TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS 1 (TAA1) and YUCCA (YUC) enzymes (Mashiguchi et al., 2011; Won et al., 2011). Our ChIP-seq analysis showed that BBM binds to the second intron and last exon/3'UTR regions of *TAA1* and to the promoter regions of *YUC3* and *YUC8*. BBM also binds to the promoter region of *STYLISH1* (*STY1*; Figure 8A), which encodes a RING zinc-finger domain containing protein that was reported to directly activate *YUC4* expression and down-regulate gibberellic acid biosynthesis (Eklund et al., 2010).

TAA1, *YUC3* and *STY1* were upregulated after BBM activation in one day-old imbibed seeds, which produce somatic embryos by direct SE (Figure 8B; Chapter 4). In contrast, *YUC8* expression was activated by BBM in five day-old seedlings (Figure B), which produce somatic embryos via indirect SE (Chapter 4). These results show that BBM activates expression of different auxin biosynthesis genes at different developmental stages.

We used the *DR5* auxin response marker (Benkova et al., 2003) to follow the timing and spatial localization of auxin signalling during BBM-mediated direct SE.

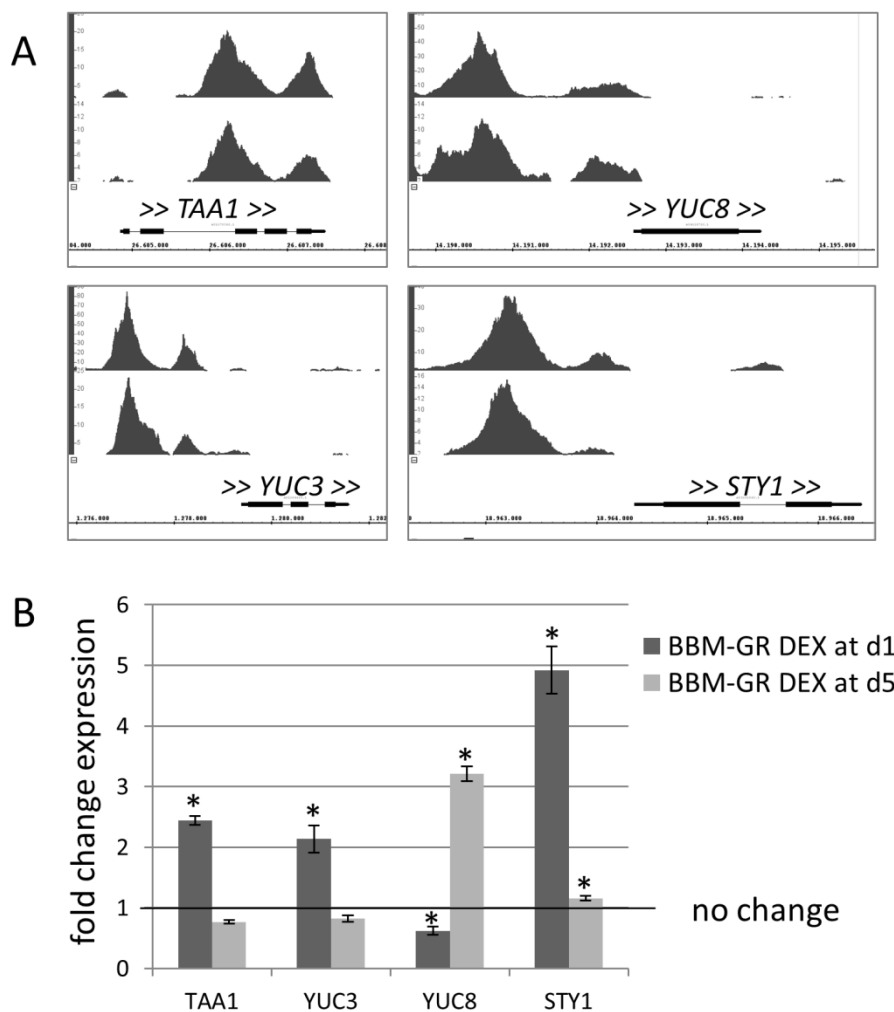


Figure 8. BBM stimulates auxin biosynthesis genes during SE

(A) ChIP-seq BBM binding profiles for auxin biosynthesis genes and *STY1*. The binding profiles from the *p35S::BBM-GFP* (upper profile) and *pBBM::BBM-YFP* (lower profile) ChIP-seq experiments are shown. The x-axis shows the nucleotide position of DNA binding in the selected genes (TAIR 10 annotation), the y-axis shows the ChIP-seq score and the brackets indicate the direction of gene transcription.

(B) The relative expression of auxin-related genes was determined by qPCR for DEX+CHX treated *p35S::BBM-GR* seedlings as described in the Materials and methods. Error bars indicate standard errors of the three biological replicates. Statistically significant differences (*) between *p35S::BBM-GR* and Col-0 were determined using a Student's *t*-test ($p < 0.05$).

GFP expression was first observed in *p35S::BBM-GR* seedlings three days after DEX was applied to imbibed seeds, where it was localized to the margin on the adaxial side of cotyledons (Figure 9A, +3). The *GFP* signal spread throughout the adaxial cotyledon surface during the following days, and was later excluded from the region at the tip of the cotyledon where the first somatic embryos developed (Figure 9A, +4/+5). Later, other regions that formed growth protrusions, likely sites of somatic embryogenesis, also lacked *DR5::GFP* expression (Figure 9A, +6/+7). Our results confirm that BBM induces an auxin response in cotyledons, and suggest that this response becomes reduced at sites where somatic embryos develop.

We used a pharmacological approach to investigate whether auxin biosynthesis via YUC proteins is required for BBM-mediated SE. *p35S::BBM-GR* and Col-0 seedlings were grown in the presence of DEX and/or the YUC inhibitor yucasin (Nishimura et al., 2014). Yucasin blocks the conversion of indole-3-pyruvic acid (IPyA) to IAA, by competing with YUC for binding to IPyA (Nishimura et al., 2014). Unlike DEX-treated cotyledons, we did not observe *GFP* in cotyledons of *DR5::GFP;p35S::BBM-GR* seedlings treated with both DEX and yucasin, showing that the auxin biosynthesis is repressed to a level below that detected by *DR5* (Figure 9B). SE was also abolished in *p35S::BBM-GR* seedlings grown in the presence of DEX and yucasin (Figure 9B). Instead, cotyledons became yellowish and callus formed at the shoot apex, from which root-like structures grew (Figure 9B). This phenotype was specific for DEX-activated BBM overexpression lines, since yucasin did not induce a similar phenotype in Col-0 seedlings (Figure 9B). Our data suggest that auxin biosynthesis via the YUC pathway is crucial for BBM-mediated SE.

BBM induces common and unique SE pathways

Transcription factors other than BBM/AIL proteins (Chapter 4), also promote SE in *Arabidopsis*. Several of the “LAFL” (for *LEAFY COTYLEDON 1 (LEC1)/LEC1-LIKE (L1L)*; *ABSCISIC ACID (ABA)-INSENSITIVE3 (ABI3)*; *FUSCA3 (FUS3)*; *LEC2*; (Jia et al., 2013)) seed maturation genes, namely *LEC1*, *L1L/NUCLEAR FACTOR Y subunit A6 (NF-YA6)* and *LEC2* can induce spontaneous SE in seedlings when overexpressed (Lotan et al., 1998; Stone et al., 2001; Mu et al., 2013). Overexpression of *FUS3* does not induce SE, but leads to cotyledon-like leaves (Gazzarrini et al., 2004), showing that it does ectopically activate embryogenesis traits. In addition to *L1L/NF-YA6*, overexpression of other NF-Y subunits, A1, 5 and 9, has the same SE-inducing effect (Mu et al., 2013). Finally, the *LEC2* target *AGAMOUS-LIKE 15 (AGL15)* can enhance SE from immature zygotic embryos (Harding et al., 2003). ChIP experiments have been performed to elucidate the *in vivo* DNA binding sites of some of these factors.

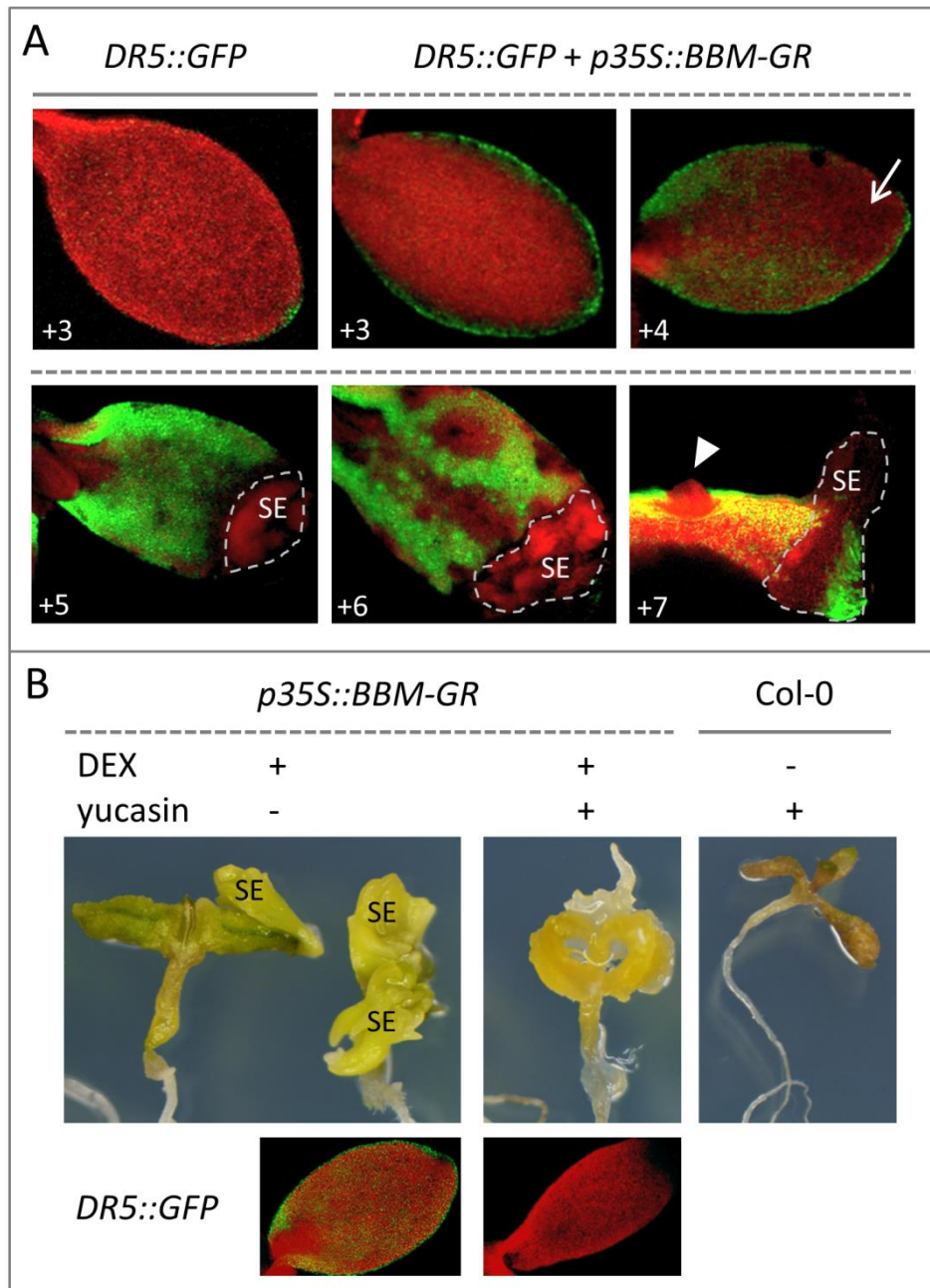


Figure 9. BBM-induced auxin biosynthesis is required for SE induction

(A) *DR5::GFP* and *DR5::GFP + p35S::BBM-GR* seeds were treated with 10 μ M DEX at the start of culture (d0) and cultured for seven days. The day of culture is indicated on the bottom left of each picture. The images show the adaxial sides of cotyledons. GFP was observed three days after BBM-GR activation at the cotyledon margin and the signal spread in the days thereafter. The arrow at +4 indicates the GFP-negative area on the distal end of the cotyledon blade. SE, somatic embryos; arrowhead at +7, a GFP-negative protrusion that will grow into a somatic embryo. Red autofluorescence was used to delineate the tissue. The outline of somatic embryos is shown with dashed lines.

(B) *p35S::BBM-GR* and Col-0 seedlings treated for 14 days with 10 μ M DEX and/or 100 μ M yucasin. SE, somatic embryos formed on *p35S::BBM-GR* seedlings treated with DEX. DEX+yucasin treatment of *p35S::BBM-GR* seedlings leads the development of root-like structures from callus formed at the shoot apex. The lower images show cotyledons of six day-old *DR5::GFP + p35S::BBM-GR* seedlings that were treated with 10 μ M DEX and/or 100 μ M yucasin at the start of culture. The GFP signal observed in the DEX-treated seedlings was absent in the DEX plus yucasin-treated seedlings.

The targets of AGL15 and FUS3 were determined by ChIP-chip (Zheng et al., 2009; Wang and Perry, 2013), while those of LEC1 were determined through ChIP followed by hybridization to an array containing active promoters (Junker et al., 2012). To obtain insight into the extent to which these transcription factor-induced SE pathways overlap, we compared the top 1000 targets from the *pBBM* ChIP with the targets identified in the LEC1, FUS3 and AGL15 ChIP experiments (Figure 10). This analysis shows that extensive cross- and auto-regulatory loops exist between BBM/AILs and the LAFL network (Figure 10, also described in Chapter 4). In addition, *NF-YA* genes were bound by all four proteins.

Only one target gene was found in all four ChIP experiments; *CELLULOSE 1 (CEL1)*, which encodes an endo-1,4-beta-glucanase involved in cell elongation (Shani et al., 2006). This was unexpected, since cellulases have not been reported to play an important role in SE. However, *CEL1* does not appear to be a specific target for SE-inducing factors, as it was also bound by, for example, the flower-specific transcription factors APETALA1 (AP1), AP2, PISTILLATA, SEPALLATA3 and LEAFY (Kaufmann et al., 2009; Kaufmann et al., 2010; Yant et al., 2010; Winter et al., 2011; Wuest et al., 2012). Therefore, *CEL1* might be a common target of transcription factors that steer growth and development, which is reinforced by its broad expression in actively dividing tissues (Shani et al., 2006).

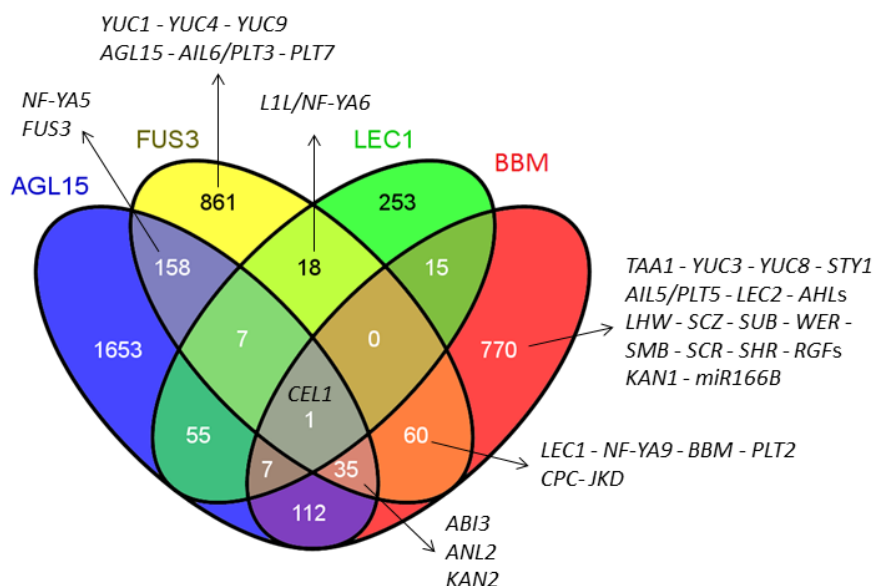


Figure 10. The overlap between ChIP targets of BBM and LAFL(-related) proteins

Venn diagram showing the overlap between the top 1000 targets of the *pBBM* ChIP-seq and the published AGL15, FUS3 and LEC1 target genes identified by ChIP (Zheng et al., 2009; Junker et al., 2012; Wang and Perry, 2013). The common and unique targets related to auxin biosynthesis, SE induction, root development and the determination of abaxial/adaxial polarity are indicated, with similar genes grouped by dashes according to the processes they are involved in (also see Table 1). *CEL1*, *CELLULOSE 1*.

The majority of root-specific target genes were unique for BBM, which is not surprising considering BBM's expression pattern and role in root meristem maintenance. The targeting of abaxial/adaxial specifying factors is not specific for BBM, as FUS3 and AGL15 also bind to *KAN2*. The auxin biosynthesis genes *YUC3* and *YUC8* are only bound by BBM (discussed above), but other members of the *YUC* gene family are bound by FUS3. The AGL15 and LEC1 ChIP targets did not include *YUC* genes, but it has been shown that LEC1 binds and regulates *YUC10* (Junker et al., 2012). In addition, LEC2 also activates the expression of *TAA1* and *YUC* genes (Braybrook et al., 2006; Stone et al., 2008; Wojcikowska et al., 2013). Therefore, YUC-mediated auxin biosynthesis emerges as a common theme among SE factors, except for AGL15. In contrast, the analysis showed that the *AHL* genes are specific targets of BBM.

Although transcription factor-DNA binding does not guarantee transcriptional regulation, these data suggest that there is cross-regulation between the different SE factors and that BBM targets both unique and overlapping classes of target genes.

Discussion

It was previously shown that BBM is a potent inducer of SE in Arabidopsis (Boutillier et al., 2002), but the downstream components of BBM-mediated SE were largely unknown. Using ChIP-seq in somatic embryos we have identified *in vivo* BBM target genes, which are involved in a wide variety of developmental processes. We showed that BBM stimulates the expression of auxin biosynthesis and *AHL* genes, and that chemical or genetic inhibition of these pathways reduced or eliminated BBM-induced SE. Our results show that in addition to the *LAFL* genes (Chapter 4), auxin biosynthesis and *AHL* genes are also important downstream components of the BBM SE pathway.

Endogenous BBM expression versus overexpression

We used two different experimental set-ups for the ChIP-seq experiments: endogenous *BBM* expression versus overexpression, 2,4-D induced somatic embryos versus BBM-induced seedlings and different negative controls. Yet, there was a substantial overlap between the target genes, making these overlapping genes highly reliable BBM targets as they are bound irrespective of the setup. The genes that are bound by BBM in only one of the ChIP experiments could reflect tissue differences, BBM expression level differences or general variability between

ChIP experiments (non-specific targets). At this point, we cannot distinguish between these possibilities, but including biological replicates could reduce the variance within each ChIP.

In zygotic embryos, *BBM* is expressed throughout the embryo at early embryo stages and then becomes progressively more basally expressed during embryo development (Chapter 3). However, the *pBBM::BBM-YFP* expression pattern may differ in 2,4-D induced SE cultures as *AIL* expression was shown to be auxin-inducible (Aida et al., 2004). Conversely, the *p35S* promoter is considered to confer high and constitutive expression, but during embryogenesis it is mainly expressed from the torpedo stage onward (Johnson et al., 2005). Expression analysis of *pBBM::BBM-YFP* in 2,4-D induced somatic embryos and of *p35S::BBM-GFP* in embryogenic seedlings could reveal tissue-specific expression in each system, as well as any quantitative differences in BBM-YFP/GFP levels.

BBM binding versus gene expression regulation

Previously, we identified direct BBM targets that were upregulated after DEX-activation of BBM-GR in seedlings (Passarinho et al., 2008). Here, we show that there is overlap between these microarray data and the ChIP-seq data, but that the differentially expressed BBM targets in the microarray experiment were not the strongest BBM-bound genes in the ChIP experiment. Poor correlations between ChIP and microarray experiments are not uncommon, as transcription factor binding does not necessarily lead to gene expression changes (Oh et al., 2009; Zheng et al., 2009; Pajoro et al., 2014). In addition, the limited overlap in target genes may be partially due to the use of different tissues, as the ChIP-seq was performed on somatic embryo tissue, while the microarray experiment was performed on four day-old seedlings. In Chapter 4, we show that BBM induces indirect SE in seedlings of this stage, while BBM induces direct SE at earlier developmental stages. The microarray data may therefore have identified target genes that regulate the early events of indirect SE (e.g. dedifferentiation, callus formation). Consistently, the *LAFL* (Chapter 4) and auxin biosynthesis genes that are found in the ChIP experiment, were mostly upregulated in one day-old imbibed seeds and not in five day-old seedlings. However, some ChIP target genes were upregulated in five day-old seedlings (e.g. *YUC8*, *AHL15*, *AHL20*), but were missed in the microarray study. The gene expression change of these targets may have been below the used cut-off, or probes for these genes may not have been present on the used microarray.

BBM DNA binding sites

The *in vivo* BBM-bound DNA sequences showed enrichment of a DNA motif that resembles the ANT motif, which was obtained via an *in vitro* method (SELEX) (Nole-Wilson and Krizek, 2000), and which is also bound by AIL5/PLT5 (Yano et al., 2009). This suggests that different AIL proteins bind similar DNA sequences and that the *in vitro* DNA binding reflects *in vivo* AIL DNA binding. However, whether AILs display subtle differences in DNA binding would require analysis of the different AILs using the same experimental setup. Potential AIL-DNA binding specificity may stem from differences in their DNA-binding domains or from different protein-protein interactions. Moreover, a single AIL may even display differential DNA binding specificity in different tissues depending on the protein partners that are present.

We also identified a second motif within the BBM binding peaks that was not centrally enriched, which indicates that this motif is not bound by the immunoprecipitated protein (BBM), but by other proteins that often bind in close vicinity of a BBM DNA binding event. The motif resembles the DNA binding sequence of BASIC PENTACYSTEIN (BPC) proteins (Meister et al., 2004; Kooiker et al., 2005) and it was shown that BPC1 regulates the expression of *SEEDSTICK* (*STK*) by binding to and inducing conformational changes of its promoter region (Kooiker et al., 2005; Simonini et al., 2012). It was suggested that BPC proteins recruit transcription factor complexes through protein-protein interactions (Simonini et al., 2012). BPC proteins were also shown to regulate the expression of *INNER NO OUTER* (*INO*), which requires ANT for its expression (Balasubramanian and Schneitz, 2002). These data, together with the occurrence of BPC binding sites in BBM-bound DNA regions, indicate that BPC proteins may also mediate AIL binding to the regulatory regions of their target genes. Future protein-protein interaction studies may reveal whether AIL proteins also physically bind to BPC proteins.

Role of auxin biosynthesis in BBM-mediated SE

We showed that BBM binds and directly activates *TAA1*, *YUC3* and *YUC8*, which encode enzymes in the auxin biosynthesis pathway. In addition, by activating *LEC1*, *LEC2* (Chapter 4) and *STY1*, BBM may also indirectly stimulate *YUC1*, *YUC2*, *YUC4* and *YUC10* expression (Stone et al., 2008; Eklund et al., 2010; Junker et al., 2012; Wojcikowska et al., 2013). Previously, it was shown that PLT5/AIL5 directly activates *YUC4* in order to control phyllotaxis in the shoot (Pinon et al., 2013), indicating that the *YUC* genes may be conserved targets among AIL transcription factors. Wójcikowska *et al* (Wojcikowska et al., 2013) showed that the *yuc2* and *yuc4* mutants have a reduced SE efficiency in 2,4-D induced SE from immature zygotic embryos (IZEs). 2,4-D induced

secondary SE also relies on YUCs, as the *yuc1;yuc4;yuc10;yuc11* mutant shows a severely reduced SE-response (Bai et al., 2013). Our finding that chemical inhibition of YUCs by yucasin eliminates BBM-mediated SE again demonstrates the requirement of YUC activity for SE induction and suggests that the SE pathways induced by 2,4-D, LEC2 and BBM converge at the level of *TAA1/YUC* gene activation. Still, *YUC* overexpression in seedlings is not sufficient to trigger SE, but rather results in epinastic growth of cotyledons (Woodward et al., 2005; Cheng et al., 2006; Kim et al., 2007). However, *YUC* overexpression in IZEs has not been investigated and might induce SE in the absence of other SE inducers, given the inherent embryogenic competence of the explant.

Our analysis of *in vivo* BBM binding sites showed that there is extensive cross-regulation between known inducers of SE, and that the pathways converge at and rely on auxin biosynthesis. In addition, the ChIP revealed the *AHL* genes as specific downstream components of the BBM SE pathway, and many more BBM target genes, some with links to SE (Tang et al., 2012). Future work should focus on whether these target genes also play a role in BBM-mediated SE.

Supplemental Table 1: Primers used in this study

Primer used for	description	5' to 3' sequence
genotyping <i>bbm</i>	Fw primer	TGCATAGAACAAGACCAAGACTC
	Rv primer	TCAAGAACTTACCCAGATAAACTGAA
	T-DNA primer	ACGTCCGCAATGTGTTAT
genotyping <i>plt2</i>	Fw primer	AAGTGGCTGATTTCTTAGGAGTG
	Rv primer	GTAGAGGGACCCCAATATTTAAGTG
	T-DNA primer	ACGGTCGGGAACTAGCTCTAC
gene expression analysis	Fw <i>SAND</i>	AACTCTATGCAGCATTTGATCCACT
	Rv <i>SAND</i>	TGATTGCATATCTTTATCGCCATC
	Fw <i>AHL15</i>	AAGAGCAGCCGCTTCAACTA
	Rv <i>AHL15</i>	TGTTGAGCCATTTGATGACC
	Fw <i>AHL19</i>	CTCTAACGCGACTTACGAGAGATT
	Rv <i>AHL19</i>	ATATTATACACCGGAAGTCCTTGGT
	Fw <i>AHL20</i>	CAAGGCAGGTTTGAAATCTTATCT
	Rv <i>AHL20</i>	TAGCGTTAGAGAAAGTAGCAGCAA
	Fw <i>TAA1</i>	TTCGTGGTCAATCTGGATCATGG
	Rv <i>TAA1</i>	ACCACGTATCGTCACCGTACAC
	Fw <i>YUC3</i>	ATGGTCGTTTCGTAGCGCTGTTC
	Rv <i>YUC3</i>	GCGAGCCAAACGGGCATATACTTC
Fw <i>YUC8</i>	TGCGGTTGGGTTTACGAGGAAAG	
Rv <i>YUC8</i>	GCGATCTTAACCGGTCCATTG	
Fw <i>STY1</i>	TCGCATACCTTCTCATTGAGGGCT	
Rv <i>STY1</i>	CACCTAACACCGCCGATGAACT	

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.
- Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M.** (1998). A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO journal* **17**, 5484-5496.
- Bai, B., Su, Y.H., Yuan, J., and Zhang, X.S.** (2013). Induction of Somatic Embryos in Arabidopsis Requires Local YUCCA Expression Mediated by the Down-Regulation of Ethylene Biosynthesis. *Mol Plant* **6**, 1247-1260.
- Bailey, T.L., and Elkan, C.** (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology* (California: AAAI Press), pp. 28-36.
- Balasubramanian, S., and Schneitz, K.** (2002). NOZZLE links proximal-distal and adaxial-abaxial pattern formation during ovule development in Arabidopsis thaliana. *Development* **129**, 4291-4300.
- Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G., and Friml, J.** (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* **115**, 591-602.
- Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K., and Scheres, B.** (2005). The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. *Nature* **433**, 39-44.
- Boutillier, K., Offringa, R., Sharma, V.K., Kieft, 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. *The Plant cell* **14**, 1737-1749.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3468-3473.
- Burgos-Rivera, B., Ruzicka, D.R., Deal, R.B., McKinney, E.C., King-Reid, L., and Meagher, R.B.** (2008). ACTIN DEPOLYMERIZING FACTOR9 controls development and gene expression in Arabidopsis. *Plant molecular biology* **68**, 619-632.
- Cheng, Y., Dai, X., and Zhao, Y.** (2006). Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. *Genes & development* **20**, 1790-1799.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R.** (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant physiology* **139**, 5-17.
- Eklund, D.M., Staldal, V., Valsecchi, I., Cierlik, I., Eriksson, C., Hiratsu, K., Ohme-Takagi, M., Sundstrom, J.F., Thelander, M., Ezcurra, I., and Sundberg, E.** (2010). The Arabidopsis thaliana STYLISH1 Protein Acts as a Transcriptional Activator Regulating Auxin Biosynthesis. *The Plant cell* **22**, 349-363.
- 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.

- Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M., and McCourt, P.** (2004). The transcription factor FUSCA3 controls developmental timing in Arabidopsis through the hormones gibberellin and abscisic acid. *Developmental cell* **7**, 373-385.
- 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 physiology* **133**, 653-663.
- Heyndrickx, K.S., Van de Velde, J., Wang, C., Weigel, D., and Vandepoele, K.** (2014). A Functional and Evolutionary Perspective on Transcription Factor Binding in Arabidopsis thaliana. *The Plant cell*.
- Horstman, A., Willemsen, V., Boutilier, K., and Heidstra, R.** (2014). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends in plant science* **19**, 146-157.
- Jia, H.Y., McCarty, D.R., and Suzuki, M.** (2013). Distinct Roles of LAFL Network Genes in Promoting the Embryonic Seedling Fate in the Absence of VAL Repression. *Plant physiology* **163**, 1293-1305.
- Johnson, K.L., Degan, K.A., Ross Walker, J., and Ingram, G.C.** (2005). AtDEK1 is essential for specification of embryonic epidermal cell fate. *The Plant journal : for cell and molecular biology* **44**, 114-127.
- Junker, A., Monke, G., Rutten, T., Keilwagen, J., Seifert, M., Thi, T.M., Renou, J.P., Balzergue, S., Viehover, P., Hahnel, U., Ludwig-Muller, J., Altschmied, L., Conrad, U., Weisshaar, B., and Baumlein, H.** (2012). Elongation-related functions of LEAFY COTYLEDON1 during the development of Arabidopsis thaliana. *The Plant journal : for cell and molecular biology* **71**, 427-442.
- Kaufmann, K., Muino, J.M., Jauregui, R., Airoidi, C.A., Smaczniak, C., Krajewski, P., and Angenent, G.C.** (2009). Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the Arabidopsis flower. *PLoS biology* **7**, e1000090.
- Kaufmann, K., Wellmer, F., Muino, J.M., Ferrier, T., Wuest, S.E., Kumar, V., Serrano-Mislata, A., Madueno, F., Krajewski, P., Meyerowitz, E.M., Angenent, G.C., and Riechmann, J.L.** (2010). Orchestration of floral initiation by APETALA1. *Science* **328**, 85-89.
- Kim, J.I., Sharkhuu, A., Jin, J.B., Li, P., Jeong, J.C., Baek, D., Lee, S.Y., Blakeslee, J.J., Murphy, A.S., Bohnert, H.J., Hasegawa, P.M., Yun, D.J., and Bressan, R.A.** (2007). yucca6, a dominant mutation in Arabidopsis, affects auxin accumulation and auxin-related phenotypes. *Plant physiology* **145**, 722-735.
- Kooiker, M., Airoidi, C.A., Losa, A., Manzotti, P.S., Finzi, L., Kater, M.M., and Colombo, L.** (2005). BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic arabidopsis gene SEEDSTICK. *The Plant cell* **17**, 722-729.
- Krizek, B.** (2009). AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. *Plant physiology* **150**, 1916-1929.
- Krizek, B.A.** (2003). AINTEGUMENTA utilizes a mode of DNA recognition distinct from that used by proteins containing a single AP2 domain. *Nucleic acids research* **31**, 1859-1868.
- Krizek, B.A.** (2011a). Auxin regulation of Arabidopsis flower development involves members of the AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT) family. *Journal of experimental botany* **62**, 3311-3319.
- Krizek, B.A.** (2011b). Aintegumenta and Aintegumenta-Like6 regulate auxin-mediated flower development in Arabidopsis. *BMC Res Notes* **4**, 176.
- Lotan, T., Ohto, M., Yee, K.M., West, M.A.L., 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.

- Maere, S., Heymans, K., and Kuiper, M.** (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* **21**, 3448-3449.
- Mashiguchi, K., Tanaka, K., Sakai, T., Sugawara, S., Kawaide, H., Natsume, M., Hanada, A., Yaeno, T., Shirasu, K., Yao, H., McSteen, P., Zhao, Y.D., Hayashi, K., Kamiya, Y., and Kasahara, H.** (2011). The main auxin biosynthesis pathway in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 18512-18517.
- Matsuzaki, Y., Ogawa-Ohnishi, M., Mori, A., and Matsubayashi, Y.** (2010). Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. *Science* **329**, 1065-1067.
- Meister, R.J., Williams, L.A., Monfared, M.M., Gallagher, T.L., Kraft, E.A., Nelson, C.G., and Gasser, C.S.** (2004). Definition and interactions of a positive regulatory element of the Arabidopsis INNER NO OUTER promoter. *Plant Journal* **37**, 426-438.
- Mu, J., Tan, H., Hong, S., Liang, Y., and Zuo, J.** (2013). Arabidopsis transcription factor genes NF-YA1, 5, 6, and 9 play redundant roles in male gametogenesis, embryogenesis, and seed development. *Mol Plant* **6**, 188-201.
- Nicol, J.W., Helt, G.A., Blanchard, S.G., Jr., Raja, A., and Loraine, A.E.** (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**, 2730-2731.
- Nishimura, T., Hayashi, K., Suzuki, H., Gyohda, A., Takaoka, C., Sakaguchi, Y., Matsumoto, S., Kasahara, H., Sakai, T., Kato, J., Kamiya, Y., and Koshiba, T.** (2014). Yucasin is a potent inhibitor of YUCCA, a key enzyme in auxin biosynthesis. *Plant Journal* **77**, 352-366.
- Nole-Wilson, S., and Krizek, B.A.** (2000). DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic acids research* **28**, 4076-4082.
- Oh, E., Kang, H., Yamaguchi, S., Park, J., Lee, D., Kamiya, Y., and Choi, G.** (2009). Genome-wide analysis of genes targeted by PHYTOCHROME INTERACTING FACTOR 3-LIKE5 during seed germination in Arabidopsis. *The Plant cell* **21**, 403-419.
- Pajoro, A., Madrigal, P., Muino, J.M., Matus, J.T., Jin, J., Mecchia, M.A., Debernardi, J.M., Palatnik, J.F., Balazadeh, S., Arif, M., O'Maoileidigh, D.S., Wellmer, F., Krajewski, P., Riechmann, J.L., Angenent, G.C., and Kaufmann, K.** (2014). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome biology* **15**, R41.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M.W., 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 molecular biology* **68**, 225-237.
- Pinon, V., Prasad, K., Grigg, S.P., Sanchez-Perez, G.F., and Scheres, B.** (2013). Local auxin biosynthesis regulation by PLETHORA transcription factors controls phyllotaxis in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1107-1112.
- Prasad, K., Grigg, S.P., Barkoulas, M., Yadav, R.K., Sanchez-Perez, G.F., Pinon, V., Blilou, I., Hofhuis, H., Dhonukshe, P., Galinha, C., Mahonen, A.P., Muller, W.H., Raman, S., Verkleij, A.J., Snel, B., Reddy, G.V., Tsiantis, M., and Scheres, B.** (2011). Arabidopsis PLETHORA transcription factors control phyllotaxis. *Current biology : CB* **21**, 1123-1128.
- Riechmann, J.L., and Meyerowitz, E.M.** (1998). The AP2/EREBP family of plant transcription factors. *Biol Chem* **379**, 633-646.
- Schneider, T.D., and Stephens, R.M.** (1990). Sequence Logos - a New Way to Display Consensus Sequences. *Nucleic acids research* **18**, 6097-6100.

- Shani, Z., Dekel, M., Roiz, L., Horowitz, M., Kolosovski, N., Lapidot, S., Alkan, S., Koltai, H., Tsabary, G., Goren, R., and Shoseyov, O.** (2006). Expression of endo-1,4-beta-glucanase (cel1) in *Arabidopsis thaliana* is associated with plant growth, xylem development and cell wall thickening. *Plant cell reports* **25**, 1067-1074.
- Simonini, S., Roig-Villanova, I., Gregis, V., Colombo, B., Colombo, L., and Kater, M.M.** (2012). Basic pentacysteine proteins mediate MADS domain complex binding to the DNA for tissue-specific expression of target genes in *Arabidopsis*. *The Plant cell* **24**, 4163-4172.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11806-11811.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 3151-3156.
- Tang, X., Bian, S., Tang, M., Lu, Q., Li, S., Liu, X., Tian, G., Nguyen, V., Tsang, E.W., Wang, A., Rothstein, S.J., Chen, X., and Cui, Y.** (2012). MicroRNA-mediated repression of the seed maturation program during vegetative development in *Arabidopsis*. *PLoS genetics* **8**, e1003091.
- Wang, F., and Perry, S.E.** (2013). Identification of direct targets of FUSCA3, a key regulator of *Arabidopsis* seed development. *Plant physiology* **161**, 1251-1264.
- Winter, C.M., Austin, R.S., Blanvillain-Baufume, S., Reback, M.A., Monniaux, M., Wu, M.F., Sang, Y., Yamaguchi, A., Yamaguchi, N., Parker, J.E., Parcy, F., Jensen, S.T., Li, H., and Wagner, D.** (2011). LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Dev Cell* **20**, 430-443.
- Wojcikowska, B., Jaskola, K., Gasiorek, P., Meus, M., Nowak, K., and Gaj, M.D.** (2013). LEAFY COTYLEDON2 (LEC2) promotes embryogenic induction in somatic tissues of *Arabidopsis*, via YUCCA-mediated auxin biosynthesis. *Planta* **238**, 425-440.
- Won, C., Shen, X.L., Mashiguchi, K., Zheng, Z.Y., Dai, X.H., Cheng, Y.F., Kasahara, H., Kamiya, Y., Chory, J., and Zhao, Y.D.** (2011). Conversion of tryptophan to indole-3-acetic acid by TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS and YUCCAs in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 18518-18523.
- Woodward, C., Bemis, S.M., Hill, E.J., Sawa, S., Koshiba, T., and Torii, K.U.** (2005). Interaction of auxin and ERECTA in elaborating *Arabidopsis* inflorescence architecture revealed by the activation tagging of a new member of the YUCCA family putative flavin monooxygenases. *Plant physiology* **139**, 192-203.
- Wuest, S.E., O'Maoileidigh, D.S., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., Lohan, A.J., Loftus, B., Graciet, E., and Wellmer, F.** (2012). Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13452-13457.
- Yano, R., Kanno, Y., Jikumaru, Y., Nakabayashi, K., Kamiya, Y., and Nambara, E.** (2009). CHOTTO1, a putative double APETALA2 repeat transcription factor, is involved in abscisic acid-mediated repression of gibberellin biosynthesis during seed germination in *Arabidopsis*. *Plant physiology* **151**, 641-654.
- Yant, L., Mathieu, J., Dinh, T.T., Ott, F., Lanz, C., Wollmann, H., Chen, X., and Schmid, M.** (2010). Orchestration of the floral transition and floral development in *Arabidopsis* by the bifunctional transcription factor APETALA2. *The Plant cell* **22**, 2156-2170.

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. *The Plant cell* **21**, 2563-2577.

Chapter 6

Microarray-based identification of transcription factor target genes

Maartje Gorte^{1,*}, Anneke Horstman^{2,*}, Robert B. Page^{3,*},
Renze Heidstra¹, Arnold Stromberg⁴ and Kim Boutilier²

* These authors contributed equally to this work

¹ Molecular Genetics Group, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

² Plant Research International, P.O. Box 619, Bode 53, 6700 AP Wageningen, The Netherlands.

³ Department of Biology, Auburn University at Montgomery, P.O. Box 244023, Montgomery, AL
36124, U.S.A.

⁴ Department of Statistics, University of Kentucky, 817 Patterson Office Tower, Lexington, KY
40506-0027, U.S.A.

Plant Transcription Factors: Methods and Protocols. Yuan, L., Perry, S.E. - :
Humana Press, 2011 (Methods in Molecular Biology 754): 119 - 141.

Abstract

Microarray analysis is widely used to identify transcriptional changes associated with genetic perturbation or signalling events. Here we describe its application in the identification of plant transcription factor target genes with emphasis on the design of suitable DNA constructs for controlling TF activity, the experimental set-up, statistical analysis of the microarray data and validation of target genes.

1. Introduction

Elucidating the signal transduction cascades activated by transcription factors (TFs) is an essential step toward understanding TF function. Analysis of loss- and gain-of-function mutants with altered phenotypes often provide the first clues to a TF's function, however additional approaches are required to identify the specific gene expression cascades that lead to the observed phenotypes. One way to obtain insight into these signalling cascades is through transcriptional profiling. Transcription profiling can be applied to stable loss- and gain-of-function TF mutants to identify global expression changes that are associated with the mutant phenotype, thereby facilitating placement of the TF in a developmental pathway or process. Ultimately, one would also like to know the direct targets activated by TF binding to be able to distinguish between primary targets and secondary downstream signalling events. A number of techniques can be used to identify the primary targets of TF binding, including microarrays (Gregory and Belostotsky, 2009), high throughput transcriptome sequencing (Wang et al., 2002) and chromatin immunoprecipitation (ChIP; (Collas, 2010)), each with its associated advantages and disadvantages. While microarrays and mRNA-seq provide information about direct and indirect transcriptional changes that take place upon TF binding, ChIP identifies TF binding sites and does not provide any information about whether a target gene is expressed. Ultimately, a combined approach is required to identify gene expression changes and DNA binding sites (Kirmizis and Farnham, 2004). Here we focus on microarray-based transcriptional profiling as an initial method to identify TF target genes due to its high throughput nature, low cost, technical ease and comparatively well-established framework for data analysis.

The chapter that follows is not a detailed step-by-step protocol for identification of TF target genes using microarray technology, but rather a compendium of points that need to be addressed when setting up the experiments and analysing the data.

2. Materials

2.1. Construct Development

Measurement of the direct transcriptional response of a transcription factor (TF) requires control over its function in both time and space. Temporal regulation can be achieved by using a chemically-inducible form of the TF, while spatial regulation can be achieved by using cell- or tissue-specific promoters. Construct design is an essential first step in the identification of TF targets, as the choice of both the promoter and inducible system will influence the results and

should therefore be adapted to answer the research question. Here we discuss several one- and two-component chemical induction systems that can be used to identify TF target genes, as well as approaches to ensure sufficiently high expression and/or activity of the TF of interest (TOI) for microarray analysis.

2.1.1 One-Component Chemical Induction Systems

The frequently used one-component inducible expression systems make use of a TOI fused in frame to a tag that ensures chemical inducibility. These systems commonly use (parts of) animal nuclear receptors (NRs), which are TFs that are translocated to the nucleus upon binding of hormones or other molecules (reviewed in (Aranda and Pascual, 2001)). These NRs have a modular structure containing a DNA binding domain (DBD) and a ligand binding domain (LBD). In the absence of its ligand, the NR is bound by heat shock proteins (HSPs) in the cytosol. However, when the ligand is present, the NR dissociates from the HSP complex and moves to the nucleus, where it binds to DNA and regulates gene transcription. This mechanism has been exploited to create chemically-inducible LBD-TF chimeras in many different organisms.

A disadvantage of any one-component system is that fusion of additional protein sequence, such as an LBD, to a TF may negatively influence its activity or reduce stability. The functionality of the fusion protein must therefore be evaluated in advance, using N- and C-terminal fusions. A fusion protein can be considered functional if it either rescues a mutant phenotype or, in the absence of a mutant phenotype, generates the expected overexpression phenotype.

Glucocorticoid Receptor

The rat glucocorticoid receptor (GR) is an example of an NR that is widely used to create inducible TFs. The synthetic glucocorticoid dexamethasone (DEX) is commonly used as the inducing agent. The GR system has been used to identify targets of *AGAMOUS* and *LEAFY* during flower development (Wagner et al., 2004; Gomez-Mena et al., 2005), and *BABY BOOM* during somatic embryo development (Passarinho et al., 2008). In this approach, the TOI is fused to the GR LBD. In the absence of DEX, the TOI:GR fusion protein is sequestered by HSPs in the cytosol (Rohila et al., 2004), thereby preventing TOI activity. When DEX is added, the interaction with HSPs is disrupted, allowing the TOI to enter the nucleus and activate gene expression. DEX treatment does not appear to affect plant development or gene expression (Craft et al., 2005; Hanson et al., 2008).

The manner in which DEX is applied, the concentration of DEX used and the amount of time required for TF target gene expression after DEX induction, are all points that need to be considered when adjusting the system to a specific TOI. DEX can be applied by spraying or watering. When plants are watered with DEX, it is taken up via the roots and transported rapidly to the shoot (Aoyama and Chua, 1997). The spraying method requires the addition of soap to the medium to lower surface tension (Aoyama and Chua, 1997). Alternatively, when using seedlings, these can be grown *in vitro* on a nylon mesh on agar-containing Petri dishes, which facilitates the transfer of the seedlings to dishes containing induction medium supplemented with DEX (Levesque et al., 2006; Passarinho et al., 2008). The concentration of DEX that is used varies greatly between studies, but 5 to 20 μM DEX is commonly used to achieve full TF induction. One way to identify the optimal DEX concentration in your experimental system is to test the dose-response of the TOI-induced phenotype or complementation (Hay et al., 2003). The amount of time needed to induce nuclear localization and transcriptional activation by TFs has not been rigorously studied in plants, and must therefore be assessed on a case-by-case basis. In one study, an APETALA3 (AP3):GR fusion protein was used to identify AP3 target genes in inflorescences, with down regulation of its target *APETALA1* being observed within two hours of DEX-induction (Sundstrom et al., 2006). In another example, activation of a SHORT ROOT (SHR):GR fusion in the root led to up regulation of SHR target genes within six hours (Levesque et al., 2006). Nuclear localization can also be confirmed using antibodies to the TOI and/or the GR LBD.

Estradiol Receptor

A similar approach that is also widely used in plants employs the LBD of the human estradiol receptor (ER) fused to the TOI. In this system, the chimeric TF is translocated to the nucleus in response to estradiol (Zuo et al., 2000). The ER system has been used to identify the target genes of a number of plant TFs, including ENHANCER OF SHOOT REGENERATION 2 (ESR2; Ikeda et al., 2006) and ARR2 (Che et al., 2008).

As in the GR system, the manner in which estradiol is applied, the dosage and the induction period need to be considered. Estradiol is efficiently taken up by seedlings, but its accumulation appears to be slower (see (Zuo et al., 2000) below), as it is relatively immobile, i.e. it does not move from treated leaves to newly developing leaves, as is the case for DEX (Tornerio et al., 2002). Nevertheless, the system can be used if induction of the fusion protein gives the expected phenotype and generates a transcriptional response within several hours.

Estradiol treatment does not appear to affect plant development (Zuo et al., 2000), however it has been suggested that the ER-system is unsuitable for plants species, such as soybean, that produce large quantities of phytoestrogens as these can activate the ER LBD (Zuo and Chua, 2000). The estradiol concentration used to induce TF expression is in the range used for the GR-based system. In the study on ESR2 target genes, *ESR2:ER* was activated for 1 hour with 10 μ M estradiol before RNA was isolated from root explants for expression profiling (Ikeda et al., 2006).

2.1.2 Two-Component Chemical Induction Systems

A chemically inducible two-component system can be used to control the activity of a TF in cases where the tagged TOI is not functional. The first component in these systems is usually an artificial TF, consisting of an activation domain and a DNA binding domain, which is made inducible through the addition of an NR LBD. Expression of the artificial transcription factor can be placed under control of any suitable promoter. The second component in these systems is a promoter that is specifically bound by the artificial TF and that is placed upstream of the gene encoding the TOI. Upon chemical induction, the artificial TF binds to the promoter and activates transcription of the TOI. This system is therefore similar to the one-component system, but with one major difference: in one-component systems the TOI is present, but sequestered in an inactive form in the cytoplasm awaiting chemical induction, while in two-component systems the TOI is transcribed and translated only after chemical induction of, and activation by, the artificial TF. This fundamental difference has implications for the identification of direct target genes in combination with the translational inhibitor cycloheximide (described in **Section 3.1.1**). A variety of two-component systems exists, including ones that use alcohol, tetracycline or copper to control gene expression (Gatz and Lenk, 1998). Here we focus on NR-based systems, as only these have been extensively used to identify TF gene targets.

One example of an NR-based two-component induction system that has been used to identify TF target genes is the GVG system. The first component of this system is a chimeric TF (GVG), consisting of the DNA-binding domain of the yeast GAL4 TF, the herpes virus VP16 transactivation domain, and a GR LBD (Aoyama and Chua, 1997). Expression of this artificial TF can be controlled by any promoter of choice. The second component is the TOI, which is transcribed from a promoter containing multiple copies of the GAL4 upstream activating sequence (UAS). Induction by DEX results in translocation of GVG to the nucleus where it activates transcription of the TOI. The kinetics of this system were studied using a luciferase

(Luc)-based construct (Aoyama and Chua, 1997). *Luc* mRNA was first detected one hour after DEX-induction, and reached a maximum level after four hours.

The GVG system was used to identify targets of MYB46 and REPRESSOR OF *ga1-3* (RGA), which are involved in secondary wall biosynthesis and GA signalling, respectively (Zentella et al., 2007; Ko et al., 2009). MYB46 mRNA was detected 30 min after spraying with 10 μ M DEX and its targets were up regulated within 3-6 hours of induction. The same treatment induced RGA protein synthesis within one hour, while RGA target transcripts accumulated after 2-4 hours.

It has been reported that activation of GVG itself can cause developmental growth defects in Arabidopsis, rice and lotus (Kang et al., 1999; Ouwerkerk et al., 2001; Andersen et al., 2003). In Arabidopsis, high level GVG expression was correlated with severe phenotypes (Kang et al., 1999). This effect appears to be due to the presence of a GR domain in the GVG protein, as overexpression of the GAL4-VP16 (GV) protein lacking the GR domain does not cause growth defects (Weijers et al., 2003). For experimental purposes it may be sufficient to select lines that do not show growth defects, however such a selection process may itself influence the inferences that are drawn.

An additional drawback of the GVG system is that expression of the transgene may be variable from generation to generation (Weijers et al., 2003), or even lost over generations (Galweiler et al., 2000; Engineer et al., 2005). In tobacco and Arabidopsis, GVG-based expression could be restored by treatment with the cytosine-methylation inhibitor 5-azacytidine, suggesting that the loss of expression is caused by methylation of the UAS sequence.

The second example of an inducible two-component system is the XVE system. It is comparable to the GVG system, but uses a different artificial TF that is composed of the DNA-binding domain of the bacterial repressor LexA and the viral VP16 transactivation domain, and is fused to the ER LBD (Zuo et al., 2000). Once estradiol is added, XVE activates the second component, consisting of multiple LexA operator sequences fused to the 35S minimal promoter, which is placed upstream of the gene encoding the TOI.

Zuo et al. (Zuo et al., 2000) used a GFP reporter to evaluate the XVE system. GFP transcripts were first detected 30 minutes after estradiol treatment and reached a maximum after 24 hours, suggesting that the XVE system is slower than the GVG system. The minimal concentration of estradiol required to induce the system was 8 nM and the response was saturated at 5 μ M.

The XVE system was used to identify target genes of AGAMOUS-LIKE 24 (AGL24; (Liu et al., 2008)), ENHANCER OF SHOOT REGENERATION 1 (ESR1; (Matsuo et al., 2009)),

OCTADECANOIDRESPONSIVE ARABIDOPSIS AP2/ERF 59 (ORA59; (Pre et al., 2008)), and the NAC transcription factor ANAC092 (Balazadeh et al., 2010). *AGL24*, *ESR1* and *ANAC092* expression was induced using 10 μ M estradiol, whereas 2 μ M was used to induce *ORA59* expression. The induction period ranged from 5 hours for *ANAC092* and 24 hours for *ESR1*.

2.1.3 Enhancing Transcription Factor Activity

An important point to consider when designing the construct is whether the TOI will induce gene expression changes at a sufficiently high level to be detected on a microarray. There are two aspects to consider here: the expression level of the TOI and the transcriptional activation or repression strength of the TOI.

When the endogenous promoter is too weakly expressed to allow microarray analysis, the TOI can be expressed at a higher level by placing it under the control of a semi-constitutive promoter, such as the Cauliflower Mosaic Virus 35S promoter or the maize *ubiquitin* promoter (*ubi*). In fact, most studies make use of the 35S promoter, probably because its high expression level in a wide range of tissues provokes a strong transcriptional response. It is important to remember however, that a TF may display ectopic functions upon ubiquitous expression compared to the wild-type situation, probably leading to identification of a different set of target genes. The choice of promoter will depend therefore on the TF that is being studied and on the biological question. Another way in which the expression level of the TOI can be increased is to add a translation enhancer to the DNA construct, although again, care must be taken to avoid generating over-expression phenotypes. Commonly used translation enhancers include the 5' non-coding leader sequences of the tobacco mosaic virus genomic RNA, alfalfa mosaic virus RNA 4, and tobacco etch virus (reviewed in (Turner and Foster, 1995)). Translation enhancers can also be used to compensate for reduced protein stability caused by protein tags (such as the fusions described in **Section 2.1.1**).

The native promoter was used to identify SHR targets (Levesque et al., 2006). In this experiment, a *pSHR::SHR:GR* construct was used that also included a translational enhancer element from tobacco etch virus to counterbalance a possible decrease in protein function due to fusion of the GR domain. The experiment was performed in a *shr* mutant background to avoid generating over-expression phenotypes.

3. Methods

3.1 Microarray Experimental Set-Up

3.1.1 Direct versus Indirect Target Genes

Analysis of the gene regulatory networks downstream of TFs requires identification of direct targets. However, induction of a TF will result in the activation of many downstream genes, some of which will be TFs that will induce expression of secondary targets. Microarray based transcriptional profiling of TF-activated cells or tissues identifies gene expression changes, but, unlike ChIP, does not differentiate between direct and indirect targets. Two ways are commonly used to enrich for direct targets in microarray analyses: a time course analysis and application of the translation inhibitor cycloheximide.

Time Course Analysis

A time course analysis gives insight into gene expression patterns, allows researchers to study and model gene dynamics and regulatory networks, and enables researchers to correlate early and late genes with their possible function. Generally, sampling of earlier time points will yield a set of genes that is enriched in direct targets, while later time points will identify gene sets that are both direct and indirect targets. It is not only important to decide on the duration of the time course, but also on the developmental stage at which the analysis will take place. An additional consideration when complementing mutant lines is that the induction times for the TF:NR needs to take place within the time frame of full complementation or ectopic phenotype formation.

An example of an extensive time course microarray analysis is the study by Wagner *et al.*, (Wagner *et al.*, 2004), who treated *35S::LFY:GR* root explants with DEX for 0, 2, 4, 6, 8, 10, 12 and 22h. Sixty-one genes were differentially expressed, most of which were up regulated. There was no full overlap between the targets identified in this experiment and previous studies, which the authors contribute to the difference in LFY activity in the different target tissues that were used. While the experimental setup does not differentiate between direct and indirect targets, the use of a time course allowed the authors to make regulation profiles for the target genes, retrieve slow-responding known targets and identify novel quick-responding TFs as targets (Wagner *et al.*, 2004).

Another example is the study by Ko *et al.* on MYB46 (Ko *et al.*, 2009). This TF was previously shown to induce biosynthesis and deposition of cellulose, xylan and lignin (Zhong *et*

al., 2007), important components of secondary cell wall thickening. The authors studied DEX-induced MYB46 expression in 2 week old leaves at 0, 1, 3 and 6 hours after DEX application. Of the 282 genes that were more than 3-fold up regulated after 6 hours, 42 were TFs. Thirty-two of these TFs were already up regulated after 3 hours. Because the authors used a two-component system, they determined putative direct targets based on speed of regulation, and later confirmed *in vitro* binding of MYB46 to their target promoters using EMSA (Ko et al., 2009).

Cycloheximide

In plant biology, cycloheximide (CHX) is often used together with an inducible TF:NR fusion, to discriminate between direct and indirect target genes. CHX is a drug that inhibits protein synthesis by preventing the transfer of amino acids from aminoacyl-tRNA to nascent proteins (McKeehan and Hardesty, 1969). Transcriptional activation by the TF will occur in the presence of CHX, but subsequent translation of the primary target genes will be blocked, thereby preventing the activation of secondary, indirect, target genes. It is not possible to use cycloheximide with two-component systems that first require translation of the TOI (see **Section 2.1.2** above).

Two important issues to consider when using CHX are the concentration and the length of application. On one hand, failure to completely block protein synthesis will contaminate the candidate direct target gene set with secondarily-induced genes, while completely blocking protein synthesis for too long, while other cellular processes continue, will eventually cause cell death. In general, a concentration of 10 μM is used in liquid culture systems (e.g. (Wagner et al., 2004)). However Lee et al. (Lee et al., 2009) used concentrations up to 50 μM to treat plants grown on agar plates, because of the relatively small contact surface between the tissue and medium. CHX pre-incubation times of 30 minutes prior to DEX induction have been used (Ravni et al., 2006), as have induction times of up to 8 hours in the presence of DEX (Passarinho et al., 2008). Reversibility of the CHX treatment can be used to test for cell-lethal effects of CHX during the sampling period. Cell cultures or seedlings can be grown in the presence of CHX for the desired time and then transferred to CHX-free medium and assayed for recovery (Passarinho et al., 2008).

Even a non-toxic CHX dose will affect the transcriptional profile of a cell. We have observed that 10 μM CHX alters the mRNA levels of roughly one third of the *Arabidopsis* genome. Many short-lived proteins that repress gene expression, like the PS-IAA4/5 related *Arabidopsis* early-response Aux/IAAs, need to be constantly supplied through *de novo* protein

synthesis (Theologis et al., 1985; Abel et al., 1995). CHX is also known to stabilise certain short-lived mRNAs (Herrick et al., 1990). Nevertheless, even though CHX causes genome-wide effects on gene expression, our observations indicate that array data from CHX treatments cluster together. CHX treated samples are therefore different from non-CHX treated samples in a specific, non-random way that enables researchers to discriminate between genes regulated only by CHX and genes regulated by both CHX and the TOI. Given that CHX treatment induces significant changes in gene expression, it is essential to include control samples in which CHX is applied, but the TOI is not induced.

To summarize, when using an inducible TF:NR fusion in combination with CHX and time course experiments, the ideal situation would be to perform the fully replicated analysis, with and without CHX. This allows for separation of early and late direct targets, and early and late indirect targets, and therefore aids in generating hypotheses concerning downstream effects (Kiddle et al., 2010).

3.1.2 *Enrichment*

Identification of TF target genes becomes more difficult when the TOI is expressed in a limited number of cells (dilution problem) or when it is expressed in many different cell types (specificity problem). A number of methods exist for identification of TF targets in specific cell-types, provided the cells of interest can be isolated or marked. An excellent method for purifying fluorescently marked cells or tissues is fluorescence-activated cell sorting (FACS). FACS is a type of flow cytometry, where cells are sorted into two or more containers based on the amount and type of laser scattering and fluorescence (Herzenberg et al., 1976).

Plant lines that express a fluorescent protein (FP) can be used for sorting, but first require tissue dissociation into single cells (protoplasts) by enzymatic digestion of their cell walls (Bargmann and Birnbaum, 2009). The power of this approach is exemplified by the transcriptional profiling performed for individual tissues both in the Arabidopsis root and shoot (Birnbaum et al., 2003; Birnbaum et al., 2005; Yadav et al., 2009). The transcriptional identity of protoplasts appears rather stable, but can be monitored as a quality control (Sheen, 2001). In addition, because of the limited amount of cells, amplification of the RNA population is required before analysis on microarrays.

Cells can be marked either by a promoter::FP reporter or directly by the expression of the TOI translationally fused to FP. Plants carrying the TOI:FP fusion protein can be used both to sort the specific cells and to study the transcriptional consequences of expressing the protein in

these cells, while the corresponding promoter::FP reporter can be used as a control to sort wild-type cells. For example, Levesque et al. (Levesque et al., 2006) sorted the epidermis and lateral root cap cell types ectopically expressing the SHR:GFP fusion protein from the *WEREWOLF* (*WER*) promoter. A *WER::GFP* line was used to sort the same cell types as a control sample. This approach allowed the authors to focus specifically on the transcriptome of tissues that were previously shown to respond to ectopic SHR activation.

In a slightly different setup this method can be used to identify direct TF targets. A TF:NR fusion protein can be expressed in a specific tissue, while a promoter::FP transcriptional fusion can be used simultaneously to mark the cells of interest. Application of the NR inducer and CHX followed by cell sorting and microarray hybridisation of labelled transcripts will identify regulated genes when compared to the mock treated control sample. Although technically feasible, this particular approach has not yet been used to identify plant TF direct targets.

3.2 Data Validation

The result of a microarray data analysis is a long list of candidate target genes that are significantly up- or down regulated to different extents under the chosen experimental conditions. The next challenge is to verify the differential expression of the identified genes using an independent technique, a step referred to as validation. Here we discuss quantitative real-time RT-PCR (qRT-PCR) as a first step to validating the expression profiles of the significantly regulated genes observed in the microarray analysis. A brief description of CHIP-based techniques, which can be used after initial validation by qRT-PCR to verify TF binding to the promoter of the identified target genes, is also included.

3.2.1 Quantitative Real-Time RT-PCR

qPCR is a method of quantifying transcript abundances that is often used to confirm microarray results (Rajeevan et al., 2001; Chuaqui et al., 2002; Canales et al., 2006). The primary strengths of qRT-PCR are its unprecedented sensitivity and large dynamic range, which make it particularly well suited for quantifying low abundance transcripts and transcripts that vary widely in abundance between groups of interest (Bustin, 2002; Pfaffl, 2006; Karlen et al., 2007). However, despite these attractive features, there are a number of difficulties associated with generating high quality qRT-PCR data. A number of previous reviews have described in detail the issues surrounding generating sound qRT-PCR data including: ensuring that RNA is of a sufficient quality and purity (Bustin and Nolan, 2004; Bustin et al., 2009), the pros and cons of the various

approaches to generating cDNA via reverse transcription (Bustin and Nolan, 2004; Bustin et al., 2005), the importance of proper sample storage (Bustin et al., 2009), the need for careful primer design and assay validation (Bustin et al., 2009), and the advantages and drawbacks to different detection chemistries (Bustin and Nolan, 2004; Wong and Medrano, 2005).

With respect to biological materials, there are three general points to consider: the samples used to validate the microarray data, the endogenous reference gene (ERG) used to normalize the data (see also **Section 3.3.2** for a detailed discussion) and primer design. Strictly speaking, for validation purposes, qRT-PCR only needs to be performed on the same RNA samples that were used for the microarray analysis. Nonetheless, qRT-PCR analysis of candidate target genes using a new set of biological replicates is an important step that can be used to determine the generality of the results (Chuaqui et al., 2002). For greatest accuracy, the expression level of the ERG should be similar to that of the gene of interest (GOI), so that distortions that occur at different stages in the PCR are comparable (Bustin, 2002). The ERG should also be stably expressed in the tissues and conditions under study. So-called ‘housekeeping’ genes defined in the pre-genomic era, e.g. *ACTIN* and *UBIQUITIN*, were thought to be ubiquitous or invariant in their expression. However, recent studies have shown that the expression of classic plant ERGs is too variable to be used in qRT-PCR experiments, and as a result, more stable ERGs have been identified (Czechowski et al., 2005; Gutierrez et al., 2008). Finally, qRT-PCR primers should be designed to be gene-specific and their standard curves and efficiency should be determined. Mispriming, primer-dimer formation (Bustin, 2002) and amplification bias (Chuaqui et al., 2002) introduce inaccuracies in the reaction that will be amplified exponentially (Freeman et al., 1999).

Even when qRT-PCR experiments are robustly designed, the results may not corroborate the microarray data. Whereas the qualitative validation (i.e., up- or down regulation) usually correlates well, quantitative correlation (i.e., similar fold change estimates between the two methods) may be lacking. These differences are often due to inherent differences between the qRT-PCR and microarray methodologies. For example, candidate GOIs showing very low levels of expression are more difficult to validate with qRT-PCR (Freeman et al., 1999; Beckman et al., 2004). Nonetheless, many potentially interesting GOIs will show low fold changes. The researcher should consider whether low fold changes can be explained by other factors (e.g. restricted expression of the candidate gene in the tissue sample) and whether they can be handled appropriately in subsequent steps of experimentation.

3.2.2 Chromatin Immunoprecipitation

ChIP allows researchers to determine TF binding sites by pulling down only those DNA fragments that are bound by the TOI. In the plant research community, complementing or overexpressing transgenic lines (Leibfried et al., 2005), and antibodies against a protein tag, such as FP or GR (Schlereth et al.; Levesque et al., 2006) are commonly used, in contrast to a specific antibody against the TOI, which is the standard in other fields (Ye et al., 2010).

An important caveat of ChIP is that it only validates TF binding, not TF functionality - binding does not imply that the expression of the gene is regulated by this TF in this particular assay or time frame (Wang et al., 2002). A combined approach using microarrays and ChIP can be used to identify a TFs direct target genes with confidence (Kirmizis and Farnham, 2004). Within the context of microarrays, ChIP is used to validate whether a certain target gene is directly or indirectly regulated. ChIP can be combined with PCR of selected target genes, tiling microarrays or next-generation parallel sequencing to determine TF binding at a genome-wide level (Kaufmann et al., 2010). For details on the procedure, controls and considerations for ChIP, see Chapters 17 and 18.

3.3 Statistical Analysis

3.3.1 Microarray Data

In the plant sciences community, the most commonly used microarray platforms for TF target gene identification are single channel whole genome printed oligonucleotide arrays and these will be the focus of our discussion. However, most of what is said here also applies to two-channel spotted arrays carrying oligonucleotides, cDNAs or PCR fragments. This section summarizes part of Fu et al. (Fu et al., 2010).

In a non-microarray experiment, sample size is typically determined by estimating variability in the data, determining the difference you want to detect, and deciding on alpha, the probability of a false positive and beta, the probability of a false negative. In a microarray experiment, which is really many thousands of individual experiments all done at the same time, the variability in expression will be different for each gene, and a traditional power analysis would result in thousands of different sample sizes. Genes that have similar, but not identical, expression in two groups would require very large sample sizes to detect the minor difference, while genes with dramatic differences can be detected with very small sample sizes. That said, here are some general guidelines. One array per group does not typically allow any assessment of variability, and is therefore rarely useful. Two arrays per group are typically used for pilot data

to see which genes change the most. Although some studies have been published using two arrays per group, in our experience, this is rare. Using three arrays per group is generally the minimum for publication, and is reasonable for samples with similar genetic backgrounds where between subject variability is small. Studies where variability is larger benefit from at least five to ten subjects per group. Obviously, a larger sample size is always better.

The use of replicates in microarray experiments is under constant debate. Technical replicates use the same mRNA sample on multiple chips. They are useful for establishing the reliability of the platform, but they cannot be used to increase the sample size for statistical calculations. Biological replicates use different mRNA samples on each chip, and contribute to the overall statistical sample size for the experiment. In general, for professionally produced microarrays, technical replicates are not useful, as the reliability of the platform has already been well established.

A related issue is the use of pooling, which means putting more than one mRNA sample on each microarray. Pooling reduces individual variability, and thus increases power, but at the price of not being able to use individual covariates in the statistical model. When pooling, it is essential to extract RNA from every sample and then combine equal amounts of RNA from each sample to go on each chip. Pooling is discussed extensively by Peng et al. (Peng et al., 2003).

Once the RNA has been appropriately extracted, hybridized to chips and scanned, it is necessary to normalize the chips so that data between chips can be “fairly” compared. Although plenty of options are available for chip normalization, the most common are MAS 5 from Affymetrix, and gcRMA from Bioconductor (www.bioconductor.org). MAS 5 is the easiest to use and will be the focus of this discussion. gcRMA will typically yield similar results, unless gene expression is very low. In that case, gcRMA is likely better than MAS 5.

Using MAS 5 results in an output file for each chip that contains the probe set ID, the probe set expression, the presence/absence call and the presence/absence p-value. The presence/absence p-value is used to declare each probe set either “P” for present ($p < 0.05$), “A” for absent ($p > 0.065$), or “M” for marginal ($0.05 < p < 0.065$). The p-value cut-offs for each label are adjustable. Technically, the assumptions of independence that the statistical test makes are not satisfied, but the P/A call is still useful as we will see below.

The next step in the statistical analysis is data reduction. If there are probe sets that are not of interest, then statistical calculations should not be done on those probe sets. On most, if not all Affymetrix chips, the first approximately fifty probe sets are quality control probe sets used by the MAS 5 software. Typically, there is no need to do statistics on these probe sets.

Another group of probe sets that are typically removed are probe sets that are labeled as absent (A) on all the chips in the experiment. If the P/A call determines that the probe set is not expressed on any chip in the experiment, there is no reason to do statistical analysis on that probe set. At this stage, researchers should also identify any subsets of the probe sets that are of particular interest. These subgroups can be statistically analyzed separately as well as together with the rest of the probe sets.

The final step prior to statistical analysis of the data is to decide whether or not to take the log of the data. For most microarray data sets, the probe sets with larger expression levels benefit from a log transformation, but the smaller expression levels should not be logged. Most researchers choose to log their data, but many do not. Typically, for one color microarrays, there is not much difference in gene lists with or without logging the data.

The end result of the statistical analysis will be one or more p-values for each probe set. The overall p-value for each row tests whether there are any statistical differences between the rows. A histogram of these p-values provides useful information. If that histogram looks like that of a uniform distribution (a rectangle), then there may be little if any differences between the treatment groups. On the other hand, a histogram with a large peak for low p-values indicates that large differences exist between the treatment groups. Histograms with a low or moderate peak for small p-values indicate that more chips would likely result in smaller p-values for probe sets that are actually differentially expressed.

The next decision is how to determine the list of probe sets that have changed. Traditionally, a p-value less than 0.05 rejects the null hypothesis of no change. In a microarray experiment in which say, 10,000 tests are done, using a p-value cut off of 0.05 could mean as many as $0.05 \times 10,000 = 500$ false positives. The false discovery rate (FDR) of Benjamini and Hochberg (Benjamini and Hochberg, 1995) chooses the cut off by a user-specified expected proportion of false positives on this list. Ten or twenty percent are common choices. For experimental conditions that cause differential expression, but no large changes, the FDR method may not find any genes that change. As an alternative to using $p=0.05$ or the FDR method, many researchers simply use $p=0.01$ as the cut-off.

Once the overall p-values are used to determine the list of genes that changed, many researchers attempt to use cluster analysis to determine genes that are responding similarly to the experimental conditions. To avoid excess noise in the gene clusters, be sure to cluster only genes that are determined by statistical methods to be differentially expressed. Many different types of cluster analysis are possible, and they often yield results that are hard to interpret.

Statistical pattern matching (e.g. (Liu et al., 2005)) is an alternative that can be used to divide the list into sub-lists of genes that changed similarly. For example, if two sample t-tests were used to generate the overall p-values, then the list should be sorted into upregulated and downregulated genes. For more complicated experimental designs, the patterns will be more complicated. Consult with a statistician to determine appropriate patterns.

The biological interpretation of the resulting list(s) is done by first annotating the gene lists using the manufacturer's web site. There may be obvious biological conclusions that can be made at this point. A more statistical approach is to provide the list plus a larger list, say the entire chip, to a statistical software package that statistically determines gene ontology categories that are over represented on the smaller list compared to the larger list.

3.3.2 *Quantitative Real Time RT-PCR Data*

Compared to the technical issues discussed in **Section 3.2.1**, little effort has been spent reviewing the issues associated with the statistical analysis of qRT-PCR data. Nevertheless, it is now clear that applying objective statistical methods to qRT-PCR data poses special challenges (Skern et al., 2005; Pfaffl, 2006; Yuan et al., 2006; Karlen et al., 2007; Yuan et al., 2008), and over the past decade, a number of approaches to analyzing qRT-PCR data have been suggested in the literature. We briefly describe the nature of the numerical data generated via most relative expression qRT-PCR experiments and then discuss issues that researchers who seek to draw valid statistical inferences from such data are likely to face.

The Quantification Cycle: The Central Value of qRT-PCR

The strategy underlying qRT-PCR is to record the accumulation of fluorescent dyes that label a specific nucleic acid product throughout the course of a PCR. The amount of product yielded by a PCR approximates a logistic (S-shaped) curve when it is plotted as a function of the number of reaction cycles completed. Thus, setting a threshold within the exponential phase of the amplification curve and recording the number of fractional cycles required to eclipse this threshold provides a correlate to the initial amount of template, known as the quantification cycle (C_q ; lower C_q values correspond to more starting template). However, while C_q is the value of interest in the majority of qRT-PCR experiments, its determination requires exclusion of ground phase cycles (also known as the background or baseline) as well as determination of where along the y-axis, within the exponential phase, the threshold should be placed. Determination of the baseline and threshold is usually handled by proprietary software that

comes with real-time PCR hardware. Because this section focuses on how to analyze C_q values rather than how to ensure that they are valid, we do not discuss baseline and threshold determination further and refer interested readers to Bustin and Nolan (Bustin and Nolan, 2004) for discussions of when to adjust the baseline and threshold manually.

The Relative Expression Quantification Strategy

There are two general approaches to conducting qRT-PCR experiments. The first, known as absolute quantification, is based on calibration to a standard curve generated from a known external source (e.g., recombinant DNA) that enables one to express data in terms of transcripts per biological unit (e.g., copies/ μg of tissue). The second, known as relative quantification, seeks to describe expression in arbitrary units that are based on comparisons to a calibrator sample or a series of calibrator samples (e.g., RNA isolated from control or un-manipulated sources). Because the relative quantification approach makes fewer assumptions, is less labor intensive, and is sufficient for most applications (Pfaffl, 2006), it is the method most frequently used in basic research.

The traditional, and still broadly used, approach to relative expression qRT-PCR is to enter the relevant C_q values, or their averages, into one of a number of mathematical models that generate relative expression ratios (R_E) describing expression in non-calibrator samples in terms of fold change relative to calibrator samples (Pfaffl, 2006). Usually, R_E is normalized to one or more ERGs (Pfaffl, 2006) because, in principle, this approach enables one to correct for variations in the amount and/or quality of starting template that are introduced during upstream phases of the workflow. The simplest and most widely used approach to calculating R_E was proposed by Livak and Schmittgen (Livak and Schmittgen, 2001) and is known as the $2^{-\Delta\Delta C_t}$ method.

Although the $2^{-\Delta\Delta C_t}$ method is popular due to its simplicity and ease of calculation, it is only valid if a number of implicit assumptions are met. First and foremost, the expression level of the ERG must be invariant across the groups being considered. Second, there must be a doubling of the reaction product following every cycle. Finally, the reaction efficiencies (E ; varies between 1 and 2) must be equal among all reactions that go into the calculation of R_E . In practice, any combination of these assumptions can be violated with the end result being an inaccurate estimate of R_E , spurious statistical significance, or both (Pfaffl, 2006; Gutierrez et al., 2008; Yuan et al., 2008).

Addressing Assumptions that are Likely to be Violated

Perhaps the most critical assumption of all relative expression qRT-PCR analyses is that the ERG used for normalization is invariant across all of the treatments/categories being considered. The reason for this is that variation in the ERG will bias estimates of R_E and may give misleading results (Pfaffl, 2006; Gutierrez et al., 2008). As mentioned in **Section 3.2.1**, it is unlikely that any genes are universally suitable as ERGs across all tissues and research paradigms, and it is therefore important to verify that ERGs are invariant each time one wishes to investigate a new experimental system or tissue (Bustin et al., 2005; Wong and Medrano, 2005). However, ERG validation is often challenging and can become a circular problem (Vandesompele et al., 2002; Andersen et al., 2004).

Several software packages are available for assessing ERG stability. A widely used approach to normalization put forth by Vandesompele et al. (Vandesompele et al., 2002), and implemented in the geNorm software package, is to assume that candidate ERGs are not co-regulated and compute metrics of stability for each of these candidates based on pair-wise calculations between the candidates and a set of RNA samples corresponding to the groups to be compared. These metrics are in turn used to arrive at a subset of candidate ERGs from which a normalization factor based on the geometric mean of this subset is calculated. Another strategy put forth by Andersen et al. (Andersen et al., 2004) and implemented by the NormFinder software package uses a model-based approach to arrive at estimates of intragroup and intergroup variation in gene expression for candidate ERGs. In turn, these estimates are used to derive “stability values” (i.e., metrics) for each candidate that enable researchers to identify the candidate ERGs with the lowest intergroup and intragroup variation.

Another assumption of the $2^{-\Delta\Delta Ct}$ model that is likely to be violated is the assumption of 100% reaction efficiency (i.e., $E = 2$; (Yuan et al., 2008)). Implicit to this assumption is the additional assumption that the E values that go into the calculation of R_E are all equal. Thus, while situations in which E differs from two, but is more-or-less equal among the reactions used to calculate R_E will result in inaccurate estimates of R_E , they are not likely to result in erroneous inferences about differences between groups. However, cases in which the E values of the reactions used to calculate R_E are qualitatively different will lead to poor estimates of R_E and may lead to erroneous inferences about differences between groups. Models that incorporate the concept of E into the calculation of R_E have been put forward by Pfaffl (Pfaffl, 2001) and Hellemans et al. (Hellemans et al., 2007). However, it is important to note that while these models account for differences in E between the GOI and ERG (i.e., allow for gene specific

efficiencies), they assume that the efficiencies of the GOI and ERG(s) do not vary between calibrator and non-calibrator samples (i.e., they do not allow for group/treatment specific efficiencies). Nevertheless, it is advisable to investigate the validity of this assumption and methods for doing so have been presented by Burns et al. (Burns et al., 2005) and Yuan et al. (Yuan et al., 2006; Yuan et al., 2008).

The introduction of E into relative quantification models means that E must be empirically estimated. The most commonly used approach for doing this is to estimate the average E from a series of reactions that were setup using a variety of cDNA template concentrations (i.e., a dilution series; (Pfaffl, 2006; Yuan et al., 2006)). However, despite being the most commonly used method, estimating E from dilution series data has several drawbacks. Most obviously, the dilution series method requires considerable amounts of RNA and is laborious. Hence, for large experiments, it may not be feasible to estimate E for every sample and gene combination to be investigated. In addition, the dilution series approach does not estimate reaction specific efficiencies, but rather the average E across several reactions (see below). Finally, the dilution series method occasionally yields estimates of E that are > 2 , suggesting that it is prone to overestimating E (Pfaffl, 2006).

The second method for estimating E is to use the cycle by cycle fluorescence data that are collected during the course of a real-time PCR. This approach has the advantage of being able to yield an estimate of E for every reaction. Furthermore, unlike the dilution series approach, it does not require additional labor as fluorescence data are acquired during the course of conducting an experiment. A number of strategies for estimating E from fluorescence data have been suggested in the literature, however, the most straightforward approaches involve identifying the exponential phase of the amplification curve (Peirson et al., 2003; Ramakers et al., 2003), and regressing the resulting \log_{10} (Peirson et al., 2003; Ramakers et al., 2003) or \log_2 (Yuan et al., 2008) transformed subset of fluorescence values against cycle number.

While estimating E from fluorescence data has several advantages over the dilution series method (see above), there are also some drawbacks. An obvious concern is that for large experiments, involving thousands of reactions, using fluorescence based approaches creates a considerable informatics problem. Another concern that arises when using fluorescence based methods that rely on linear regression, is that estimates of E will be based on small sample sizes due to the exclusion of a large number of reaction cycles from the analysis. Further, it is not clear that using the estimates of E generated for every reaction in a dataset is the most appropriate

use of this information as analyzing data based on reaction-specific E values may introduce considerable noise into a dataset (Peirson et al., 2003; Cikos et al., 2007). Peirson et al. (Peirson et al., 2003) and Cikos et al. (Cikos et al., 2007) have suggested that analyses based on averaged efficiencies provide more robust results and that reaction-specific E values should be used primarily to exclude reactions that have outlying E s.

Statistical Inference

By far the most common way in which qRT-PCR data are analyzed is via the use of standard parametric statistical tests (i.e., t-test, ANOVA, etc.). As described in the previous sections, there are a number of situations that can render such analyses invalid. Nevertheless, if care is taken to ensure that the assumptions that are essential to calculating unbiased estimates of R_E are met, the application of objective statistical methods to relative expression qRT-PCR data is acceptable. However, meeting the assumptions of relative expression models is an altogether distinct issue from meeting the assumptions of conventional parametric statistical tests (e.g., constant and normally distributed error variance, independence of observations, symmetrically scaled continuous response variable, etc). One particularly egregious error that occurs far too often during the analysis of qRT-PCR experiments is the use of R_E values as a response variable in standard parametric tests. The reasons for this being inappropriate are simple. First, R_E is not symmetrically scaled, as up regulated values of R_E lie on one scale ($1 < R_E < \infty$) and down regulated values lie on another scale ($1 > R_E > 0$). Second, R_E is rarely if ever normally distributed and is unlikely to be described adequately by linear functions. For these reasons, it is necessary to logarithmically transform R_E before conducting standard parametric analyses. We suggest analyzing $\log_2(R_E)$ as transformation to this base results in a straightforward interpretation that has a long history of use in the microarray literature. It is important to note that analyzing $\Delta\Delta C_q$ values (as opposed to $2^{-\Delta\Delta C_q}$) is essentially the same as analyzing $\log_2(R_E)$. Because $\log_2(R_E)$ is symmetrically scaled and is usually approximately normally distributed, it has a far greater chance of meeting the assumptions of conventional parametric tests than R_E .

Summary of Data Processing and Error Propagation

It has been repeatedly demonstrated that the way in which qRT-PCR data are processed and analyzed can strongly influence the biological conclusions drawn from the data (Skern et al., 2005). Although a large number of processing procedures have been described in the literature,

there is currently no consensus on which methods are the most robust. The quality control and processing procedures that are essential to relative expression qRT-PCR experiments consist of several steps including: baseline determination, threshold determination, control gene validation, estimation of E , removal of reactions with outlying C_q (Burns et al., 2005) and/or E values, calculation of R_E or some other measure of expression, and a statistical analysis of these calculated expression measures. One of the most important issues associated with data processing is the issue of error propagation (Hellemans et al., 2007). Of particular concern within the context of the approaches to data analysis discussed in this section is that all of the components that go into the calculation of R_E such as estimates of E and C_q are themselves measured with uncertainty. Thus, there is a pressing need to develop processing procedures that account for this uncertainty, as well as user-friendly software implementations of these procedures that make them readily accessible to biologists.

Acknowledgements

M. G. was supported by a Netherlands Genomic Institute Horizon grant. A.H. was supported by a Technology Top Institute Green Genetics grant.

References

- Abel, S., Nguyen, M.D., and Theologis, A. (1995). The PS-IAA4/5-like family of early auxin-inducible mRNAs in *Arabidopsis thaliana*. *J Mol Biol* **251**, 533-549.
- Andersen, C.L., Jensen, J.L., and Orntoft, T.F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* **64**, 5245-5250.
- Andersen, S.U., Cvitanich, C., Hougaard, B.K., Roussis, A., Gronlund, M., Jensen, D.B., Frokjaer, L.A., and Jensen, E.O. (2003). The glucocorticoid-inducible GVG system causes severe growth defects in both root and shoot of the model legume *Lotus japonicus*. *Mol Plant Microbe Interact* **16**, 1069-1076.
- Aoyama, T., and Chua, N.H. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J* **11**, 605-612.
- Aranda, A., and Pascual, A. (2001). Nuclear hormone receptors and gene expression. *Physiol Rev* **81**, 1269-1304.
- Balazadeh, S., Siddiqui, H., Allu, A.D., Matallana-Ramirez, L.P., Caldana, C., Mehrnia, M., Zanor, M.I., Kohler, B., and Mueller-Roeber, B. (2010). A gene regulatory network controlled by the NAC transcription factor ANAC092/AtNAC2/ORE1 during salt-promoted senescence. *Plant Journal* **62**, 250-264.
- Bargmann, B.O., and Birnbaum, K.D. (2009). Fluorescence activated cell sorting of plant protoplasts. *J Vis Exp*.
- Beckman, K.B., Lee, K.Y., Golden, T., and Melov, S. (2004). Gene expression profiling in mitochondrial disease: assessment of microarray accuracy by high-throughput Q-PCR. *Mitochondrion* **4**, 453-470.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* **57**, 289-300.
- 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.
- Birnbaum, K., Jung, J.W., Wang, J.Y., Lambert, G.M., Hirst, J.A., Galbraith, D.W., and Benfey, P.N. (2005). Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. *Nat Methods* **2**, 615-619.
- Burns, M.J., Nixon, G.J., Foy, C.A., and Harris, N. (2005). Standardisation of data from real-time quantitative PCR methods - evaluation of outliers and comparison of calibration curves. *BMC Biotechnol* **5**, 31.
- Bustin, S.A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* **29**, 23-39.
- Bustin, S.A., and Nolan, T. (2004). Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* **15**, 155-166.
- Bustin, S.A., Benes, V., Nolan, T., and Pfaffl, M.W. (2005). Quantitative real-time RT-PCR--a perspective. *J Mol Endocrinol* **34**, 597-601.
- Bustin, S.A., Benes, V., Garson, J.A., Hellems, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., and Wittwer, C.T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* **55**, 611-622.
- Canales, R.D., Luo, Y., Willey, J.C., Austermiller, B., Barbacioru, C.C., Boysen, C., Hunkapiller, K., Jensen, R.V., Knight, C.R., Lee, K.Y., Ma, Y., Maqsoodi, B., Papallo, A., Peters, E.H.,

- Poulter, K., Ruppel, P.L., Samaha, R.R., Shi, L., Yang, W., Zhang, L., and Goodsaid, F.M.** (2006). Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat Biotechnol* **24**, 1115-1122.
- Che, P., Lall, S., and Howell, S.H.** (2008). Acquiring competence for shoot development in Arabidopsis: ARR2 directly targets A-type ARR genes that are differentially activated by CIM preincubation. *Plant Signal Behav* **3**, 99-101.
- Chuaqui, R.F., Bonner, R.F., Best, C.J., Gillespie, J.W., Flaig, M.J., Hewitt, S.M., Phillips, J.L., Krizman, D.B., Tangrea, M.A., Ahram, M., Linehan, W.M., Knezevic, V., and Emmert-Buck, M.R.** (2002). Post-analysis follow-up and validation of microarray experiments. *Nat Genet* **32 Suppl**, 509-514.
- Cikos, S., Bukovska, A., and Koppel, J.** (2007). Relative quantification of mRNA: comparison of methods currently used for real-time PCR data analysis. *BMC Mol Biol* **8**, 113.
- Collas, P.** (2010). The Current State of Chromatin Immunoprecipitation. *Molecular Biotechnology* **45**, 87-100.
- Craft, J., Samalova, M., Baroux, C., Townley, H., Martinez, A., Jepson, I., Tsiantis, M., and Moore, I.** (2005). New pOp/LhG4 vectors for stringent glucocorticoid-dependent transgene expression in Arabidopsis. *Plant Journal* **41**, 899-918.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R.** (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* **139**, 5-17.
- Engineer, C.B., Fitzsimmons, K.C., Schmuke, J.J., Dotson, S.B., and Kranz, R.G.** (2005). Development and evaluation of a Gal4-mediated LUC/GFP/GUS enhancer trap system in Arabidopsis. *BMC Plant Biol* **5**, 9.
- Freeman, W.M., Walker, S.J., and Vrana, K.E.** (1999). Quantitative RT-PCR: pitfalls and potential. *Biotechniques* **26**, 112-122, 124-115.
- Fu, W.J., Stromberg, A.J., Viele, K., Carroll, R.J., and Wu, G.** (2010). Statistics and bioinformatics in nutritional sciences: analysis of complex data in the era of systems biology. *J Nutr Biochem* **21**, 561-572.
- Galweiler, L., Conlan, R.S., Mader, P., Palme, K., and Moore, I.** (2000). Technical advance: the DNA-binding activity of gal4 is inhibited by methylation of the gal4 binding site in plant chromatin. *Plant J* **23**, 143-157.
- Gatz, C., and Lenk, I.** (1998). Promoters that respond to chemical inducers. *Trends in Plant Science* **3**, 352-358.
- Gomez-Mena, C., de Folter, S., Costa, M.M.R., Angenent, G.C., and Sablowski, R.** (2005). Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. *Development* **132**, 429-438.
- Gregory, B.D., and Belostotsky, D.A.** (2009). Whole-genome microarrays: applications and technical issues. *Methods Mol Biol* **553**, 39-56.
- Gutierrez, L., Mauriat, M., Guenin, S., Pelloux, J., Lefebvre, J.F., Louvet, R., Rusterucci, C., Moritz, T., Guerineau, F., Bellini, C., and Van Wuytswinkel, O.** (2008). The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol J* **6**, 609-618.
- Hanson, J., Hanssen, M., Wiese, A., Hendriks, M.M., and Smeekens, S.** (2008). The sucrose regulated transcription factor bZIP11 affects amino acid metabolism by regulating the expression of ASPARAGINE SYNTHETASE1 and PROLINE DEHYDROGENASE2. *Plant J* **53**, 935-949.

- Hay, A., Jackson, D., Ori, N., and Hake, S.** (2003). Analysis of the competence to respond to KNOTTED1 activity in Arabidopsis leaves using a steroid induction system. *Plant Physiol* **131**, 1671-1680.
- 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.
- Herrick, D., Parker, R., and Jacobson, A.** (1990). Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* **10**, 2269-2284.
- Herzenberg, L.A., Sweet, R.G., and Herzenberg, L.A.** (1976). Fluorescence-activated cell sorting. *Sci Am* **234**, 108-117.
- Ikeda, Y., Banno, H., Niu, Q.W., Howell, S.H., and Chua, N.H.** (2006). The ENHANCER OF SHOOT REGENERATION 2 gene in Arabidopsis regulates CUP-SHAPED COTYLEDON 1 at the transcriptional level and controls cotyledon development. *Plant Cell Physiol* **47**, 1443-1456.
- Kang, H.G., Fang, Y., and Singh, K.B.** (1999). A glucocorticoid-inducible transcription system causes severe growth defects in Arabidopsis and induces defense-related genes. *Plant J* **20**, 127-133.
- Karlen, Y., McNair, A., Perseguers, S., Mazza, C., and Mermod, N.** (2007). Statistical significance of quantitative PCR. *BMC Bioinformatics* **8**, 131.
- Kaufmann, K., Muino, J.M., Osteras, M., Farinelli, L., Krajewski, P., and Angenent, G.C.** (2010). Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nat Protoc* **5**, 457-472.
- Kiddle, S.J., Windram, O.P., McHattie, S., Mead, A., Beynon, J., Buchanan-Wollaston, V., Denby, K.J., and Mukherjee, S.** (2010). Temporal clustering by affinity propagation reveals transcriptional modules in Arabidopsis thaliana. *Bioinformatics* **26**, 355-362.
- Kirmizis, A., and Farnham, P.J.** (2004). Genomic approaches that aid in the identification of transcription factor target genes. *Exp Biol Med (Maywood)* **229**, 705-721.
- Ko, J.H., Kim, W.C., and Han, K.H.** (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *Plant J* **60**, 649-665.
- Lee, D.J., Park, J.W., Lee, H.W., and Kim, J.** (2009). Genome-wide analysis of the auxin-responsive transcriptome downstream of *iaa1* and its expression analysis reveal the diversity and complexity of auxin-regulated gene expression. *J Exp Bot* **60**, 3935-3957.
- Leibfried, A., To, J.P., Busch, W., Stehling, S., Kehle, A., Demar, M., Kieber, J.J., and Lohmann, J.U.** (2005). WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* **438**, 1172-1175.
- 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, C., Chen, H., Er, H.L., Soo, H.M., Kumar, P.P., Han, J.H., Liou, Y.C., and Yu, H.** (2008). Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. *Development* **135**, 1481-1491.
- Liu, H., Tarima, S., Borders, A.S., Getchell, T.V., Getchell, M.L., and Stromberg, A.J.** (2005). Quadratic regression analysis for gene discovery and pattern recognition for non-cyclic short time-course microarray experiments. *BMC Bioinformatics* **6**.
- Livak, K.J., and Schmittgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.

- Matsuo, N., Mase, H., Makino, M., Takahashi, H., and Banno, H.** (2009). Identification of ENHANCER OF SHOOT REGENERATION 1-upregulated genes during in vitro shoot regeneration. *Plant Biotechnology* **26**, 385-393.
- McKeehan, W., and Hardesty, B.** (1969). The mechanism of cycloheximide inhibition of protein synthesis in rabbit reticulocytes. *Biochem Biophys Res Commun* **36**, 625-630.
- Ouwerkerk, P.B., de Kam, R.J., Hoge, J.H., and Meijer, A.H.** (2001). Glucocorticoid-inducible gene expression in rice. *Planta* **213**, 370-378.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M.W., 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.
- Peirson, S.N., Butler, J.N., and Foster, R.G.** (2003). Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Res* **31**, e73.
- Peng, X., Wood, C.L., Blalock, E.M., Chen, K.C., Landfield, P.W., and Stromberg, A.J.** (2003). Statistical implications of pooling RNA samples for microarray experiments. *BMC Bioinformatics* **4**, 26.
- Pfaffl, M.A.** (2006). Relative quantification. In *Real-time PCR*, M.T. Dorak, ed (New York: Taylor and Francis), pp. 63-82.
- Pfaffl, M.W.** (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45.
- Pre, M., Atallah, M., Champion, A., De Vos, M., Pieterse, C.M.J., and Memelink, J.** (2008). The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. *Plant Physiology* **147**, 1347-1357.
- Rajeevan, M.S., Ranamukhaarachchi, D.G., Vernon, S.D., and Unger, E.R.** (2001). Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. *Methods* **25**, 443-451.
- Ramakers, C., Ruijter, J.M., Deprez, R.H., and Moorman, A.F.** (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* **339**, 62-66.
- Ravni, A., Eiden, L.E., Vaudry, H., Gonzalez, B.J., and Vaudry, D.** (2006). Cycloheximide treatment to identify components of the transitional transcriptome in PACAP-induced PC12 cell differentiation. *J Neurochem* **98**, 1229-1241.
- Rohila, J.S., Chen, M., Cerny, R., and Fromm, M.E.** (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant Journal* **38**, 172-181.
- Schlereth, A., Moller, B., Liu, W., Kientz, M., Flipse, J., Rademacher, E.H., Schmid, M., Jurgens, G., and Weijers, D.** MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. *Nature* **464**, 913-916.
- Sheen, J.** (2001). Signal transduction in maize and Arabidopsis mesophyll protoplasts. *Plant Physiol* **127**, 1466-1475.
- Skern, R., Frost, P., and Nilsen, F.** (2005). Relative transcript quantification by quantitative PCR: roughly right or precisely wrong? *BMC Mol Biol* **6**, 10.
- Sundstrom, J.F., Nakayama, N., Glimelius, K., and Irish, V.F.** (2006). Direct regulation of the floral homeotic APETALA1 gene by APETALA3 and PISTILLATA in Arabidopsis. *Plant J* **46**, 593-600.
- Theologis, A., Huynh, T.V., and Davis, R.W.** (1985). Rapid induction of specific mRNAs by auxin in pea epicotyl tissue. *J Mol Biol* **183**, 53-68.

- Tornero, P., Chao, R.A., Luthin, W.N., Goff, S.A., and Dangl, J.L.** (2002). Large-scale structure-function analysis of the Arabidopsis RPM1 disease resistance protein. *Plant Cell* **14**, 435-450.
- Turner, R., and Foster, G.D.** (1995). The Potential Exploitation of Plant Viral Translational Enhancers in Biotechnology for Increased Gene-Expression. *Molecular Biotechnology* **3**, 225-236.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F.** (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, RESEARCH0034.
- Wagner, D., Wellmer, F., Dilks, K., William, D., Smith, M.R., Kumar, P.P., Riechmann, J.L., Greenland, A.J., and Meyerowitz, E.M.** (2004). Floral induction in tissue culture: a system for the analysis of LEAFY-dependent gene regulation. *Plant J* **39**, 273-282.
- Wang, H., Tang, W., Zhu, C., and Perry, S.E.** (2002). A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. *Plant J* **32**, 831-843.
- 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.
- Wong, M.L., and Medrano, J.F.** (2005). Real-time PCR for mRNA quantitation. *Biotechniques* **39**, 75-85.
- Yadav, R.K., Girke, T., Pasala, S., Xie, M., and Reddy, G.V.** (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proc Natl Acad Sci U S A* **106**, 4941-4946.
- Ye, Q., Zhu, W., Li, L., Zhang, S., Yin, Y., Ma, H., and Wang, X.** (2010). Brassinosteroids control male fertility by regulating the expression of key genes involved in Arabidopsis anther and pollen development. *Proc Natl Acad Sci U S A* **107**, 6100-6105.
- Yuan, J.S., Wang, D., and Stewart, C.N., Jr.** (2008). Statistical methods for efficiency adjusted real-time PCR quantification. *Biotechnol J* **3**, 112-123.
- Yuan, J.S., Reed, A., Chen, F., and Stewart, C.N., Jr.** (2006). Statistical analysis of real-time PCR data. *BMC Bioinformatics* **7**, 85.
- Zentella, R., Zhang, Z.L., Park, M., Thomas, S.G., Endo, A., Murase, K., Fleet, C.M., Jikumaru, Y., Nambara, E., Kamiya, Y., and Sun, T.P.** (2007). Global analysis of della direct targets in early gibberellin signaling in Arabidopsis. *Plant Cell* **19**, 3037-3057.
- Zhong, R., Richardson, E.A., and Ye, Z.H.** (2007). The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell* **19**, 2776-2792.
- Zuo, J., and Chua, N.H.** (2000). Chemical-inducible systems for regulated expression of plant genes. *Curr Opin Biotechnol* **11**, 146-151.
- Zuo, J., Niu, Q.W., and Chua, N.H.** (2000). Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J* **24**, 265-273.

Chapter 7

A cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies

Anneke Horstman¹, Isabella Antonia Nougalli Tonaco¹,

Kim Boutilier¹, and Richard G.H. Immink^{1,2}

¹ Wageningen University and Research Centre, Bioscience, and ² Wageningen University,
Physiology of Flower Bulbs, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

Abstract

Since its introduction in plants 10 years ago, the bimolecular fluorescence complementation (BiFC) method, or split-YFP, has gained popularity within the plant biology field as a method to study protein-protein interactions. BiFC is based on the restoration of fluorescence after the two non-fluorescent halves of a fluorescent protein are brought together by a protein-protein interaction event. The major drawback of BiFC is that the fluorescent protein halves are prone to self-assembly independent of a protein-protein interaction event. To circumvent this problem, several modifications of the technique have been suggested, but these modifications have not lead to improvements in plants. Therefore, it remains crucial to include appropriate internal controls. Our literature survey among recent BiFC studies in plants shows that most studies use inappropriate controls, and a qualitative rather than quantitative read-out of fluorescence. Therefore, we provide a cautionary note and beginner's guideline for the setup of BiFC experiments, discussing each step of the protocol, including vector choice, plant expression systems, negative controls and signal detection. In addition, we present our experience with BiFC with respect to self-assembly, peptide linkers and incubation temperature. With this note, we aim to provide a guideline that will improve the quality of plant BiFC experiments.

1. Introduction

The vast majority of proteins encoded by a genome function in multi-protein complexes (Braun et al., 2013). Identifying these protein-protein interactions can provide insight into the functions of individual proteins, as well as the biological processes they control. A large variety of high-throughput technologies have been developed in the past 20 years to identify protein-protein interactions, including a toolbox of techniques to detect or confirm putative interactions *in vivo* under physiologically relevant conditions (Immink and Angenent, 2002; Braun et al., 2013). The bimolecular fluorescence complementation (BiFC) assay, also referred to as “split-fluorescent protein” technology (e.g. split-YFP), is one of the most popular and frequently used methods in the plant scientific field to study protein-protein interactions *in vivo* (reviewed in (Kodama and Hu, 2012)). BiFC is based on the *in vivo* reconstitution of fluorescence after two non-fluorescent halves of a fluorescent protein (FP) are brought together by a protein-protein interaction event (Figure 1). As such, BiFC not only provides information on whether two proteins interact, but can also be used to determine the cellular and sub-cellular site of a protein-protein interaction event. The possibility to split a FP into two halves and to use this for the detection of interactions between molecules was first described in 2000 for the GREEN FLUORESCENT PROTEIN (GFP) (Ghosh et al., 2000). Shortly thereafter, this method was used to detect *in vivo* protein-protein interactions in COS-1, NIH3T3, and HeLa cells (Hu et al., 2002) and later in plants (Bracha-Drori et al., 2004; Walter et al., 2004). The ease of implementation of the technology without the need for sophisticated equipment to detect the fluorescence signal has made BiFC a popular technology. In many cases, BiFC is the first method of choice for testing potential protein-protein interactions *in planta* and to confirm the outcomes of large-scale yeast-based or *in vitro* protein-protein interaction studies. The popularity of BiFC inspired researchers to optimize and modify the method to make it suitable for additional applications, including the development of multicolor BiFC for studying competition between interacting protein pairs or to simultaneously visualize multiple interactions in the same cell (Lee et al., 2008; Waadt et al., 2008; Kerppola, 2013), and BiFC-FRET (Fluorescence Resonance Energy Transfer) for the detection of higher-order protein complex formation (Shyu et al., 2008b, a). In the majority of plant studies, the split YFP-tagged proteins are overexpressed transiently or stably in isolated cells (protoplasts) or cell cultures; however, the BiFC method was recently used to study protein-protein interactions in intact plant tissues using native promoters to drive expression of the tagged proteins (Smaczniak et al., 2012).

Despite its widespread use, the BiFC method does have a number of shortcomings for the detection and visualization of protein-protein interactions. The major drawback of the system is the ability of the two FP halves to reassemble in the absence of a *bona fide* protein-protein interaction. This so-called ‘self-assembly’ of the FP halves can result in a high background signal, leading to a false-positive BiFC signal for a protein-protein interaction. To address this and other problems, a myriad of technical modifications have been implemented, including changing the split position in eYFP or YFP Venus from amino acid (AA)155 to AA173 or AA210 (Ohashi et al., 2012), introduction of point mutations to suppress self-assembly of the two FP halves (Kodama and Hu, 2010; Nakagawa et al., 2011) and the use of negative controls, including point-mutated versions of the proteins under study (Kerppola, 2006). Based on these observations, standard protocols have been developed (Kerppola, 2006; Fang and Spector, 2010; Kodama and Hu, 2012) and additional optimization steps have been proposed to generate a more reliable and robust assay. Unfortunately, none of the proposed changes to improve the robustness and reliability of the method appear to be generally applicable in plants (Kodama and Hu, 2012). Similarly, improvements developed for a mammalian expression system did not result in a more reliable read-out in *Xenopus* (Saka et al., 2008).

We performed an inventory of the recent literature in the plant BiFC field and conclude that despite the awareness of shortcomings in the BiFC technology, that the majority of researchers fail to include the correct internal controls and also incorrectly evaluate the results. Therefore, we present a guideline for BiFC use in plants, highlighting the most critical steps in the protocol and provide practical considerations for each individual step.

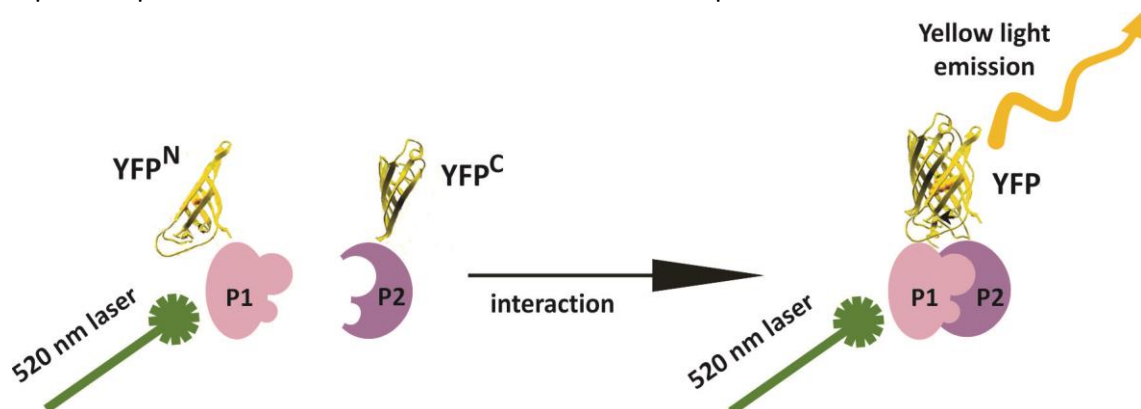


Figure 1. Schematic representation of the split-YFP/BiFC method based on the Yellow Fluorescent Protein (YFP). YFP is split into two non-fluorescent ‘halves’, an N-terminal part/half of the protein (YFP^N) and a C-terminal part/half of the protein (YFP^C), which are then fused to the proteins of interest (P1 and P2). The YFP molecule is reconstituted upon interaction between P1 and P2, resulting in yellow fluorescence when the molecules are excited with the correct wavelength.

2. Results and Discussion

2.1. Inventory of BiFC use in plant studies

Ten years ago, the first publications appeared showing the potential of the BiFC method for the detection and confirmation of protein-protein interactions in living plant cells (Bracha-Drori et al., 2004; Walter et al., 2004). Since then, a range of novel vectors and proposed improved protocols, mainly based on studies in mammalian cells, has been developed, with the goal to reduce the false-discovery rate and to improve the robustness of this technique. We performed a literature survey to determine which vectors, internal controls, and expression systems are used by the plant scientific community. A PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) search was performed in February 2014 using the terms 'BiFC' and 'plant' and the 100 most recent experimental papers were selected for analysis. From these studies, we extracted information about how the BiFC assay was performed (Supplementary Table 1). Analysis of this dataset revealed that the majority of recent BiFC studies were carried out using the original vectors or using home-made vectors with the split position in the YFP molecule around amino acid 155 (AA155) (Bracha-Drori et al., 2004; Walter et al., 2004). Our analysis revealed that the newly developed vectors for the mammalian field have not been implemented in the plant field. One of the reasons might be that for plant systems, these new vectors do not solve the problem of the high false-discovery rate, as discussed below in 2.2.1. Remarkably and more problematic, the majority of BiFC experiments were conducted without any or with inappropriate internal controls (Figure 2A, see discussion below in 2.2.3). Furthermore, we noticed that in more than 90% of the studies only qualitative measurements of fluorescence signal were performed (Figure 2B, see further discussion below in 2.2.5).

Based on this survey, we conclude that the BiFC method is generally not executed in the proper manner. We therefore provide a guideline that can be followed for the design of an optimal BiFC experiment in the following pages. This guideline is not meant to replace existing protocols (e.g. (Kerppola, 2006; Fang and Spector, 2010; Kodama and Hu, 2012)), but rather, to provide additional information and notes on critical points in the method based on published experiments and unpublished studies from our lab. This guideline will help the plant community to perform high quality BiFC studies, with an ensuing improvement in the quality of protein-protein interaction data.

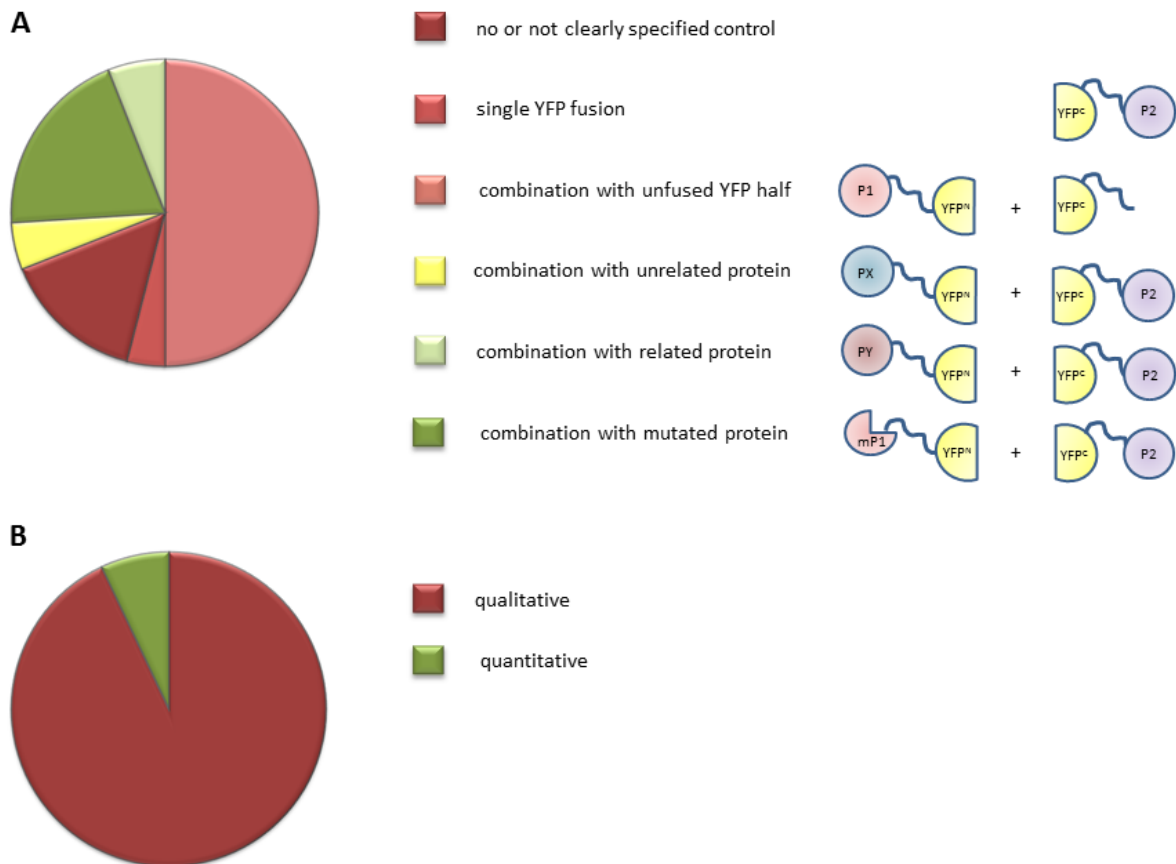


Figure 2. Survey of experimental set-up used in plant BiFC experiments. **(A)** Negative controls used in recent BiFC experiments. The suitability of a control is scaled using shading from dark red (the worst) to dark green (the best). The red classes indicate controls with a higher incidence of detecting false positives. The green classes represent suitable controls. The yellow class indicates a control of intermediate quality. Schematic representations of the controls are shown to the right. P1 and P2 represent the two proteins of interest, PX and PY indicate proteins that are related and unrelated, respectively to the protein of interest, and mP1 represents a mutant or truncated version of P1 **(B)** Percentage of BiFC experiments in which a qualitative (red) or quantitative (green) read-out of the fluorescence signal was measured.

2.2. Overview of the BiFC method

An overview of the BiFC method is presented in Figure 3. Prior to the start of a BiFC experiment, a number of choices have to be made, including the selection of vectors and negative and positive controls, and the expression system, each of which influences the outcome and quality of the experiment. In this section, we discuss the most important considerations for each individual step of the BiFC protocol.

2.2.1. Selection of vectors

When the BiFC method was introduced in plants in 2004, a small number of vector sets were available that all split eYFP in the loop between the seventh and eighth β -sheet (around AA155), and that expressed the fusion proteins from the strong constitutive Cauliflower Mosaic Virus 35S RNA promoter (CaMV35S) (Bracha-Drori et al., 2004; Walter et al., 2004). Unfortunately, this split position results in irreversible self-interaction capacity of the two non-fluorescent fragments, which may result in the detection of false positive protein-protein interactions if inappropriate controls are used. A few years later, new vectors were developed that split the YFP between β -sheet nine and ten (AA173), and the resulting YFP^N (1-173) fragment was combined with YFP^C (156-239) (Waadt et al., 2008).

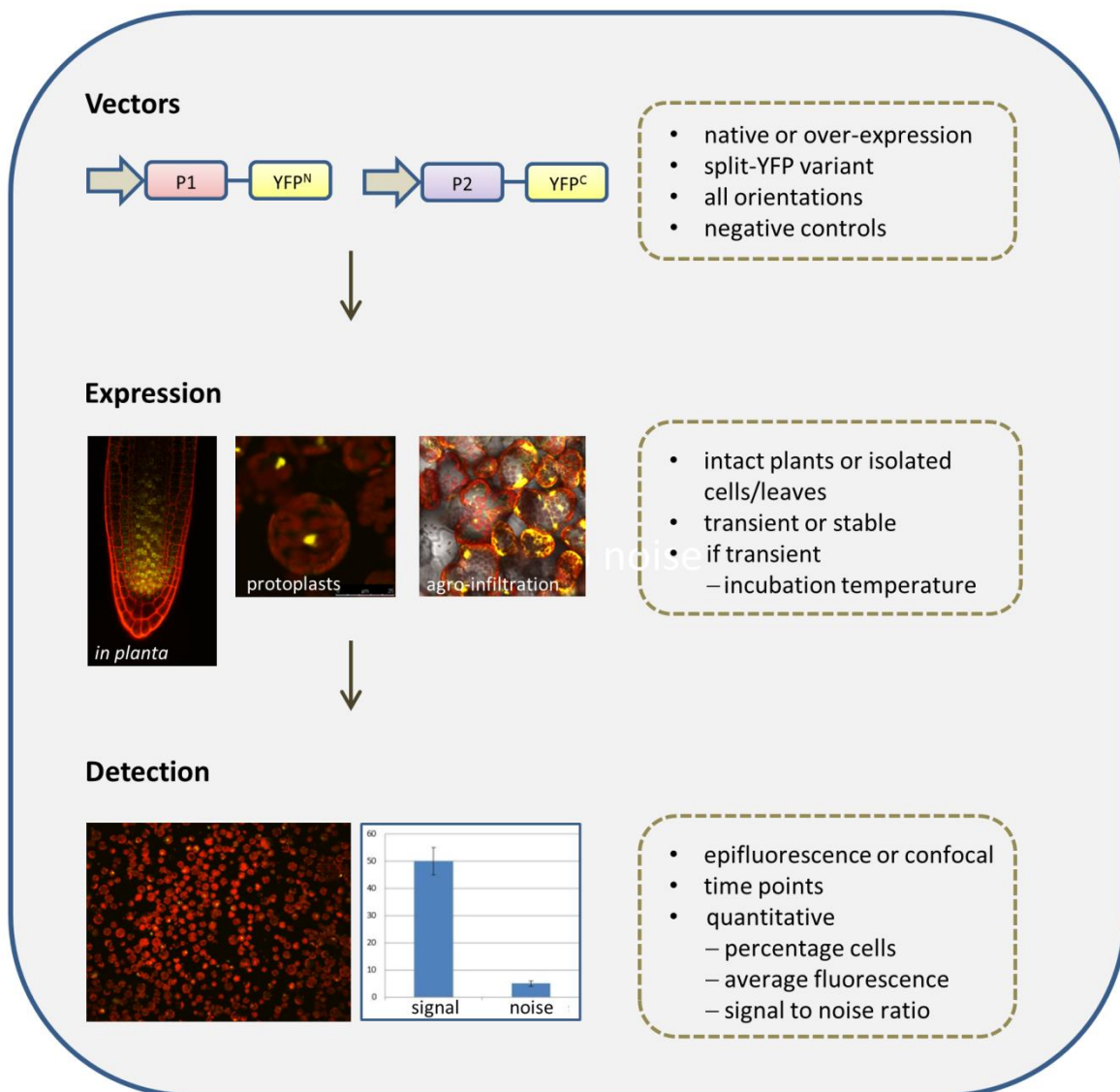


Figure 3. Flow diagram representing the steps and critical points in a BiFC experiment

This novel split position and combination of YFP fragments resulted in an increased signal for a tested positive control protein-protein interaction. However, the signal due to self-assembly also increased, which did not improve the signal (tested interaction)-to-noise (self-assembly) ratio (Waadt et al., 2008). Hence, this novel split position and combination of YFP fragments did not circumvent the problem of self-assembly and the accompanying high fluorescence signal. Aiming to minimize self-assembly, point mutations were introduced in the YFP halves, but with limited and varying success rates. Of all the reported mutations (Kodama and Hu, 2012), only the amino acid change I152L in the Venus YFP protein seems to give a consistent better signal-to-noise ratio when used in animal cells (Kodama and Hu, 2010). However, in plant cells this change results in a very weak fluorescence signal even with strongly interacting proteins (Lee et al., 2012).

It has also been proposed that the sequence and length of the peptide linker between the protein of interest and the YFP fragment could influence the complementation capacity of the split YFP fragments by affecting the flexibility and/or folding of the fused proteins, which in turn might be required for complex formation (Arai et al., 2001). We tested three different vector sets varying substantially in the sequence and length of the peptide linker (Table 1) and obtained similar fluorescence complementation signals for the interacting petunia MADS domain proteins FLOWERING BINDING PROTEIN2 (FBP2) and FBP11. Our results suggest that this specific protein-protein interaction is not influenced by the characteristics of the peptide linker, but we cannot exclude that the peptide linker is of importance for proper folding and detection of interactions of other fusion proteins. Nonetheless, the observation that peptide linkers in commonly used BiFC vectors vary substantially, but still allow BiFC (see Supplementary Table 1), suggests that the peptide linker sequence is not a critical factor for the success of a BiFC experiment.

In conclusion, our results and survey of the plant BiFC literature suggests that there is no evidence for the superiority of a particular BiFC vector set. Rather, it appears that reconstruction of the FP halves through protein-protein interaction depends more on the characteristics of the fused proteins than on the sequence of the YFP halves, the linker region, or the vector. Furthermore, BiFC efficiency differences have been observed between species, indicating that the cell type and the accompanying incubation conditions have a larger effect on BiFC than the vector itself. As discussed below, incorporating proper negative controls and experimental conditions seems to be of more importance for the success of a BiFC experiment.

BiFC vector set	promoter	protein of interest	peptide linker	Length (AA)	YFP part	terminator
1 (Welch et al., 2007)	CaMV35S	FBP2	Gateway	17	YFP ^N	NosT
		FBP11	Gateway	18	YFP ^C	
2		FBP2	RSIAT	15	YFP ^N	
		FBP11	KQKVMNH	17	YFP ^C	
3 (Walter et al., 2004)		FBP2	Myc-c tag	26	YFP ^N	
		FBP11	HA-tag	25	YFP ^C	

Table 1. Overview of BiFC constructs with different peptide linker sequences between the coding regions of the YFP halves and the coding regions of two interacting petunia MADS domain proteins. The indicated linker lengths include amino acids (AA) encoded by parts of the multiple cloning site. Vector combinations were tested upon transient transfection of petunia protoplasts. All expression cassettes were embedded in a *pUC* vector backbone. Peptide linker sequences: Gateway-based cloning linker sequence (Hartley et al., 2000); RSIAT/KQKVMNH (Hu et al., 2002); myc and HA tag (Walter et al., 2004).

2.2.2. Fusion orientations

One factor that influences the ability to detect protein-protein interactions in BiFC assays is the effect of the YFP fusion on the protein of interest. Protein-protein interactions are mediated by specific protein domains and fusing other (fluorescent) proteins to a protein of interest can interfere with the interaction capacity of these domains by steric hindrance or due to mis-folding (Busso et al., 2005). In addition, the three dimensional structure of a protein complex can also inhibit the reconstitution of the FP by spatial restrictions. It was shown by Bracha-Drori *et al.* that the fusion orientation can affect the amount of BiFC signal (Bracha-Drori et al., 2004). Therefore, to exclude false negative combinations, it is recommended to generate and test all of the eight combinations of constructs in which the N- and C-terminal fragments of the FP are fused to the N- and C-terminus of the proteins of interest (Figure 4A). The functionality of these fusion proteins can also be tested by genetic complementation, provided a mutant phenotype is available for the protein of interest. We believe that a single positive combination can provide sufficient proof of protein-protein interaction, as long as suitable negative controls are included and a correct experimental set-up is followed.

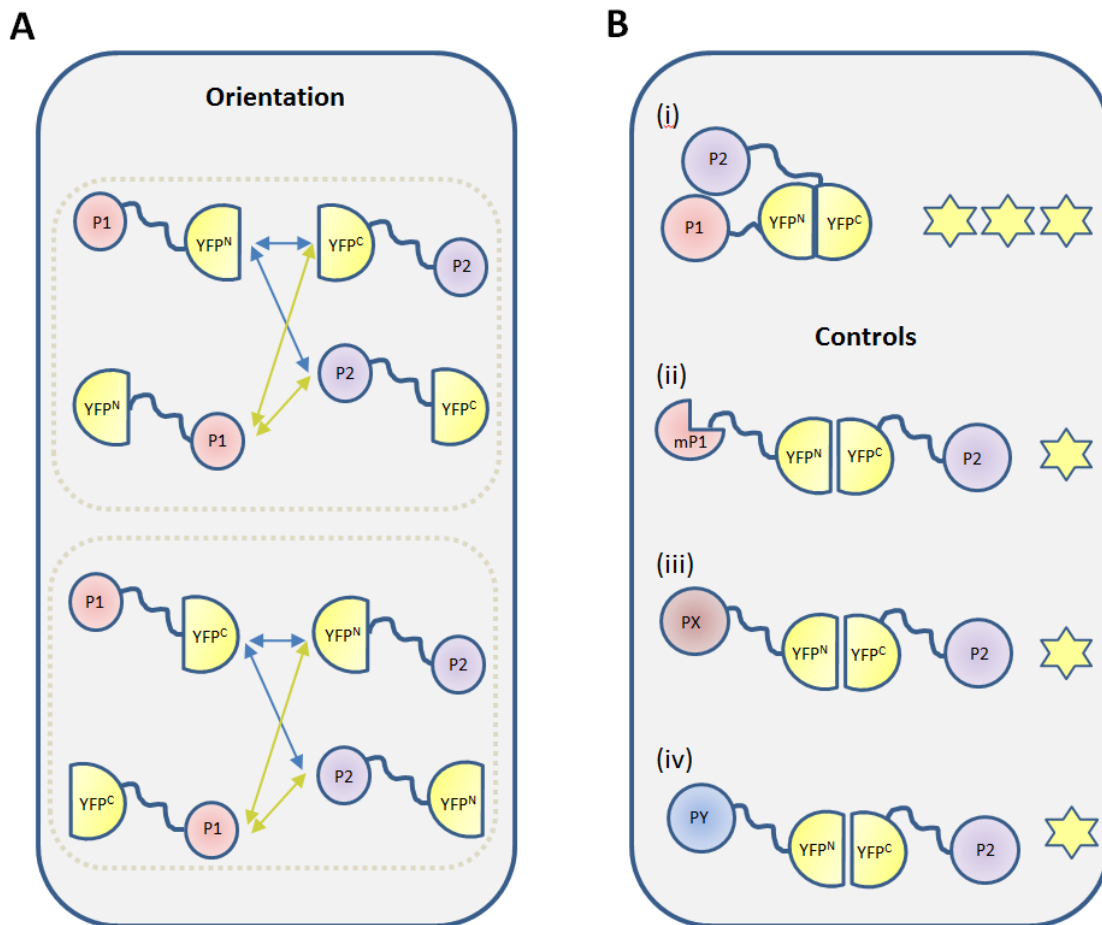


Figure 4. Two key elements of a BiFC experiment: fusion orientation and controls. **(A)** Each of the two YFP halves (YFP^N and YFP^C) can be fused at either its N- or C-terminus, with the protein of interest. Likewise, the protein of interest can also be fused to the YFP half via its N- or C-terminal end. This creates four possible YFP-protein combinations for each protein of interest and eight combinations that should be tested for each interaction pair. **(B)** In addition to the protein-protein interaction to be tested (i), appropriate negative controls should be incorporated in a BiFC experiment. These negative controls include substitution for one of the protein of interest by (ii) a mutated protein (mP1), (iii) a related protein (PX) that does not interact, and (iv) an unrelated protein (PY) with the same subcellular localization. The stars indicate YFP fluorescence due to self-assembly (one star) and expression due to a *bona fide* interaction between the test proteins (three stars).

2.2.3. Negative controls

The major disadvantage of BiFC as a method to detect protein-protein interactions is the signal that results from aspecific and irreversible interaction of the N- and C-terminal parts of the FP in the absence of interaction between the fused proteins of interest. For this reason, choosing a proper negative control is a critical step in the design a BiFC experiment. The ideal negative control in a BiFC experiment is a translational fusion between one half of the FP and a truncated or mutated version of the protein of interest that is unable to bind to its interaction partner (Figure 4B; e.g. (Jang et al., 2013; Park and Kim, 2013)). Development of this type of negative

control implies that interaction domains or amino acids have been identified using an *in vitro* approach such as Y2H, using bioinformatics predictions or genetic complementation, and that the changes to the protein do not negatively affect its stability or folding. In lieu of this, a fusion with a protein that is related to the protein of interest, but that does not interact would be a good alternative (Figure 4B; e.g. (Dong et al., 2010; Ren et al., 2013)). If a non-interacting protein family member is not known or available, one could opt for the use of an unrelated, non-interacting protein with the same cellular localisation (Figure 4B). We noticed that in a few cases in the literature, separate transfection with a plasmid comprising one protein of interest fused to YFP^N or YFP^C was used as a control (Supplementary Table 1). However, this control does not report the aspecific interactions that might occur between split-YFP fragments, because only one of the two YFP fragments is expressed. Empty vector controls comprising either half of the split YFP molecule, but lacking the protein of interest, were used most commonly in combination with the expression of a protein of interest fused to the complementary YFP fragment. Although it has been suggested that the expression levels of these unfused non-fluorescent fragments is higher than the expression of fusion proteins (Kodama and Hu, 2012), which would provide a conservative background level estimation, the subcellular localization of these split-YFP fragments might differ from that of the fusion between the split FP and protein of interest, thereby abolishing any potential for aspecific interaction and subsequent underestimation of the background. Note that the FP halves of negative controls should theoretically self-assemble, but may not do so due to interference by the fused test protein. It is therefore important to always use negative control that consistently exhibits a detectable fluorescence signal to be able to make a conservative estimate of the signal-to-noise ratio.

To determine the background fluorescence levels caused by self-assembly of AA155-based BiFC vectors (Welch et al., 2007), we generated fusion constructs between TagRED FLUORESCENT PROTEIN (TagRFP) and either the N- or the C-terminal YFP fragments (TagRFP-YFP^N and TagRFP-YFP^C). TagRFP is a monomeric fluorescent protein (Merzlyak et al., 2007) and its use enabled us to confirm the expression of the individual fusion proteins. TagRFP expression was observed as early as eight hours after co-transfection of Arabidopsis protoplasts with single TagRFP fusion constructs or TagRFP-YFP^C and YFP^N-BBM, a fusion of YFP^N with the transcription factor BABY BOOM (BBM; (Boutilier et al., 2002)). We noticed that in the double transfection, TagRFP fluorescence always coincided with weak YFP fluorescence. This fluorescence was not caused by bleed-through of TagRFP into the YFP channel, as no YFP signal was observed upon single transfections with TagRFP-YFP^C or TagRFP-YFP^N constructs alone (Figure 5).

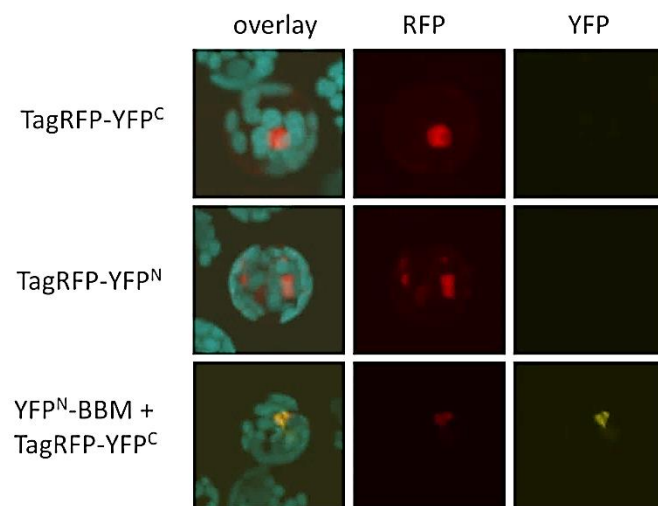


Figure 5. Self-assembly of YFP fragments. Confocal images of Arabidopsis protoplasts eight hours after transfection with either single TagRFP-YFP^C or -YFP^N plasmids, or with TagRFP-YFP^C and YFP^N-BBM plasmid combinations. The first column shows an overlay image of the background autofluorescence signal (blue), the TagRFP signal (red) and the YFP signal (yellow). The second and third columns show only the TagRFP and YFP signals, respectively. Co-transfection of plasmids containing YFP^N-BBM and TagRFP-YFP^C fragments results in a weak YFP signal, showing the re-constitution of the YFP molecule in the absence of a protein-protein interaction (self-assembly), and indicating the need for suitable negative controls and a quantitative read-out of the fluorescent signal in BiFC experiments.

Since there is no indication that BBM and TagRFP proteins interact, this YFP signal likely reflects that from YFP self-assembly. Although the YFP signal is weak, this experiment shows that BiFC results should be interpreted with caution. BiFC experiments cannot be performed without a proper negative control: transfection with a single plasmid is not sufficient. Additionally, a quantitative read-out should be used to distinguish between a fluorescence complementation by a true protein-protein interaction and signal due to self-assembly of the two FP halves, which can be scored based on the inclusion of a suitable negative control.

2.2.4. The BiFC assay: expression systems

BiFC experiments in plants are almost exclusively carried out using transient expression systems, usually *Agrobacterium* infiltration of tobacco leaf cells or PEG-mediated transfection of leaf protoplasts. Several protocols have been published for both methods (e.g. (Kerppola, 2006; Waadt and Kudla, 2008; Schutze et al., 2009; Boevink et al., 2014)), and neither is considered superior. Alternatively, the BiFC assay can be performed in whole plants (*in situ*) using stable transformants in which the promoters of the protein of interest are used to drive expression of the YFP-fusion (Smaczniak et al., 2012). Studying protein-protein interactions in their native context ensures that any additional proteins that are required for a protein-protein interaction will be co-expressed in the correct tissue context, and that quantitative relationship between the target protein and endogenous partner proteins is maintained. This approach can give insight

about when and where a specific protein-protein interaction first occurs, however, the irreversibility of the re-assembly of the YFP halves excludes the possibility to study protein-protein interaction dynamics. Since the genomic integration site of a transgene can affect its expression, it is important to determine whether the expression level and tissue specificity/localization of the fusion protein corresponds to that of the endogenous gene, and that expression of the transgene does not confer any mutant phenotypes. Testing of multiple independent transgenic lines is therefore required. BiFC analyses *in planta* using native protein expression levels in the native cellular environment is the most elegant approach, but these type of studies are still in their infancy and need to be more thoroughly analyzed to determine their robustness and reliability.

Another factor that needs to be considered in relation to the expression system is the culture temperature of the protoplasts or plant tissues prior to detection of fluorescence signal. While performing BiFC experiments by transient transfection of petunia protoplasts, we noticed a strong negative effect of the culture temperature on fluorescence complementation. Petunia protoplasts are routinely cultured at 28 °C (Denecke et al., 1989; Immink et al., 2002). At this temperature, no BiFC signal was obtained for the FBP2-FBP11 protein combination. In contrast, at least 30% of the protoplasts showed a fluorescence signal when the protoplasts were incubated at 23 °C. This difference is not due to mis-folding of the petunia MADS domain proteins at 28 °C, because strong fluorescence signals were obtained in protoplasts cultured at 28 °C when the FBP2 and FBP11 proteins were tagged with a complete eYFP fluorophore (Immink et al., 2002; Tonaco et al., 2006). The high-temperature sensitivity therefore seems to be specifically associated with the BiFC method. In support of this, a four hour pre-incubation of mammalian cells at lower temperatures prior to BiFC imaging significantly increases the fluorescence signal (Shyu et al., 2006).

Regardless of the expression method used, it is important to determine whether the proteins of interest and the control proteins are expressed. This not only provides information about the level of protein expression, which can greatly influence the results, but also indicates if the fusion between the protein of interest-split YFP is intact.

2.2.5. Detection methods

Fluorescence complementation in a BiFC experiment is usually detected using an epifluorescence microscope or a confocal laser scanning microscope (CLSM) (Supplementary Table 1). This qualitative analysis a BiFC experiment is problematic because of the self-assembly

capacity of the two FP halves. Consequently, simply showing images of fluorescent cells from the protein combination of interest and non-fluorescent cells from a control transfection is insufficient proof of protein-protein interaction. Rather, a quantitative comparison should be made between the signal obtained with the protein combination under study and the signal obtained with a proper negative control combination.

When using transient overexpression, it is important to realize that the irreversible nature of FP complementation leads to accumulation of the fluorescence signal in time, both for the tested interaction and the control experiment. The complementation of the FP by a protein-protein interaction and accumulation of fluorescence signal will proceed faster than the self-assembly in the control experiment due to the higher binding affinity of the interacting fusion proteins. Therefore, it is advisable to check fluorescence signals at different time points after transfection, as saturation of fluorescence signals influences the signal-to-noise ratio. Unfortunately, the vast majority of BiFC experiments is performed in a qualitative manner (Figure 2; Supplementary Table 1). One method for BiFC quantification is to determine the percentage of positive cells by counting for both the controls and the tested protein-protein interaction. Alternatively, the signal (tested interaction)-to-noise (self-assembly signal) ratio can be determined by measuring fluorescence intensities. Signal intensities can be determined from fluorescence images, but this approach is time-consuming, as it requires measurement of many cells one-by-one to determine average signal intensity. A faster way to analyse average fluorescence intensity in a population of cells at the same time is by fluorometry or by flow cytometry (Berendzen et al., 2012). Because the amount of expressed fusion proteins greatly influences the results, it is important to determine the expression levels of the different fusions within the population of cells by Western blotting. Subsequently, the BiFC signal intensity of the cell population can be normalized against the amount of fusion protein.

3. Conclusions

Split-YFP/BiFC is a widely used method for the detection and confirmation of protein-protein interactions in living plant cells. Nevertheless, the usefulness of this technology is overshadowed by self-assembly of the two halves of the FP, which results in the detection of fluorescent signal regardless of an interaction between the proteins of interest. Consequently, the introduction of control experiments is essential to obtain evidence for a potential protein-protein interaction event. However, a literature survey revealed that proper controls are missing in more than half of all analysed studies. For each protein-protein interaction analysed, a

negative control should be included as discussed in 2.2.3, and the fluorescent signals should be measured quantitatively. Currently, images are presented of plant cells or tissues with a fluorescent signal, without providing a thorough quantification of the signal-to-noise ratio between the fluorescence signals from the proteins of interest and the negative controls. It is important to realize that the goal of a BiFC experiment is to obtain strong support of a protein-protein interaction and not just to obtain an image of a fluorescent cell.

The lack of reliability and robustness of the split-YFP/BiFC technology due to self-assembly of the FP halves was recognized shortly after the introduction of the method, and a plethora of modifications were suggested to overcome this problem, as discussed in 2.2.1. Various improvements have been suggested based on splitting the FP at different positions and the introduction of point mutations in the FP sequence, but these improvements only appeared to overcome the self-assembly problem for the tested protein combinations, or only under specific conditions. Therefore, we conclude that none of the currently used BiFC vectors is superior and that all can be used as long as the right controls are included and quantitative measurements are applied. Alternatively, other methods can be used, based on complementation of other types of proteins, such as split-ubiquitin and split-luciferase (for review on these techniques, see (Morsy et al., 2008)). Self-assembly of protein halves is not an issue for these proteins; however, a drawback of using ubiquitin or luciferase is that no information can be extracted about the subcellular position of a protein-protein interaction event due to the nature of the read-out in these systems. In this respect, FRET (Fluorescence Resonance Energy Transfer)-based methods are more informative than protein complementation assays, because information is observed on both the localisation pattern of the individual proteins and the position at which these proteins interact (Immink et al., 2002; Bücherl et al., 2010), but these methods require sophisticated microspectroscopy equipment.

In conclusion, a low-tech, robust, and fully reliable system for the detection of protein-protein interactions in plant cells or tissues does not exist. Nonetheless, when implemented with caution split-YFP/BiFC remains a valuable tool for studying protein-protein interactions.

4. Experimental Section

4.1 Test of different peptide linker sequences.

The effect of using different peptide linker sequences between the halves of YFP and the protein of interest was examined in petunia protoplasts. The interaction between the petunia MADS domain proteins FBP2 and FBP11 (Immink et al., 2002) was used as positive control. All vectors used for this experiment were based on the *pUC* vector backbone and are described in Table 1. Isolation and transfection of petunia W115 leaf protoplasts was performed as described previously (Denecke et al., 1989; Immink et al., 2002). After transfection, protoplasts were incubated overnight in the dark at 28 °C (according to the original protocol) or at 23 °C.

4.2 Self-assembly capacity of YFP halves.

BBM (Boutillier et al., 2002) and *TagRFP* (Merzlyak et al., 2007) cDNA entry clones (pDONR207) were used to generate the *YFP^N-BBM*, *TagRFP-YFP^N*, *TagRFP-YFP^C* plasmids for Arabidopsis protoplast transfection. The plasmids were cloned using recombination into Gateway-compatible BiFC vectors (Welch et al., 2007). Arabidopsis protoplast isolation followed the procedures described in (Denecke and Vitale, 1995), except that leaves of three to four week-old Col-0 seedlings were used. Protoplast transfections were carried out as described in (Bucherl et al., 2010), but with a transfection time of 10 minutes. Fluorescence was viewed 8 hours after transfection by CLSM.

Author Contributions

A.H. and I.A.T.N performed the experiments. R.G.H.I performed the literature study. R.G.H.I and K.B. contributed to the experimental design, and A.H., K.B. and R.G.H.I. wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Sander van der Krol for providing an image of an *Agrobacterium* infiltrated tobacco leaf shown in Figure 3.

References and Notes

- Arai, R., Ueda, H., Kitayama, A., Kamiya, N., and Nagamune, T.** (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering* **14**, 529-532.
- Berendzen, K.W., Bohmer, M., Wallmeroth, N., Peter, S., Vesic, M., Zhou, Y., Tiesler, F.k.E., Schleifenbaum, F., and Harter, K.** (2012). Screening for in planta protein-protein interactions combining bimolecular fluorescence complementation with flow cytometry. *Plant Methods* **8**.
- Boevink, P., McLellan, H., Bukharova, T., Engelhardt, S., and Birch, P.** (2014). In Vivo Protein-Protein Interaction Studies with BiFC: Conditions, Cautions, and Caveats. *Methods in molecular biology* (Clifton, N.J.) **1127**, 81-90.
- Boutilier, 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.M., and van Lookeren Campagne, M.M.** (2002). Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. *The Plant Cell Online* **14**, 1737-1749.
- Bracha-Drori, K., Shichrur, K., Katz, A., Oliva, M., Angelovici, R., Yalovsky, S., and Ohad, N.** (2004). Detection of protein-protein interactions in plants using bimolecular fluorescence complementation. *The Plant Journal* **40**, 419-427.
- Braun, P., Aubourg, S., Van Leene, J., De Jaeger, G., and Lurin, C.** (2013). Plant Protein Interactomes. *Annual Review of Plant Biology* **64**, 161-187.
- Bucherl, C., Aker, J., de Vries, S., and Borst, J.W.** (2010). Probing protein-protein Interactions with FRET-FLIM. *Methods in molecular biology* (Clifton, N.J.) **655**, 389-399.
- Bücherl, C., Aker, J., Vries, S., and Borst, J.W.** (2010). Probing Protein-Protein Interactions with FRET-FLIM. In *Plant Developmental Biology* (Humana Press), pp. 389-399.
- Busso, D., Dleagoutte-Busso, B., and Moras, D.** (2005). Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. *Analytical Biochemistry* **343**, 313-321.
- Denecke, J., and Vitale, A.** (1995). The use of protoplasts to study protein synthesis and transport by the plant endomembrane system. *Methods in cell biology* **50**, 335-348.
- Denecke, J., Gossele, V., Botterman, J., and Cornelissen, M.** (1989). Quantitative analysis of transiently expressed genes in plant cells. *Methods in Molecular and Cellular Biology* **1**, 19-27.
- Dong, C.H., Jang, M., Scharein, B., Malach, A., Rivarola, M., Liesch, J., Groth, G., Hwang, I., and Chang, C.** (2010). Molecular association of the Arabidopsis ETR1 ethylene receptor and a regulator of ethylene signaling, RTE1. *The Journal of biological chemistry* **285**, 40706-40713.
- Fang, Y., and Spector, D.L.** (2010). BiFC Imaging Assay for Plant Protein-Protein Interactions. *Cold Spring Harbor Protocols* **2010**, pdb.prot5380.
- Ghosh, I., Hamilton, A.D., and Regan, L.** (2000). Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *Journal of the American Chemical Society* **122**, 5658-5659.
- Hartley, J.L., Temple, G.F., and Brasch, M.A.** (2000). DNA cloning using in vitro site-specific recombination. *Genome Res* **10**, 1788-1795.
- Hu, C.-D., Chinenov, Y., and Kerppola, T.K.** (2002). Visualization of Interactions among bZIP and Rel Family Proteins in Living Cells Using Bimolecular Fluorescence Complementation. *Molecular Cell* **9**, 789-798.

- Immink, R.G.H., and Angenent, G.C.** (2002). Transcription factors do it together: the hows and whys of studying protein–protein interactions. *Trends in Plant Science* **7**, 531-534.
- Immink, R.G.H., Gadella, T.W.J., Ferrario, S., Busscher, M., and Angenent, G.C.** (2002). Analysis of MADS box protein–protein interactions in living plant cells. *Proceedings of the National Academy of Sciences* **99**, 2416-2421.
- Jang, C., Seo, E.Y., Nam, J., Bae, H., Gim, Y.G., Kim, H.G., Cho, I.S., Lee, Z.W., Bauchan, G.R., Hammond, J., and Lim, H.S.** (2013). Insights into Alternanthera mosaic virus TGB3 Functions: Interactions with Nicotiana benthamiana PsbO Correlate with Chloroplast Vesiculation and Veinal Necrosis Caused by TGB3 Over-Expression. *Frontiers in plant science* **4**, 5.
- Kerppola, T.K.** (2006). Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nature Protocols* **1**, 1278-1286.
- Kerppola, T.K.** (2013). Multicolor Bimolecular Fluorescence Complementation (BiFC) Analysis of Protein Interactions with Alternative Partners. *Cold Spring Harbor Protocols* **2013**, pdb.top077164.
- Kodama, Y., and Hu, C.-D.** (2010). An improved bimolecular fluorescence complementation assay with a high signal-to-noise ratio. *BioTechniques* **49**, 793-805.
- Kodama, Y., and Hu, C.-D.** (2012). Bimolecular fluorescence complementation (BiFC): A 5-year update and future perspectives. *BioTechniques* **53**, 285-298.
- Lee, L.-Y., Fang, M.-J., Kuang, L.-Y., and Gelvin, S.** (2008). Vectors for multi-color bimolecular fluorescence complementation to investigate protein-protein interactions in living plant cells. *Plant Methods* **4**, 24.
- Lee, L.Y., Wu, F.H., Hsu, C.T., Shen, S.C., Yeh, H.Y., Liao, D.C., Fang, M.J., Liu, N.T., Yen, Y.C., Dokladal, L., Sykorova, E., Gelvin, S.B., and Lin, C.S.** (2012). Screening a cDNA library for protein-protein interactions directly in planta. *The Plant cell* **24**, 1746-1759.
- Merzlyak, E.M., Goedhart, J., Shcherbo, D., Bulina, M.E., Shcheglov, A.S., Fradkov, A.F., Gaintzeva, A., Lukyanov, K.A., Lukyanov, S., Gadella, T.W.J., and Chudakov, D.M.** (2007). Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nature Methods* **4**, 555-557.
- Morsy, M., Gouthu, S., Orchard, S., Thorneycroft, D., Harper, J.F., Mittler, R., and Cushman, J.C.** (2008). Charting plant interactomes: possibilities and challenges. *Trends in Plant Science* **13**, 183-191.
- Nakagawa, C., Inahata, K., Nishimura, S., and Sugimoto, K.** (2011). Improvement of a Venus-Based Bimolecular Fluorescence Complementation Assay to Visualize bFos-bJun Interaction in Living Cells. *Bioscience, Biotechnology, and Biochemistry* **75**, 1399-1401.
- Ohashi, K., Kiuchi, T., Shoji, K., Sampei, K., and Mizuno, K.** (2012). Visualization of cofilin-actin and Ras-Raf interactions by bimolecular fluorescence complementation assays using a new pair of split Venus fragments. *BioTechniques* **52**, 45-50.
- Park, M.-R., and Kim, K.-H.** (2013). Molecular characterization of the interaction between the N-terminal region of Potato virus X (PVX) coat protein (CP) and Nicotiana benthamiana PVX CP-interacting protein, NbPCIP1 **46**, 517-523.
- Ren, X.L., Qi, G.N., Feng, H.Q., Zhao, S., Zhao, S.S., Wang, Y., and Wu, W.H.** (2013). Calcineurin B-like protein CBL10 directly interacts with AKT1 and modulates K⁺ homeostasis in Arabidopsis. *The Plant journal : for cell and molecular biology* **74**, 258-266.
- Saka, Y., Hagemann, A.I., and Smith, J.C.** (2008). Visualizing protein interactions by bimolecular fluorescence complementation in *Xenopus*. *Methods* **45**, 192-195.

- Schutze, K., Harter, K., and Chaban, C.** (2009). Bimolecular fluorescence complementation (BiFC) to study protein-protein interactions in living plant cells. *Methods in molecular biology* (Clifton, N.J.) **479**, 189-202.
- Shyu, Y.J., Suarez, C.D., and Hu, C.-D.** (2008a). Visualization of AP-1–NF- κ B ternary complexes in living cells by using a BiFC-based FRET. *Proceedings of the National Academy of Sciences* **105**, 151-156.
- Shyu, Y.J., Suarez, C.D., and Hu, C.-D.** (2008b). Visualization of ternary complexes in living cells by using a BiFC-based FRET assay. *Nature Protocols* **3**, 1693-1702.
- Shyu, Y.J., Liu, H., Deng, X., and Hu, C.-D.** (2006). Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. *BioTechniques* **40**, 61-66.
- Smaczniak, C., Immink, R.G.H., Muiño, J.M., Blanvillain, R., Busscher, M., Busscher-Lange, J., Dinh, Q.D., Liu, S., Westphal, A.H., Boeren, S., Parcy, F., Xu, L., Carles, C.C., Angenent, G.C., and Kaufmann, K.** (2012). Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. *Proceedings of the National Academy of Sciences*.
- Tonaco, I.A.N., Borst, J.W., de Vries, S.C., Angenent, G.C., and Immink, R.G.H.** (2006). In vivo imaging of MADS-box transcription factor interactions. *Journal of Experimental Botany* **57**, 33-42.
- Waadt, R., and Kudla, J.** (2008). In *Planta Visualization of Protein Interactions Using Bimolecular Fluorescence Complementation (BiFC)*. CSH protocols **2008**, pdb prot4995.
- Waadt, R., Schmidt, L.K., Lohse, M., Hashimoto, K., Bock, R., and Kudla, J.** (2008). Multicolor bimolecular fluorescence complementation reveals simultaneous formation of alternative CBL/CIPK complexes in planta. *The Plant Journal* **56**, 505-516.
- Walter, M., Chaban, C., Schütze, K., Batistic, O., Weckermann, K., Näke, C., Blazevic, D., Grefen, C., Schumacher, K., Oecking, C., Harter, K., and Kudla, J.** (2004). Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *The Plant Journal* **40**, 428-438.
- Welch, D., Hassan, H., Blilou, I., Immink, R., Heidstra, R., and Scheres, B.** (2007). Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORT-ROOT action. *Genes & Development* **21**, 2196-2204.



Chapter 8

General discussion

Somatic embryogenesis (SE) is an intriguing illustration of plant cell totipotency. Application of stress, exogenous growth regulators or ectopic overexpression of certain transcription factors can induce SE from a variety of plant tissues (Chapter 1). The AIL transcription factor BBM triggers somatic embryo formation in *Arabidopsis*, tobacco, sweet pepper and white poplar (Boutilier et al., 2002; Srinivasan et al., 2007; Deng et al., 2009; Heidmann et al., 2011). Overexpression of PLT5/AIL5 also induces SE (Tsuwamoto et al., 2010), while this and other AILs are reported to affect organ size, or to induce other types of regeneration (Krizek, 1999; Krizek and Eaddy, 2012). Several other genes, including the *LEC* genes, are also able to induce SE when overexpressed (Lotan et al., 1998; Stone et al., 2001), but even though these phenotypes resemble the BBM overexpression phenotype, no genetic or molecular intersection between these SE pathways has been described. Direct BBM target genes were identified by microarray analysis after BBM-GR activation, but again, no clear link with known SE regulators was identified (Passarinho et al., 2008).

In this thesis, I examined the requirements for AIL-induced SE, identified BBM-interacting proteins and direct BBM target genes, and studied BBM's molecular genetic relation to other SE factors.

AIL overexpression phenotypes are dosage-dependent

The *Arabidopsis AIL* family consists of eight genes, with overlapping and distinct expression patterns within the growing tissues of the plant, where they are required to maintain meristematic activity. This renders them indispensable for many developmental processes, including embryogenesis, root, shoot and floral meristem maintenance, and organ development (Chapter 2). Molecular genetic studies identified AILs as components of many different gene regulatory networks, with conserved genetic modules found among these networks (Chapter 2).

We showed that all AIL proteins except ANT and AIL1 can induce SE, but that this process requires a relatively high AIL concentration (Chapter 4). We observed that an intermediate concentration induces organogenesis and a low concentration inhibits cellular differentiation. These results may explain why SE was not reported in other AIL overexpression studies. Firstly, the amount of overexpression conferred by the vector must be sufficiently high to induce embryogenesis. Secondly, sufficient transgenic lines need to be generated to observe the

phenotype, as only a small percentage of the obtained transgenics express sufficient AIL to induce SE (15-30%, *35S::BBM*; 5-10%, *35S::BBM-GR*) and finally, embryogenic transgenics are very small and slow growing, and may be overlooked.

It was previously suggested that *PLT2* regulates root meristem size and maintenance through a protein concentration gradient, with high, intermediate and low AIL concentrations instructing stem cell fate, cell division and differentiation, respectively (Galinha et al., 2007). Our overexpression data also support a dose-dependent AIL output in cotyledons and leaves, but suggest that a low AIL dose prevents rather than promotes differentiation (Chapter 4). The reported low AIL dose in the root might not actively instruct cellular differentiation, rather, it might simply be ineffective, thereby allowing cellular differentiation to take place. We combined our results with the reported AIL overexpression phenotypes to extend the model of dose-dependent AIL function (Figure 1). In this new model, the AIL concentration in the root that instructs cell division corresponds to an AIL overexpression dose that leads to enlarged organs, since the increased floral organ size in *ANT*, *PLT5/AIL5* and *AIL6/PLT3* overexpression seedlings is also due to increased cell division (Krizek, 1999; Nole-Wilson et al., 2005; Krizek and Eaddy, 2012). It was shown that higher levels of *AIL6/PLT3* (compared to those leading to enlarged organs) lead to smaller and undifferentiated sepal epidermal cells (Krizek and Eaddy, 2012), a phenotype that was also observed in *35S::BBM* flowers (Boutilier et al., 2002) and that relates to reduced leaf epidermal cell differentiation in *BBM/PLT2* overexpression lines presented in Chapter 4 (Figure 1). Finally, we have shown that increasing concentrations of AIL proteins in our system lead to organogenesis, and even higher levels to SE. Why organs and somatic embryos form at high AIL doses is not clear, but this phenotype might reflect the endogenous role of AIL

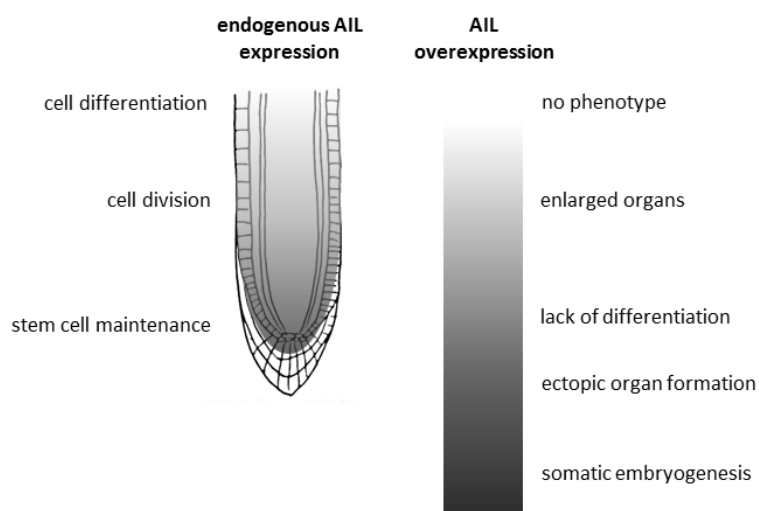


Figure 1. Extended model of dose-dependent AIL function

The model proposed by Galinha et al. (2007) to explain root meristem development (left) was combined with the AIL overexpression data presented in Chapter 4 and in other studies (right). The gradients represent the different cellular AIL concentrations from low (white) to high (dark grey).

proteins during organ and embryo development, i.e. the initiation and/or maintenance of organ/embryo identity. Transcriptome and AIL-DNA binding studies using different AIL dosages could reveal whether the different cellular outputs result from different expression levels of the same genes or from activation of specific target genes by a particular AIL dose.

ALLs induce two distinct modes of SE

In addition to being dependent on a high protein dosage, AIL-induced SE was also affected by the developmental stage of the seedling. We showed that BBM and PLT2 induce direct SE when activated in a time window surrounding seed germination, while post-germination activation leads to indirect SE from callus (Chapter 4). This callus almost exclusively formed on cotyledon regions above the vasculature, which is striking since organogenesis from above-ground tissues, which proceeds via a lateral root pathway, starts from pericycle-like cells around the vasculature (Che et al., 2007; Atta et al., 2009; Sugimoto et al., 2010). Although the AIL-induced callus did not appear to originate from such pericycle-like cells, but rather from the outer cotyledon cell layers, it would be interesting to determine whether the underlying embryogenic callus develops via a lateral root pathway.

The two different modes of SE were characterized by differential responses of BBM target genes at early and late time points; *LAFI* genes were only activated by early BBM activation (Chapter 4) and distinct auxin biosynthesis genes were activated by early and late BBM activation (Chapter 5). A previous study in which BBM target genes were identified by microarray analysis used BBM activation in four day-old seedlings (Passarinho et al., 2008) and therefore, unwittingly, identified BBM targets in the indirect SE pathway. In contrast, our CHIP-seq analysis in somatic embryo tissue and gene expression analysis in one day-old germinating seeds identified BBM targets that also play a role in the direct SE pathway. The *LAFI* genes are one such group of direct target genes. In germinating seeds and in seedlings, the *LAFI* genes are epigenetically repressed through deposition of H3K27me3 and through histone deacetylation by the chromatin remodeller PKL and the VAL transcriptional repressors (Dean Rider et al., 2003; Zhang et al., 2008; Zhang et al., 2012; Zhou et al., 2013). Therefore, these genes may not be accessible for BBM binding after germination, i.e. in the indirect pathway. Alternatively, BBM may bind to these loci after germination, but other cofactors required for *LAFI* gene activation

might be missing initially. These questions could be answered by BBM ChIP-seq experiments during and after germination.

AIL and HDG proteins interact and function antagonistically

Transcription factors function in protein complexes. To gain further insight into BBM function, we identified BBM interacting proteins (Chapter 3), among which the HDG1 and HDG11 transcription factors. We showed that these interactions extend beyond BBM and HDG1/HDG11 to other AIL and HDG protein-protein interactions. *HDG* genes are expressed in the L1 layer throughout the plant, where they function to specify epidermis identity and control development of its associated structures, such as trichomes, stomata or giant cells (Abe et al., 2003; Nakamura et al., 2006; Roeder et al., 2012; Peterson et al., 2013; Takada et al., 2013). We showed that ectopic overexpression of *HDG1* results in root and shoot meristem arrest, similar to *ail* loss-of-function phenotypes (Aida et al., 2004; Galinha et al., 2007; Mudunkothge and Krizek, 2012). We also observed other enhanced differentiation phenotypes in *HDG1* overexpression seedlings (ectopic formation of margin cells and higher ploidy levels) that were opposite to those observed in *BBM* overexpression seedlings (decreased cellular differentiation and reduced ploidy levels). Together, this suggested opposite functions of AIL and HDG transcription factors, with AILs stimulating cell proliferation, HDGs stimulating cell differentiation, and the ratio between the two determining the developmental outcome. In line with this antagonistic model, co-overexpression of BBM and HDG1 reduced the overexpression phenotypes of both proteins. Moreover, the phenotypes associated with decreasing HDG levels seem to correlate with those observed with increasing AIL levels (Chapter 4); *hdg11*, *atml1* and *hdg2* single mutants show decreased cellular differentiation (lack of giant cells, reduced stomatal differentiation) (Roeder et al., 2012; Peterson et al., 2013), similar to low AIL overexpression phenotypes; *pdf2;hdg3* and *atml1;hdg3* double mutants occasionally form ectopic shoots (Nakamura et al., 2006), similar to intermediate *AIL* overexpression phenotypes; and more extensive down-regulation of multiple *HDG* genes by amiRNAs or cosuppression led to the formation of somatic embryo-like tissue (Chapter 3), similar to high *AIL* overexpression phenotypes.

In wild-type *Arabidopsis*, the root and shoot stem cell niches are located close to the epidermis. The epidermis even produces a mobile microRNA that maintains the stem cell

identity of cells in its close vicinity, thereby serving as a stable spatial reference point (Knauer et al., 2013). Endogenous *AIL* and *HDG* expression patterns only show overlap in the epidermal cells close to the meristems and the AIL-HDG antagonism may function here to ensure that cell proliferation is limited to the meristem. The HDG proteins may maintain epidermal cell fate by counteracting AIL-driven cell proliferation in the epidermis. Notably, both *AIL* and *HDG* genes are also expressed in early embryos before the stem cell niches are established and before the protoderm is formed, thus their role in these early developmental stages requires further investigation.

How the AIL-HDG interaction works at the molecular level is currently unclear. One mechanism for proteins to exert antagonistic functions is through differential regulation of common target genes. For example, Bicoid target gene expression in *Drosophila* embryos is inhibited by the transcriptional repressor Capicua, which binds to Bicoid targets, and the ratio between the two factors determines whether a target gene is expressed or not (Lohr et al., 2009). We have identified several overlapping target genes, some of which were differentially regulated by HDG1 and BBM. However, differential regulation does not necessarily require physical interaction between interacting proteins with antagonistic functions. The limited target gene overlap between BBM and HDG1 suggests that differential target gene regulation is not the main mechanism underlying the AIL-HDG antagonism. Another mechanism for interacting proteins to exert their antagonistic functions is through competitive inhibition. In this scenario, interaction of HDG and AIL proteins inhibit their respective abilities to act as transcriptional regulators, with the balance between the amount of 'free' HDG or AIL determining the developmental outcome. This mode of action was shown for *Antirrhinum* RADIALIS, which binds and sequesters MYB-like transcription factors in the cytoplasm during flower development (Raimundo et al., 2013). However, cytoplasmic sequestration of AIL-HDG proteins seems unlikely, since both are localized in the nucleus. Alternatively, nuclear-localized transcriptional repressors can bind transcription factors and modulate their ability to bind to DNA, as was shown for SHORT VEGETATIVE PHASE and a splice form of FLOWERING LOCUS M (Pose et al., 2013), and for REVOLUTA and LITTLE ZIPPER proteins (Wenkel et al., 2007). Finally, it is possible that nuclear localization and DNA binding capacity are unaffected, but that the ability to activate or repress gene expression is reduced by binding of the antagonist. These options could be studied using EMSAs and reporter activity assays. In addition, it would be interesting to determine HDG1 binding sites by ChIP, in order to investigate the overlap between BBM and HDG1 target genes and how BBM and HDG1 dosage affects binding to common target genes.

BBM positively regulates other SE-inducing genes

We performed ChIP-seq experiments in somatic embryo tissue to identify genome-wide *in vivo* BBM binding sites (Chapter 5). This approach yielded many new BBM target genes, including some known players in somatic embryo induction, and allowed us to start unravelling the gene regulatory networks underlying BBM-induced SE (Figure 2). Until now, the relation between *BBM* and the *LEC* (*LAFL*) genes was unclear, even though their overexpression phenotypes are very similar. The FUS3 *LAFL* protein binds to the *BBM* locus, but transcriptional regulation was not investigated (Wang and Perry, 2013). Although FUS3 overexpression does not induce SE, it does confer embryonic (cotyledon) traits to leaves (Gazzarrini et al., 2004). We showed that activation of *LAFL* expression is a crucial downstream component of BBM-induced SE, as BBM-induced SE is abolished in the *lec1* and *fus3* mutants (Chapters 4). BBM also bound to *AGL15*, a downstream target of LEC2 that is partially responsible for its SE phenotype (Braybrook et al., 2006), but we did not observe a clear effect of BBM on *AGL15* expression at the time points analysed. Still, the BBM SE phenotype was reduced in the *agl15* mutant, indicating that it relies on this pathway. Whether BBM activates and requires *AGL15* expression later than the three hour time period that we examined, or *AGL15* activation is indirectly activated via LEC proteins needs to be examined.

At late embryo stages, *LEC* genes regulate seed maturation and it has been proposed that LEC2 overexpression promotes SE in vegetative tissues through dehydration stress resulting from the ectopic activation of seed maturation processes (Stone et al., 2008). However, *LEC* genes are also expressed during early embryogenesis and *lec* mutants occasionally form secondary embryos from suspensors at early embryo stages (Lotan et al., 1998). LECs may either directly suppress embryogenesis in the suspensor or they could be required for normal embryo proper development and thereby indirectly limit embryogenic suspensor growth, as was also shown for other embryogenesis mutants (Schwartz et al., 1994). This early LEC function is less well characterized, but could also underlie LEC- (and BBM-) induced SE. Transcriptional regulation of *LEC* genes by BBM/AIL in the wild-type seed is probably restricted to the early embryo stages, since this is when their expression patterns overlap, while at late embryo stages *AIL* expression is confined to the meristems, while *LECs* are expressed in the epidermis.

BBM bound to and stimulated expression of several *AHL* genes that are novel SE induction factors (Chapter 5; PhD thesis Omid Karami, unpublished). Again, downregulation of

these targets led to a reduction in BBM-mediated SE, suggesting that they are an important component of the BBM SE pathway. Interestingly, the *AHLs* appeared to be specifically targeted by BBM, as they were not identified as targets of LEC1, FUS3 or AGL15 (Chapter 5). By contrast, auxin biosynthesis genes (*TAA1* and *YUCs*) are activated by many of the somatic embryo-inducing proteins; they are targeted by the BBM, FUS3 (Wang and Perry, 2013) and LEC proteins (Braybrook et al., 2006; Stone et al., 2008; Junker et al., 2012; Wojcikowska et al., 2013), and *AHLs* also enhance *YUC* expression, although it is unclear whether they are direct *AHL* targets (data not shown, PhD thesis Omid Karami). Blocking *YUC* activity eliminated the formation of somatic embryos in BBM overexpression seedlings, and induced callus formation on the cotyledons. Similarly, SE induction by 2,4-D from immature zygotic embryos was also reduced in *yuc* mutants (Bai et al., 2013), suggesting that auxin biosynthesis is an important component of different SE systems. Future embryo marker analysis may reveal whether auxin biosynthesis is required to trigger cells to switch from vegetative to embryo cell fate, or whether it promotes the outgrowth of the embryos, similar to the auxin requirement for organ growth (Bohn-Courseau, 2010).

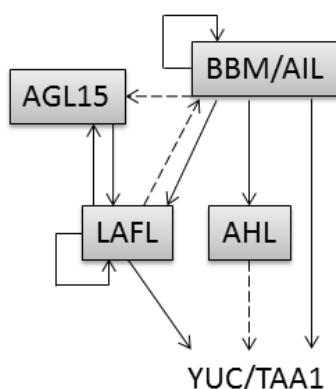


Figure 2: Transcriptional interactions between BBM/AIL and other SE factors
A schematic representation of the genetic interactions between SE-inducing genes (boxes). The solid arrows indicate DNA binding and transcriptional activation, while the dashed arrows indicate the cases where DNA binding was observed in the absence of transcriptional regulation (BBM → AGL15 and LAF1 (FUS3) → BBM/AIL) or where the regulation may be indirect (AHL → YUC/TAA1).

BOOMing business: improving BBM-based regeneration

The results presented in this thesis may help to improve BBM-based SE protocols in crop species. The dose-dependent *AIL* overexpression phenotypes that we have described for *Arabidopsis* might be useful for directing the formation of shoots or somatic embryos on demand. For example, explants might produce ectopic organs in a certain system, when somatic embryos are actually desired. In such a case, the *BBM* dose might be insufficient to induce SE. Finding ways to elevate *BBM* protein levels might lead to practical solutions for directing the type of regeneration. One way to enhance protein levels is through increased protein stability.

Previously, it was shown that AIL proteins can be stabilised by secreted RGF/GLV peptides (Matsuzaki et al., 2010) or by protein fusion to Histone 2B (Mahonen et al., 2014). In addition, more stable versions of AIL proteins might be obtained through mutagenesis. Alternatively, reducing the levels of AIL repressors, such as the HDG proteins, could elevate the effective AIL dose. Finally, the enhanced BBM response in ABA-insensitive (*abi*) mutants suggest a negative effect of ABA on the BBM phenotype, therefore it may be interesting to study whether blocking ABA synthesis, -transport or -signalling enhances SE.

References

- Abe, M., Katsumata, H., Komeda, Y., and Takahashi, T.** (2003). Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. *Development* **130**, 635-643.
- 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.
- Atta, R., Laurens, L., Boucheron-Dubuisson, E., Guivarc'h, A., Carnero, E., Giraudat-Pautot, V., Rech, P., and Chriqui, D.** (2009). Pluripotency of Arabidopsis xylem pericycle underlies shoot regeneration from root and hypocotyl explants grown in vitro. *Plant J* **57**, 626-644.
- Bai, B., Su, Y.H., Yuan, J., and Zhang, X.S.** (2013). Induction of Somatic Embryos in Arabidopsis Requires Local YUCCA Expression Mediated by the Down-Regulation of Ethylene Biosynthesis. *Mol Plant* **6**, 1247-1260.
- Bohn-Courseau, I.** (2010). Auxin: a major regulator of organogenesis. *Comptes rendus biologies* **333**, 290-296.
- Boutlier, K., Offringa, R., Sharma, V.K., Kieft, 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. *The Plant cell* **14**, 1737-1749.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3468-3473.
- Che, P., Lall, S., and Howell, S.H.** (2007). Developmental steps in acquiring competence for shoot development in Arabidopsis tissue culture. *Planta* **226**, 1183-1194.
- Dean Rider, S., Jr., 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.
- Deng, W., Luo, K.M., Li, Z.G., and Yang, Y.W.** (2009). A novel method for induction of plant regeneration via somatic embryogenesis. *Plant Sci* **177**, 43-48.
- 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.

- Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M., and McCourt, P.** (2004). The transcription factor FUSCA3 controls developmental timing in Arabidopsis through the hormones gibberellin and abscisic acid. *Dev Cell* **7**, 373-385.
- Heidmann, I., de Lange, B., Lambalk, J., Angenent, G.C., and Boutilier, K.** (2011). Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. *Plant cell reports* **30**, 1107-1115.
- Junker, A., Monke, G., Rutten, T., Keilwagen, J., Seifert, M., Thi, T.M., Renou, J.P., Balzergue, S., Viehover, P., Hahnel, U., Ludwig-Muller, J., Altschmied, L., Conrad, U., Weisshaar, B., and Baumlein, H.** (2012). Elongation-related functions of LEAFY COTYLEDON1 during the development of Arabidopsis thaliana. *Plant J* **71**, 427-442.
- Knauer, S., Holt, A.L., Rubio-Somoza, I., Tucker, E.J., Hinze, A., Pisch, M., Javelle, M., Timmermans, M.C., Tucker, M.R., and Laux, T.** (2013). A protodermal miR394 signal defines a region of stem cell competence in the Arabidopsis shoot meristem. *Dev Cell* **24**, 125-132.
- Krizek, B.A.** (1999). Ectopic expression of AINTEGUMENTA in Arabidopsis plants results in increased growth of floral organs. *Developmental genetics* **25**, 224-236.
- Krizek, B.A., and Eaddy, M.** (2012). AINTEGUMENTA-LIKE6 regulates cellular differentiation in flowers. *Plant molecular biology* **78**, 199-209.
- Lohr, U., Chung, H.R., Beller, M., and Jackle, H.** (2009). Antagonistic action of Bicoid and the repressor Capicua determines the spatial limits of Drosophila head gene expression domains. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 21695-21700.
- Lotan, T., Ohto, M., Yee, K.M., West, M.A.L., 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.
- Mahonen, A.P., ten Tusscher, K., Siligato, R., Smetana, O., Diaz-Trivino, S., Salojarvi, J., Wachsman, G., Prasad, K., Heidstra, R., and Scheres, B.** (2014). PLETHORA gradient formation mechanism separates auxin responses. *Nature* **515**, 125-129.
- Matsuzaki, Y., Ogawa-Ohnishi, M., Mori, A., and Matsubayashi, Y.** (2010). Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. *Science* **329**, 1065-1067.
- Mudunkothge, J.S., and Krizek, B.A.** (2012). Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function. *Plant J* **71**, 108-121.
- Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K.T., and Takahashi, T.** (2006). Characterization of the class IV homeodomain-leucine zipper gene family in Arabidopsis. *Plant physiology* **141**, 1363-1375.
- Nole-Wilson, S., Tranby, T.L., and Krizek, B.A.** (2005). AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant molecular biology* **57**, 613-628.
- Passarinho, P., Ketelaar, T., Xing, M., van Arkel, J., Maliepaard, C., Hendriks, M.W., 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 molecular biology* **68**, 225-237.
- Peterson, K.M., Shyu, C., Burr, C.A., Horst, R.J., Kanaoka, M.M., Omae, M., Sato, Y., and Torii, K.U.** (2013). Arabidopsis homeodomain-leucine zipper IV proteins promote stomatal development and ectopically induce stomata beyond the epidermis. *Development* **140**, 1924-1935.

- Pose, D., Verhage, L., Ott, F., Yant, L., Mathieu, J., Angenent, G.C., Immink, R.G., and Schmid, M. (2013). Temperature-dependent regulation of flowering by antagonistic FLM variants. *Nature* **503**, 414-417.
- Raimundo, J., Sobral, R., Bailey, P., Azevedo, H., Galego, L., Almeida, J., Coen, E., and Costa, M.M. (2013). A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in *Antirrhinum* flower asymmetry. *Plant J* **75**, 527-538.
- Roeder, A.H.K., Cunha, A., Ohno, C.K., and Meyerowitz, E.M. (2012). Cell cycle regulates cell type in the *Arabidopsis* sepal. *Development* **139**, 4416-4427.
- 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**, 3235-3245.
- Srinivasan, C., Liu, Z., Heidmann, I., Supena, E.D., Fukuoka, H., Joosen, R., Lambalk, J., Angenent, G., Scorza, R., Custers, J.B., and Boutilier, K. (2007). Heterologous expression of the BABY BOOM AP2/ERF transcription factor enhances the regeneration capacity of tobacco (*Nicotiana tabacum* L.). *Planta* **225**, 341-351.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11806-11811.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 3151-3156.
- Sugimoto, K., Jiao, Y.L., and Meyerowitz, E.M. (2010). *Arabidopsis* Regeneration from Multiple Tissues Occurs via a Root Development Pathway. *Dev Cell* **18**, 463-471.
- Takada, S., Takada, N., and Yoshida, A. (2013). ATML1 promotes epidermal cell differentiation in *Arabidopsis* shoots. *Development* **140**, 1919-1923.
- 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 molecular biology* **73**, 481-492.
- Wang, F., and Perry, S.E. (2013). Identification of direct targets of FUSCA3, a key regulator of *Arabidopsis* seed development. *Plant physiology* **161**, 1251-1264.
- Wenkel, S., Emery, J., Hou, B.H., Evans, M.M., and Barton, M.K. (2007). A feedback regulatory module formed by LITTLE ZIPPER and HD-ZIPIII genes. *The Plant cell* **19**, 3379-3390.
- Wojcikowska, B., Jaskola, K., Gasiorek, P., Meus, M., Nowak, K., and Gaj, M.D. (2013). LEAFY COTYLEDON2 (LEC2) promotes embryogenic induction in somatic tissues of *Arabidopsis*, via YUCCA-mediated auxin biosynthesis. *Planta* **238**, 425-440.
- Zhang, H., Bishop, B., Ringenberg, W., Muir, W.M., and Ogas, J. (2012). The CHD3 remodeler PICKLE associates with genes enriched for trimethylation of histone H3 lysine 27. *Plant physiology* **159**, 418-432.
- Zhang, H., Rider, S.D., Jr., Henderson, J.T., Fountain, M., Chuang, K., Kandachar, V., Simons, A., Edenberg, H.J., Romero-Severson, J., Muir, W.M., and Ogas, J. (2008). The CHD3 remodeler PICKLE promotes trimethylation of histone H3 lysine 27. *The Journal of biological chemistry* **283**, 22637-22648.
- Zhou, Y., Tan, B., Luo, M., Li, Y., Liu, C., Chen, C., Yu, C.W., Yang, S., Dong, S., Ruan, J., Yuan, L., Zhang, Z., Zhao, L., Li, C., Chen, H., Cui, Y., Wu, K., and Huang, S. (2013). HISTONE DEACETYLASE19 interacts with HSL1 and participates in the repression of seed maturation genes in *Arabidopsis* seedlings. *The Plant cell* **25**, 134-148.

A grayscale, close-up photograph of a white, multi-petaled flower, possibly a chrysanthemum, with the word "Summary" overlaid in the center. The petals are numerous and layered, creating a complex, textured appearance. The lighting is soft, highlighting the delicate structure of the flower against a dark, out-of-focus background.

Summary

Under appropriate tissue culture conditions, somatic plant cells can be induced to form embryos in a process called somatic embryogenesis (SE). SE provides a way to clonally propagate desirable plants and is therefore an important plant breeding tool. SE has also fascinated scientists for decades as an expression of plant ‘totipotency’, the ability to regenerate a whole new individual through embryogenesis. This thesis aims to obtain a deeper understanding of somatic embryo induction in *Arabidopsis* by the transcription factor BABY BOOM (BBM), through identification and functional analysis of BBM-binding proteins and BBM target genes.

Chapter 1 introduces the concept of somatic embryogenesis, describes the different SE systems in *Arabidopsis*, and discusses the role of the plant hormone auxin and chromatin modifying proteins in this process. An overview is presented on the current knowledge on SE-induction through ectopic overexpression of certain transcription factor genes. These include *BBM*, as well as other genes that are studied in this thesis in relation to BBM.

BBM is part of the eight member AIL subfamily of AP2/ERF domain transcription factors. **Chapter 2** reviews the role of AIL proteins during embryogenesis, stem cell niche specification, meristem maintenance and organ positioning and growth. We summarize the gene regulatory networks in which AILs function and describe how these transcription factors integrate multiple hormonal inputs, with special emphasis on the interactions between AILs and auxin. Finally, we conclude that although the functions of AILs in plant development are well described, knowledge on the molecular mode of action of AIL proteins and the identity of AIL target genes is still limited.

Transcription factors function in protein complexes and in **Chapter 3** we show that members of the HOMEODOMAIN GLABROUS (HDG) transcription factor family physically interact with BBM and other AILs. *HDG* genes are expressed in the epidermis, the outer cell layer of the plant, where they promote differentiation of cells into specialized epidermal cell types, such as trichomes or stomata. We show that ectopic overexpression of *HDG1* leads to loss of root and shoot meristems, phenotypes that had previously been reported for loss-of-function *ail* mutants. Conversely, down-regulation of *HDG* genes led to reduced cell differentiation, enhanced cell proliferation and SE phenotypes, phenotypes that resemble those found in *AIL* overexpression lines. Moreover, we found that co-overexpression of BBM and HDG1 reduces the overexpression phenotypes of both proteins. These results suggest opposite functions of AIL and HDG transcription factors, with AILs stimulating cell proliferation and HDGs stimulating cell

differentiation, with the ratio between the two proteins determining the developmental outcome. Finally, we show that HDGs and AILs regulate each other on a transcriptional level and that they share common target genes.

A variety of AIL overexpression phenotypes has been described in the literature, with BBM and PLT5/AIL5 being the only known AILs that induce SE upon overexpression. We show in **Chapter 4** that all AIL proteins except AIL1 and ANT are able to induce SE, but that this phenotype relies on a high AIL protein dosage. Using BBM and PLT2 as AIL representatives, we show that an intermediate AIL concentration induces organogenesis (ectopic root and shoot formation) and that a low concentration inhibits cellular differentiation. In addition, we show that BBM and PLT2 induce direct SE when activated at seed germination, while post-germination activation leads to indirect SE from callus. The *LEAFY COTYLEDON (LEC)/LAFL* genes, which also encode SE-inducing transcription factors, are direct targets of BBM/PLT2 during direct SE, showing that these two SE pathways are linked. Using *LAFL* gene mutants, we show that the LAFL pathway is an important downstream component of BBM-mediated SE.

Chapter 5 presents the *in vivo*, genome-wide analysis of BBM DNA binding sites in somatic embryos using chromatin immunoprecipitation followed by sequencing (ChIP-seq). Our ChIP-seq and gene expression analysis reveal that BBM binds and positively regulates auxin biosynthesis genes and the recently discovered positive regulators of SE, the *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED (AHL)* genes. Knock-out of either pathway reduced BBM-mediated SE, showing that auxin biosynthesis and the *AHL* genes are important components of the BBM pathway. We also show that BBM binds to a consensus DNA motif that resembles the reported ANT binding motif.

Chapter 6 reviews methods for identifying the direct target genes of a plant transcription factor using microarrays, as was done for HDG1 (Chapter 4). We describe which different systems can be used to control transcription factor activity, and how these can be combined with microarray analysis to identify target genes. In addition, we provide guidelines for the statistical analysis of microarray data and for the confirmation of candidate target genes.

In plant biology, protein-protein interactions are often studied using bimolecular fluorescence complementation (BiFC) or split-YFP. In my BBM-HDG interaction studies I encountered problems using this method, which lead to the cautionary note on the use of BiFC presented in **Chapter 7**. BiFC is based on the restoration of fluorescence after the two non-

fluorescent halves of a fluorescent protein are brought together by a protein-protein interaction event. However, because the fluorescent protein halves are prone to self-assembly, it is crucial to use proper controls and a quantitative read-out of fluorescence to avoid false positive interactions. We present a guideline for the setup of a BiFC experiment, discussing each step in the protocol.

Chapter 8 discusses how the results presented in this thesis contribute to our knowledge on AIL transcription factors and somatic embryo induction, as well as the questions that still remain. An extended model of dose-dependent AIL function is proposed, as well as mechanisms by which the AIL-HDG interaction could function at the molecular level. Finally, an overview is provided of the molecular-genetic intersection between the different transcription factor-induced SE pathways.



Samenvatting

Somatische plantencellen kunnen in weefselkweek onder de juiste omstandigheden geïnduceerd worden om embryo's te vormen, een proces dat somatische embryogenese (SE) genoemd wordt. SE is een manier om specifieke planten klonaal te vermeerderen en is daarom belangrijk voor de plantenveredeling. Bovendien zijn wetenschappers al decennia lang gefascineerd door SE, omdat het een interessante vorm is van 'totipotentie': het vermogen om een heel nieuw individu te regenereren via embryogenese. Dit proefschrift tracht een beter begrip te krijgen van SE inductie in *Arabidopsis* door de transcriptiefactor BABY BOOM (BBM), door middel van identificatie en functionele analyse van BBM-bindende eiwitten en BBM targetgenen.

Hoofdstuk 1 introduceert het concept somatische embryogenese, beschrijft de verschillende SE systemen in *Arabidopsis* en bespreekt de rol die het plantenhormoon auxine en chromatine modifierende eiwitten spelen in dit proces. Er wordt een overzicht gegeven van de huidige kennis over SE inductie door middel van ectopische overexpressie van bepaalde transcriptiefactoren. Hieronder vallen *BBM* en andere genen die in dit proefschrift worden bestudeerd in relatie tot *BBM*.

BBM maakt deel uit van de achtkoppige AIL subfamilie van AP2/ERF domein transcriptiefactoren. **Hoofdstuk 2** beschrijft de functies van AIL eiwitten tijdens embryogenese, stamcelspecificatie, meristeemonderhoud, en orgaanpositionering en -groei. We geven een overzicht van de genregulatiernetwerken waarin AILs betrokken zijn en beschrijven hoe deze transcriptiefactoren meerdere hormonale inputs integreren, met de focus op de interacties tussen AILs en auxine. Ten slotte concluderen we dat de functies van AILs tijdens plantenontwikkeling uitgebreid beschreven zijn, maar dat kennis over de moleculaire werking van AIL eiwitten en de identiteit van AIL targetgenen nog schaars is.

Transcriptiefactoren functioneren in eiwitcomplexen en in **Hoofdstuk 3** laten we zien dat leden van de HOMEODOMAIN GLABROUS (HDG) transcriptiefactorfamilie fysiek interacteren met BBM en andere AILs. *HDG* genen komen tot expressie in de epidermis, de buitenste cellaag van de plant, waar ze de differentiatie van cellen naar gespecialiseerde celtypen, zoals bladharen en huidmondjes, bevorderen. We laten zien dat ectopische overexpressie van *HDG1* leidt tot verlies van de wortel- en scheutmeristemen, fenotypes die voorheen beschreven waren voor *ail* loss-of-function mutanten. Aan de andere kant leidt verminderde expressie van *HDG* genen tot minder celdifferentiatie, meer celdelingen en SE, en lijkt daarmee op het effect dat *AIL*

overexpressie op planten heeft. Bovendien vonden we dat co-overexpressie van BBM en HDG1 de effecten van overexpressie van beide eiwitten vermindert. Deze resultaten wijzen op antagonistische functies van AIL en HDG transcriptiefactoren, waarbij AILs celdeling en HDGs celdifferentiatie stimuleren, en waarbij de verhouding tussen de twee de uiteindelijke uitkomst bepaalt. Tenslotte laten we zien dat HDGs en AILs elkaars transcriptie reguleren en dat ze gezamenlijke targetgenen hebben.

In de literatuur is een scala aan AIL overexpressiefenotypes beschreven, waarbij BBM en PLT5/AIL5 de enige AILs zijn die SE induceren wanneer ze tot overexpressie gebracht worden. Wij laten in **Hoofdstuk 4** zien dat alle AIL eiwitten behalve AIL1 en ANT SE kunnen induceren, maar dat dit fenotype een hoge dosis AIL eiwit vereist. Door BBM en PLT2 als representatieve AILs te nemen laten we zien dat een middelmatige AIL concentratie organogenese (ectopische wortel- en scheutvorming) induceert en dat een lage concentratie celdifferentiatie tegenhoudt. Verder laten we zien dat BBM en PLT2 directe SE induceren wanneer ze geactiveerd worden tijdens de zaadkieming, terwijl activatie na de kieming leidt tot indirecte SE vanuit callus. De *LEAFY COTYLEDON (LEC)/LAFL* genen, die ook coderen voor SE-inducerende transcriptiefactoren, zijn directe targetgenen van BBM/PLT2 tijdens directe SE. Dit bewijst dat deze twee 'SE routes' met elkaar verbonden zijn. Met behulp van mutanten in *LAFL* genen laten we zien dat de LAFLs een belangrijke schakel vormen in BBM-geïnduceerde SE.

In **Hoofdstuk 5** wordt de *in vivo*, genoom-brede analyse beschreven van de DNA-fragmenten die worden gebonden door BBM in somatische embryo's door middel van chromatine immunoprecipitatie en sequencen (ChIP-seq). Onze ChIP-seq en genexpressie analyse onthult dat BBM bindt aan auxine biosynthese genen en aan de onlangs ontdekte positieve SE regulatoren, de *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED (AHL)* genen, en dat BBM deze genen positief reguleert. Het uitschakelen van beide groepen genen vermindert BBM-geïnduceerde SE, wat impliceert dat auxine biosynthese en de *AHL* genen belangrijke componenten van het BBM proces zijn. Verder hebben we gevonden dat BBM een DNA motief bindt dat lijkt op het eerder gepubliceerde bindingsmotief van ANT.

In **Hoofdstuk 6** worden de methoden beschreven om de directe targetgenen van een plantentranscriptiefactor te identificeren met behulp van microarrays, net zoals ik heb gedaan voor HDG1 (Hoofdstuk 4). We geven een overzicht van de verschillende systemen die gebruikt kunnen worden om de activiteit van een transcriptiefactor te reguleren en hoe dit gecombineerd

kan worden met microarrays om targetgenen te identificeren. Verder geven we richtlijnen voor de statistische analyse van microarray data en voor de bevestiging van mogelijke targetgenen.

In de plantenwetenschap worden eiwit-eiwit interacties vaak onderzocht met behulp van bimolecular fluorescence complementation (BiFC) of split-YFP. Tijdens mijn onderzoek aan BBM-HDG interacties heb ik veel problemen gehad met deze techniek, wat geleid heeft tot een waarschuwendende publicatie over het gebruik van BiFC (**Hoofdstuk 7**). Deze techniek maakt gebruik van het herstel van fluorescentie van twee niet-fluorescente helften van een fluorescent eiwit nadat ze in elkaars nabijheid zijn gebracht door een eiwit-eiwit interactie. Maar aangezien deze helften geneigd zijn ook vanzelf samen te komen, is het cruciaal om goede controles mee te nemen en om de mate van fluorescentie kwantitatief te bepalen. We presenteren richtlijnen om een BiFC experiment op te zetten, waarbij we elke stap in het protocol bespreken.

Hoofdstuk 8 laat zien wat de resultaten in dit proefschrift bijdragen aan onze kennis over AIL transcriptiefactoren en de inductie van somatische embryo's, en welke vragen onbeantwoord blijven. Ik stel een uitbreiding voor op het bestaande model voor dosis-afhankelijke AIL werking, en bespreek mechanismen waarop de AIL-HDG interactie zou kunnen werken op moleculair niveau. Ten slotte wordt er een overzicht gegeven van de verbanden op moleculair en genetisch vlak tussen de verschillende transcriptiefactor-geïnduceerde SE routes.



Acknowledgements

Acknowledgements

So this is it, the final piece of writing! I consider my PhD a pleasant ride, although it was sometimes bumpy, and there are many people that made this possible.

First of all, Kim, I was so lucky to have you as my supervisor! I realize it is exceptional to have a supervisor that always has time for me, inspires me and pushes me when I need it. I'm very happy you made it possible for me to stay in your group and hope we will continue to form many more 'conspiracy theories' and to switch languages middenin een zin!

Gerco, allereerst bedankt voor het doorsturen van de vacature, anders was ik wellicht nooit bij PDS terecht gekomen! Verder waardeer ik al je wetenschappelijke input tijdens onze werkdiscussies en dat je me stimuleert om mezelf te ontwikkelen. Je hebt gezorgd voor een fijne sfeer binnen PDS en een hele goede infrastructuur voor je AIO's en ik ben blij dat ik er daar één van mocht zijn.

Verder gaat mijn dank uit naar iedereen binnen de embryogenese groep. Mieke, ik had geen efficiëntere of meer precieze analist kunnen wensen op mijn project! Ik durf niet eens te schatten hoeveel constructen je hebt gemaakt en hoeveel planten je getransformeerd hebt! En telkens wanneer Kim en ik wéér wat nieuws verzonnen dan klaagde jij nooit, maar klaarde de klus. Ik denk dat we een goed team vormen, vooral tijdens de eindeloze embryodissecties in het flow cabinet en het belachelijk maken van Martijn. Speaking of which, Tinus, hoewel je tegenwoordig je mannelijke kant aan het ontwikkelen bent buiten de embryogenese groep, hebben we tijdens het grootste deel van mijn PhD nauw samengewerkt. Jij leerde mij allerlei labtrucjes, was altijd geïnteresseerd in mijn onderzoek en leefde mee wanneer experimenten mislukten of ik gefrustreerd was. Ik heb altijd veel lol met je en ben blij dat je mijn paranimf bent. Tjitske, bedankt voor het regelen van al die gezellige borrels, epische clusteruitjes en cadeautjes voor iedereen. Hui, you were my fellow embryo PhD all those years and I'm happy to continue your work now. Thanks for all your help, your delicious cooking and all the best in China! Merche, I love how you simply could not leave us! Thanks for all the trips we made and I definitely won't stop e-mailing for advice. Iris, mijn mede-BBM'er, bedankt voor alle discussies die we hebben gehad, Ik denk dat we een hoop BBM mysteries hebben opgelost, hoewel er nog een aantal blijven bestaan. ☺

I would also like to thank all other PDS people for the fun times during coffee and lunch breaks, borrels, movie filming and clusteruitjes. Richard, Ruud, Romyana en Steven, bedankt voor alle

discussies en jullie advies. Marco, Michiel, Jacqueline, Froukje en Jan, bedankt voor al jullie hulp in en om het lab.

All the (ex-) open work space inhabitants: thanks for the fun trips, PhD dinners and other social activities. Sela - I still miss our morning ritual; Cezary - my open work space neighbor who's always helpful; Jenny - my former work-out buddy; Alice - thanks for all your CHIP-seq and thesis advice; Marian - for the amusing car rides. Leonie - my paranimph; Hilda, Suzanne, Suraj, Sam and Violeta: good luck finishing your theses!

Mijn eerste kennismaking met de AIL transcriptiefactoren was tijdens mijn masterstage in Utrecht. Renze en Viola, inmiddels Radix-buren: bedankt voor de open discussies, de prettige samenwerking en gezamenlijke publicaties in de afgelopen jaren.

Voor mijn FRET-FLIM experimenten was ik welkom op de afdeling Biochemie. Jan-Willem en Lisette, bedankt voor al jullie tijd, advies en gezelligheid in de krochten van het transitorium!

Thanks also to Remko, Omid and Cheryl from Leiden University for our regular somatic embryogenesis meetings and nice collaborations.

Al mijn lieve vriendjes en vriendinnetjes, bedankt voor de broodnodige ontspanning in de vorm van etentjes, festivals, feestjes, barbecues, spelletjesavonden en weekendjes weg.

Verder mijn familie en schoonfamilie; pap en mam, Bart en Esther, Pieter en Maaïke, Frans en Ineke, en Janneke en Eric: bedankt voor jullie geloof in mij, jullie steun en de interesse voor mijn onderzoek.

En tot slot, Bart, door jouw hulp kon ik mij richten op het afmaken van mijn proefschrift. Bedankt voor al je begrip en liefde en dat we nog maar lang samen leuke dingen mogen doen!

*Cheers,
Anneke*

Curriculum vitae

Anneke Horstman was born on 24 December 1984, in Weert, The Netherlands. In 2003, she moved to Nijmegen to study biology at the Radboud University. Her first interest in developmental plant genetics was sparked at the Genetics department in Nijmegen during an internship on the role of *Petunia* and *Arabidopsis* WUSCHEL homeobox (WOX) genes in flower and leaf development. Another internship at the Molecular Genetics department of Utrecht University focused on the analysis of



PLETHORA2 target genes in *Arabidopsis* root development. In addition, she wrote two theses on sex determination in monoecious and dioecious plant species and on polyploidy in plants. After graduating *cum laude* in 2009, she started a PhD project at Plant Research International in Wageningen. In this project, she studied the transcription factor BABY BOOM and its ability to induce somatic embryogenesis in *Arabidopsis*. She is currently working as a post-doc in the same group, studying the epigenetic control of microspore embryogenesis in *Brassica* and *Arabidopsis*.

Publications

Vandenbussche M, Horstman A, Zethof J, Koes R, Rijpkema AS, Gerats T. (2009) Differential recruitment of WOX transcription factors for lateral development and organ fusion in *Petunia* and *Arabidopsis*. *Plant Cell* **21**: 2269-83

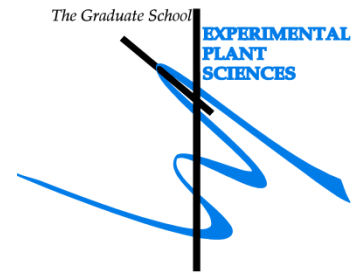
Gorte M, Horstman A, Page RB, Heidstra R, Stromberg A, Boutilier K. (2011) Microarray-based identification of transcription factor target genes. *Plant Transcription Factors: Methods and Protocols. Methods in Molecular Biology* **754**: 119 - 41.

Horstman A, Willemsen V, Boutilier K, and Heidstra R. (2014). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends in Plant Science* **19**: 146-57.

Horstman A, Tonaco IA, Boutilier K, Immink RG. (2014) A cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies. *International Journal of Molecular Sciences* **15**: 9628-43.

Horstman A, Fukuoka H, Muino JM, Nitsch L, Guo C, Passarinho P, Sanchez-Perez G, Immink R, Angenent G, Boutilier K. (2015) AIL and HDG proteins act antagonistically to control cell proliferation. *Development* **142**: 454-64.

**Education Statement of the Graduate School
Experimental Plant Sciences**



Issued to: Anneke Horstman
Date: 10 April 2015
Group: Plant Developmental Systems & Molecular Biology
University: Wageningen University & Research Centre

1) Start-up phase	<i>date</i>
▶ First presentation of your project BABY BOOM-induced somatic embryogenesis	Sep 10, 2009
▶ Writing or rewriting a project proposal Writing a review or book chapter Microarray-Based Identification of Transcription Factor Target Genes, IN ' Plant Transcription Factors: Methods and Protocols'. Humana Press 2011. pp 119-141	May 2010
▶ MSc courses	
▶ Laboratory use of isotopes	

Subtotal Start-up Phase

*7.5 credits**

2) Scientific Exposure	<i>date</i>
▶ EPS PhD student days EPS PhD student day, Wageningen EPS Get2Gether, Soest	May 20, 2011 Jan 29, 2015
▶ EPS theme symposia EPS theme 1 'Developmental Biology of Plants', Wageningen EPS theme 1 'Developmental Biology of Plants', Leiden EPS theme 1 'Developmental Biology of Plants', Wageningen EPS theme 1 'Developmental Biology of Plants', Leiden EPS theme 1 'Developmental Biology of Plants', Wageningen	Jan 08, 2010 Jan 20, 2011 Jan 19, 2012 Jan 17, 2013 Jan 24, 2014
▶ NWO Lunteren days and other National Platforms ALW meeting 'Experimental Plant Sciences', Lunteren ALW meeting 'Experimental Plant Sciences', Lunteren ALW meeting 'Experimental Plant Sciences', Lunteren ALW meeting 'Experimental Plant Sciences', Lunteren ALW meeting 'Experimental Plant Sciences', Lunteren TTI Green Genetics Networking Event	Apr 19-20, 2010 Apr 04-05, 2011 Apr 02-03, 2012 Apr 22-23, 2013 Apr 14-15, 2014 Sep 22, 2010
▶ Seminars (series), workshops and symposia Plant Science Seminar 2nd International Retreat of PhD Students in Plant Sciences, Cologne - Germany Plant Science Seminar Seminar Daniel Schubert Seminar David Baulcombe Seminar Peter Cook Workshop Microscopy by Leica Seminar Hong Ma Seminar Pierre Hilson Seminar Cristel Carles Seminar Tim Sharbel, Keygene Radix-West symposium Seminar Jose M. Seguí-Simarro EPS symposium: Omics Advances for Academia & Industry Seminar George Coupland Seminar Marcelo Dornelas	Feb 09, 2010 Apr 15-17, 2010 Jun 08, 2010 May 11, 2010 Sep 27, 2010 Oct 27, 2010 Nov 09, 2010 May 30, 2012 May 31, 2012 Jan 16, 2013 Feb 20, 2013 Okt, 2013 Sep 11, 2014 Dec 11, 2014 Jan 19, 2015 Jan 27, 2015
▶ Seminar plus	
▶ International symposia and congresses Cell-cell communication in plant reproduction, Bath Molecular Aspects of Plant Development, Vienna International conference on arabidopsis research (ICAR), Madison, USA European Frontiers of Plant Reproduction Research, Oslo	Sep 14-16, 2009 Feb 23-26, 2010 Jun 22-25, 2011 Oct 2-4, 2013

▶ Presentations	
Poster Bath	Sep 14-16, 2009
Poster Vienna	Feb 23-26 2010
Presentation Thematic Meeting Bioscience	Apr 13, 2010
Poster Lunteren	Apr 19-20, 2010
Poster Retzbach	Sep 06-08, 2010
Poster and Presentation TTI Green Genetics Networking Day	Sep 22, 2010
Presentation EPS theme 1 'Developmental Biology of Plants', Leiden	Jan 20, 2011
Presentation Thematic Meeting Bioscience	Mar 22, 2011
Poster Lunteren	Apr 04-05, 2011
Poster Madison	Jun 22-25, 2011
Poster Siena	Sep 25-28, 2012
Presentation Thematic Meeting Bioscience	Mar 20, 2012
Presentation EPS theme 1 'Developmental Biology of Plants', Leiden	Jan 17, 2013
Presentation Thematic Meeting Bioscience	Jan 22, 2013
Presentation ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 22, 2013
Poster Oslo	Oct 02-04, 2013
Presentation Thematic Meeting Bioscience	Jan 21, 2014
Presentation at INRA, Versaille	Mar 13, 2014
Presentation Thematic Meeting Bioscience	Nov 11, 2014
▶ IAB interview	
Meeting with a member of the International Advisory Board of EPS	Nov 14, 2012
▶ Excursions	
EPS Excursion to Keygene	Jan 26, 2012
<i>Subtotal Scientific Exposure</i>	
<i>32.4 credits*</i>	

3) In-Depth Studies	<i>date</i>
▶ EPS courses or other PhD courses	
3rd PhD School on Plant Development, Retzbach - Germany	Sep 06-08, 2010
EMBL Advanced course: Analysis of Next Generation Sequencing Data (Heidelberg)	Oct 16-18 2012
5th International PhD School on Plant Development, Siena - Italy	Sep 25-28, 2012
RNAseq course, Wageningen	Jan 21-23, 2015
▶ Journal club	
Participation in a literature discussion group	2009-2013
▶ Individual research training	
<i>Subtotal In-Depth Studies</i>	
<i>7.2 credits*</i>	

4) Personal development	<i>date</i>
▶ Skill training courses	
WGS Course: Information Literacy, including Introduction Endnote	Dec 07-08, 2009
Competence Assessment	Mar-Apr 2011
Course: Scientific writing	May-June 2011
EPS Career Day	Nov 18 2011
Career Perspectives	Mar-Apr 2013
▶ Organisation of PhD students day, course or conference	
▶ Membership of Board, Committee or PhD council	
<i>Subtotal Personal Development</i>	
<i>4.0 credits*</i>	

TOTAL NUMBER OF CREDIT POINTS*	51.1
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.	

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Thesis layout and cover: Anneke Horstman
Printed by Ipskamp drukkers