THE UNIVERSITY
of EDINBURGH

Understanding epidermal cell fate specification during plant embryogenesis

Rita San-Bento

Supervisor Dr Gwyneth Ingram

Thesis submitted for the Degree of Doctor in Philosophy
The University of Edinburgh
2013

Declaration

I contributed for all results described in this thesis. All collaborations are clearly indicated, specified and referenced in the manuscript.

I Table of contents

II Lay Summary

The “skin” of plants, known as the epidermal layer, is critical for plant survival, growth, and interaction with the environment. This single, continuous layer of cells first appears in very young plant embryos and is maintained throughout the entire life of the plant. In this work I used molecular biology, bioinformatics, genetics, and protein-based studies, combined with the generation of computer simulations, to understand how epidermal cells are formed, and how they are maintained in growing plants.

Two proteins, ATML1 and PDF2, had previously been shown to be important for the formation of the epidermal cell layer. In this study I show that the amount of these proteins present in the cell is very important for the formation of the epidermis in embryos. By attaching different protein labels to these two proteins in plants, I was able to show that they act in pairs to control the production of other proteins. I also demonstrated that they could potentially form pairs with five other related proteins, some of which I showed are also needed for the formation of the epidermis. This work suggests that ATML1, PDF2 and related proteins can act in many combinations, each of which potentially has a slightly different function.

To take this work further I investigated the relationship between a signalling protein, ACR4, which is involved in cell-cell communication between epidermal cells, and the regulation of ATML1/PDF2 levels. I showed that ACR4 signalling is needed to maintain high levels of ATML1 and PDF2. In contrast however, I showed that ATML1 and PDF2 act to reduce both the level of ACR4, and their own levels. These three proteins therefore regulate their own expression, either directly, or indirectly. By representing this regulation as mathematical formulae, I was able to carry out computer simulations to test how changes in the levels of specific components could affect the formation of epidermal cells. My results show that after the epidermis has been formed it is very difficult to disturb its identity. They also provide valuable clues about the signalling processes which could lead to the formation of the epidermis in the early embryo.

Finally I combined publically available information with the results of some of my experiments with the aim of discovering new proteins necessary for the formation of the plant epidermis. This part of the study will form the basis for future work in the laboratory.

III Abstract

Shoot epidermal identity is critical for plant survival, growth, and interaction with the environment. Epidermal identity is specified during very early embryogenesis, and maintained in the outermost cells of the plant throughout the entire life cycle. In this work I aimed to generate a model for the establishment of basal epidermal cell fate during embryogenesis based on the analysis of both known and novel regulators.

Loss of function of two HD-ZIP IV transcription factors, ATML1 and PDF2 had previously been shown to lead to embryo lethality due to loss of epidermal specification. In this study I uncover dosage dependency of ATML1 and PDF2 function during embryogenesis. By expressing functional ATML1 and PDF2 fusion proteins specifically in the epidermis, I developed a novel tool allowing demonstration of homo- and heterodimerization of these two transcription factors *in planta*. Using genetic and proteomic analysis I provide evidence that other HD-ZIP IV proteins are involved in epidermal specification together with ATML1 and PDF2, suggesting the presence of multiple regulatory protein complexes.

Based on previous published and unpublished work, I tested the hypothesis that ATML1 and PDF2 form part of a regulatory feedback loop necessary for maintenance of epidermal identity, and involving cell-cell signalling mediated by the receptor kinase ACR4. Using a genetic approach I confirm that ATML1 and PDF2 likely act together with ACR4 in the specification of embryonic epidermal identity. I show that ATML1 and PDF2 negatively regulate both *ACR4*, and their own expression, most likely by binding to L1 box motifs. In contrast, I provide evidence that ACR4-mediated signalling participates in maintaining expression levels of *ATML1* and *PDF2*. Mathematical modelling of the properties of the feedback loop supported by my results, suggests that it is capable of maintaining a robust epidermal cell fate, and predicts possible changes in network interactions during the process of epidermal cell fate specification.

Finally I used a combination of bioinformatics approaches to integrate *in silico* and experimental data with the aim of discovering potential novel epidermal regulators

and targets of epidermal fate specifying pathways. This work highlighted potential roles for WOX-family transcription factors in epidermal fate specification, which were further analysed genetically. In addition, bioinformatics analysis pinpointed an intriguing overlap between the targets of epidermal specification pathways and targets of abiotic stresses signalling.

IV Acknowledgments

This project took place in the University of Edinburgh and the ENS de Lyon. I would like to thank all my colleagues at IMPS and RDP laboratories for the exceptional work environment, support and sharing their passion for science. Most importantly, I want to thank my supervisor Gwyneth Ingram for the friendship, unconditional support and excellent direction of my work. All I learned and accomplished in these years I owe it to her brilliant supervision and remarkable motivation at the scientific and personal level.

I would like to thank everyone in the Marie Curie ITN SIREN project for the companionship and scientific opportunities. In particular, I want to thank Cristina Llavata-Peris and Georgios Lagiotis for their help in the CoIP experiments. In addition, I would like to thank Audrey Creff, Roberta Galletti and Etienne Farcot that contributed directly to this work.

At a personal level, there are too many people I would like to thank for being there and make this challenging experience unforgettable. The ups and downs of a project staged in two countries described a wonderful journey, full of memorable moments. I apologize in advance for the names unspoken but unforgettable which are not present in the following list. In Edinburgh, several friends took a special place in my life, namely Quique, Julie, Sarah, Andrew and Erica. Their friendship was very special to me and contributed for an amazing year in the wonderful city of Edinburgh. In France, I discovered a remarkable lab full of life and science. I would like to thank Benoit, Kai, Michiel, Thomas, Audrey, Roberta, Davide, Vincent and Julie amongst everyone in theRDP for their friendship, brilliant discussions and collaborations.

Finally, I would like to thank all my friends and family, who mean the world to me, for always being there, especially my two brilliant sisters and parents.

Thank you all! Merci à tous! Muito obrigada!

V List of Abbreviations

List of Abbreviations

µg	Microgram
µm	Micrometre
µm	Microlitre
ABA	Abscisic acid
ARE	Auxin Response Element
At	*Arabidopsis thaliana*
bHLH	Basic helix-loop-helix domain
bp	Base pair
C	Cytidine
ChIP	Chromatin ImmunoPrecipitation
cm	Centimetre
CO2	Carbon dioxide
CoA	Coenzyme A
CoIP	Co-Immunoprecipitation
Col0	Columbia Arabidopsis ecotype
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dTTP	2'-deoxythymidine 5'-triphosphate
EDTA	Ethylene diaminetetraacetic acid
EMSA	Electrophoretic mobility shift assays
ENS	École Normale Supérieure in Lyon
ER	Endoplasmatic Reticulum
ESR	Embryo surrounding region
FAE	Fatty acid elongase complex
g	Gram

G	Guanine
GC	Guanine/cytidine content
GO	Gene Ontology
GUS	β-glucuronidase
h	Hour
HD	Homeodomain
Hz	Hertz
IAA	Indole 3-acetic-acid
IP	Immunoprecipitation
kb	Kilobase (1000 DNA base pairs)
kV	Kilovolt
LB	Lauria Broth
LC	Liquid chromatography
Ler	Landsberg erecta Arabidopsis ecotype
LTQ	Label-free quantification
M	Molar
mg	Milligram
min	Minute
ml	Millilitre
mm	Millimetre
mM	Millimolar
MS	Mass spectrometry
MS	Murashige and Skoog
ng	Nanogram
nLC-MS/MS	Nano Liquid chromatography-tandem mass spectrometry
nm	Nanometre
nt	Nucleotides
ODE	Ordinary Differential Equations
Os	*Oryza Sativa*
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pH	Potential of Hydrogen

PO	Protein Ontology
QC	Quiescent centre
Q-PCR	Quantitative polymerase chain reaction
RDP	Reproduction and Development of Plants (ENS department)
RNA	Ribonucleic acid
RPKM	Reads per kilobase per million
Rpm	Revolutions per minute
RUB	Related-to-ubiquitin
S	Second
S	Cytidine or guanine
SAD	START-adjacent domain
SAM	Shoot Apical Meristem
SDS	Sodium dodecyl sulphate
SEM	Scanning electron microscope
START	Steroidogenic acute regulatory protein-related lipid transfer
TAIR	The Arabidopsis Information Resource
TF	Transcription Factor
TNFR	Tumor necrosis factor receptor
UTR	Untranslated region
VLCFAs	Very long chain fatty acids
WS	Wassilewskija Arabidopsis ecotype
WT	Wild-type
xxYy	Genotype of line homozygous for x allele and hetozygous for Y allele
Y	Cytidine or thymidine
Zm	*Zea mays*

List of genes

ABCG	*ATP-BINDING CASSETTE G*
ACR4	*ARABIDOPSIS CRINKLY 4*
AFB	*AUXIN SIGNALING F BOX PROTEIN*
AGO	*ARGONAUTE*
ALE	*ABNORMAL LEAF SHAPE*
AN	*ANGUSTIFOLIA*
ANL2	*ANTHOCYANINLESS 2*
AP3	*APETALA3*
APRR2	*PSEUDO-RESPONSE REGULATOR2*
ARF	*AUXIN RESPONSE FACTOR*
ARF-GEF	*ADP-RIBOSYLATION FACTORS AND GUANINE NUCLEOTIDE EXCHANGE FACTORS*
ARR	*ARABIDOPSIS RESPONSE REGULATOR*
AS	*ASYMMETRIC LEAVES*
ATBZIP	*BASIC LEUCINE-ZIPPER*
ATCSLC8	*CELLULOSE-SYNTHASE LIKE C8*
ATDEK1	*ARABIDOPSIS THALIANA DEFECTIVE KERNEL*
ATHB	*HOMEOBOX PROTEIN*
ATML1	*ARABIDOPSIS THALIANA MERISTEM LAYER 1*
ATPLC4	*PHOSPHATIDYLINOSITOL-SPECIWC PHOSPHOLIPASE C4*
ATSR	*ARABIDOPSIS THALIANA SERINE*
AXL	*AXR1- LIKE*
AXR	*AUXIN RESISTANT*
BCAT-1	*BRANCHED-CHAIN AMINO ACID TRANSAMINASE 1*
BDG	*BODYGUARD*
BDL	*BODENLOS*
BGLU23	*BETA-GLUCOSIDASE23*
CD2	*CUTIN DEFICIENT2*
CER	*ECERIFERUM*
CFL1	*CURLY FLAG LEAF1*
CLE	*CLV3/ESR-RELATED*
CLV	*CLAVATA*
CML38	*CALMODULIN-LIKE 38*

CNA	*CORONA*
COR	*COLD-REGULATED*
CPC	*CAPRICE*
CR4	*CRINKLY4*
CRCK1	*CALMODULIN-BINDING RECEPTOR-LIKE CYTOPLASMIC KINASE*
CRK2	*CYSTEINE-RICH RLK*
CRR	*CRINKLY4 RELATED*
CUC	*CUP-SHAPED COTYLEDON*
CUL1	*CULLIN1*
DDF	*DWARF AND DELAYED FLOWERING*
DEK1	*DEFECTIVE KERNEL1*
DRB	*DSRNA-BINDING PROTEIN*
DRN	*DORNOSCHEN*
DRNL	*DRN-LIKE*
ECR	*ENOYL-COA REDUCTASE*
ECR1	*E1 C-TERMINAL RELATED 1*
EGL3	*ENHANCER OF GL3*
EPF	*EPIDERMAL PATTERNING FACTOR*
ER	*ERECTA*
ERD	*EARLY-RESPONSIVE TO DEHYDRATION*
ERF	*ETHYLENE RESPONSE FACTOR*
ERL	*ERECTA-LIKE*
ETC1	*ENHANCER OF TRY AND CPC1*
ETR	*ETHYLENE RESPONSE*
FAS	*FATTY ACID SYNTHASE*
FATB	*ACYL ACP THIOESTEASE*
FDH	*FIDDLEHEAD*
FIL	*FILAMENTOUS FLOWER*
FLA	*FASCICLIN-LIKE ARABINOGALACTAN PROTEIN*
FLP	*FOUR LIPS*
FSD	*FE SUPEROXIDE DISMUTASE*
FWA	*FLOWERING WAGENINGEN*
GFP	*GREEN FLUORESCENT PROTEIN*
GL1	*GLABRA1*

GL2	*GLABROUS2*
GL3	*GLABRA3*
GN	*GNOM*
GOLS3	*GALACTINOL SYNTHASE 3*
GPAT	*GLYCEROL-3-PHOSPHATE SN-2-ACYLTRANSFERASE 8*
GRD	*GROUNDED*
GSO	*GASSHO*
HAN	*HANABA TARANU*
HAT3	*HOMEOBOX-LEUCINE ZIPPER PROTEIN*
HB	*HOMEOBOX PROTEIN*
HCD	*B-HYDROXYACYL-COA DESYDRATASE*
HDG	*HOMEODOMAIN GLABROUS*
HD-ZIP	*HOMEODOMAIN LEUCINE ZIPPER*
HR4	*HOMOLOG OF RPW8 4*
HS90.1	*HEAT SHOCK PROTEIN 90-1*
HVA22D, E	*HVA22 HOMOLOGUE*
HY5	*ELONGATED HYPOCOTYL 5*
IAA	*INDOLE-3-ACETIC ACID INDUCIBLE*
ICR1	*INTERACTOR OF CONSTITUTIVE ACTIVE ROPS 1*
IQD	*IQ-DOMAIN*
JAZ	*JASMONATE-ZIM-DOMAIN PROTEIN*
KAN	*KANADI*
KCR	*B-KETOACYL-COA REDUCTASE*
KCS	*B-KETOACYL-COA SYNTHETASE*
KRP1	*KIP-RELATED PROTEIN1*
LACS	*LONG CHAIN ACYL-COA SYNTHETASE*
LGO	*LOSS OF GIANT CELLS FROM ORGANS*
LOX3	*LIPOXYGENASE3*
LTP	*LIPID TRANSPORT PROTEIN*
LTPG1	*GLYCOSYL-PHOSPHATIDYL-INOSITOL-ANCHORED LIPID TRANSFER PROTEIN*
MAB2	*MACCHI BOU 2*
MAH	*MID-CHAIN ALKANE HYDROXYLASE*
MAP18	*MICROTUBULE-ASSOCIATED PROTEIN 18*
MAPK	*MITOGEN ACTIVATED PROTEIN KINASE*

MKK	*MAP KINASE KINASE*
MLP43	*MLP-LIKE PROTEIN 43*
MP	*MONOPTEROS*
MPK	*MAP KINASE*
MYB	*MYB DOMAIN PROTEIN*
NAC	*NAC DOMAIN CONTAINING PROTEIN*
NAT12	*NUCLEOBASE-ASCORBATE TRANSPORTER 12*
NGA	*NGATHA*
NIT	*NITRILASE*
NRPD1B	*NUCLEAR RNA POLYMERASE D1B*
OBE	*OBERON*
OCL	*OUTER CELL LAYER*
OFP15	*OVATE FAMILY PROTEIN 15*
OPR	*OXOPHYTODIENOATE-REDUCTASE*
PAP29	*PURPLE ACID PHOSPHATASE 29*
PAS	*PASTICCINO*
PBP1	*PINOID-BINDING PROTEIN 1*
PDCB3	*PLASMODESMATA CALLOSE-BINDING PROTEIN 3*
PDF1	*PROTODERMAL FACTOR1*
PDF2	*PROTODERMAL FACTOR2*
PEL3	*PERMEABLE LEAVES3*
PEN1	*PENETRATION1*
PHB	*PHABULOSA*
PHV	*PHAVOLUTA*
PID	*PINOID*
PIN	*PIN-FORMED*
PK1B	*PROTEIN KINASE 1B*
PLDGAMMA3	*PHOSPHOLIPASE D GAMMA 3*
PLT	*PLETHORA*
PME	*PECTIN METHYLESTERASE*
POR	*OXOPHYTODIENOATE-REDUCTASE*
PP2A	*PROTEIN PHOSPHATASE 2A*
PRS	*PRESSED FLOWER*
PSB29	*PHOTOSYSTEM II REACTION CENTER PSB29 PROTEIN*
QWRF	*QWRF DOMAIN CONTAINING*

RAB18	*RESPONSIVE TO ABA18*
RALF	*RAPID ALKALINIZATION FACTOR*
RCE1	*RUB1 CONJUGATING ENZYME 1*
REV	*REVOLUTA*
RLK	*RECEPTOR-LIKE KINASE*
RLP	*RECEPTOR LIKE PROTEIN*
RPK	*RECEPTOR-LIKE PROTEIN KINASE*
SAG	*SENESCENCE-ASSOCIATED GENE*
SCPL25	*SERINE CARBOXYPEPTIDASE-LIKE 25*
SCR	*SCARECROW*
SCR2	*SOYBEAN GENE REGULATED BY COLD-2*
SCRM	*SCREAM*
SDD1	*STOMATAL DENSITY AND DISTRIBUTION 1*
SEP	*SEPALLATA*
SHR	*SHORT ROOT*
SPCH	*SPEECHLESS*
SPL	*SQUAMOSA PROMOTER BINDING PROTEIN-LIKE*
SRO	*SIMILAR TO RCD ONE*
SSP	*SHORT SUSPENSOR*
STI	*STICHEL*
STIP	*STIMPY/WOX9*
STM	*SHOOTMERISTEMLESS*
SUC5	*SUCROSE-PROTON SYMPORTER 5*
SVB	*SMALLER WITH VARIABLE BRANCHES*
TAA	*TRYPTOPHAN AMINOTRANSFERASE*
TAR	*TRYPTOPHAN AMINOTRANSFERASE RELATED*
TCH	*TOUCH*
TIR1	*TRANSPORT INHIBITOR RESPONSE 1*
TMM	*TOO MANY MOUTHS*
TMO	*TARGET OF MONOPTEROS*
TOAD2	*TOADSTOOL 2*
TPL	*TOPLESS*
TRY	*TRYPTICON*
TTA	*TITANIA*
TTG1	*TRANSPARENT TESTA GLABRA1*

WBC	*WHITE-BROWN COMPLEX HOMOLOG PROTEIN*
WIN1	*WAX INDUCER 1*
WOX	*WUS RELATED HOMEOBOX*
WRKY	*WRK DNA-BINDING PROTEIN*
WUS	*WUSCHEL*
XTH	*XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE*
YDA	*YODA*
YUC	*YUCCA*
ZMHDZIV	*ZEA MAYS HOMEODOMAIN LEUCINE ZIPPER*
ZOU	*ZHOUPI*

Chapter 1 – Introduction

1.1 Embryogenesis

Angiosperms encompass a huge diversity of flowering plant species with different dimensions, complexity, shapes and colours. However the first steps in the development of each species are the same: the development of a seed containing the mature embryo in which the main features of the plant are already established. The objectives of embryogenesis are common in all higher plants. The first phase of embryo development involves the establishment of the apical-basal axis which will later orient the shoot apical meristem (SAM) and root meristem, and consequently the plant aerial organs and root respectively. The radial patterning of tissues (establishment of an inside and an outside) is essential for tissue organization. In *Arabidopsis*, radial patterning occurs at the globular stage and is followed by the establishment of shoot and root meristems and by cotyledon initiation at the heart stage which defines bilateral symmetry. Cotyledon elongation occurs at the torpedo stage and the embryo occupies nearly the whole seed at the maturation stage (Figure1.1).

Although embryo development is a leading actor in seed development, several other tissues work together during this process. In *Arabidopsis*, the seed consists of four main tissues: the seed coat, the endosperm, the suspensor and the embryo (Figure 1.1). The seed coat is a maternal sporophytic tissue and contains the nucellar/chalazal region, and tissues derived from the outer and inner integuments of the ovule, which form four or five cell layers surrounding the embryo sac (Figure 1.1A). In particular, the outer layer of the seed coat is the first defensive barrier of the mature seed and is composed of specialized cells that produce mucilage as an aid to seed imbibition and germination. The main functions of the mature seed coat are the protection of embryo and seed reserves against abiotic and biotic stress (Beeckman et al. 2000; Rajjou and Debeaujon 2008). However, during early seed development this tissue acts as a highway for nutrient transport to the developing embryo (Ingram 2010; Stadler et al. 2005). The suspensor is derived from the zygote (the fertilized egg cell) and is therefore genetically identical to the embryo. Its role in embryogenesis is in some ways comparable to that of the mammalian placenta, since it is involved in absorbing nutrients transported from the seed coat through the endosperm at early

stages of embryo development (Kawashima and Goldberg 2010; Yeung and Meinke 1993). Additionally, the suspensor is required for the positioning of the embryo in the seed. At later stages of seed development when it is no longer required, the suspensor degenerates.

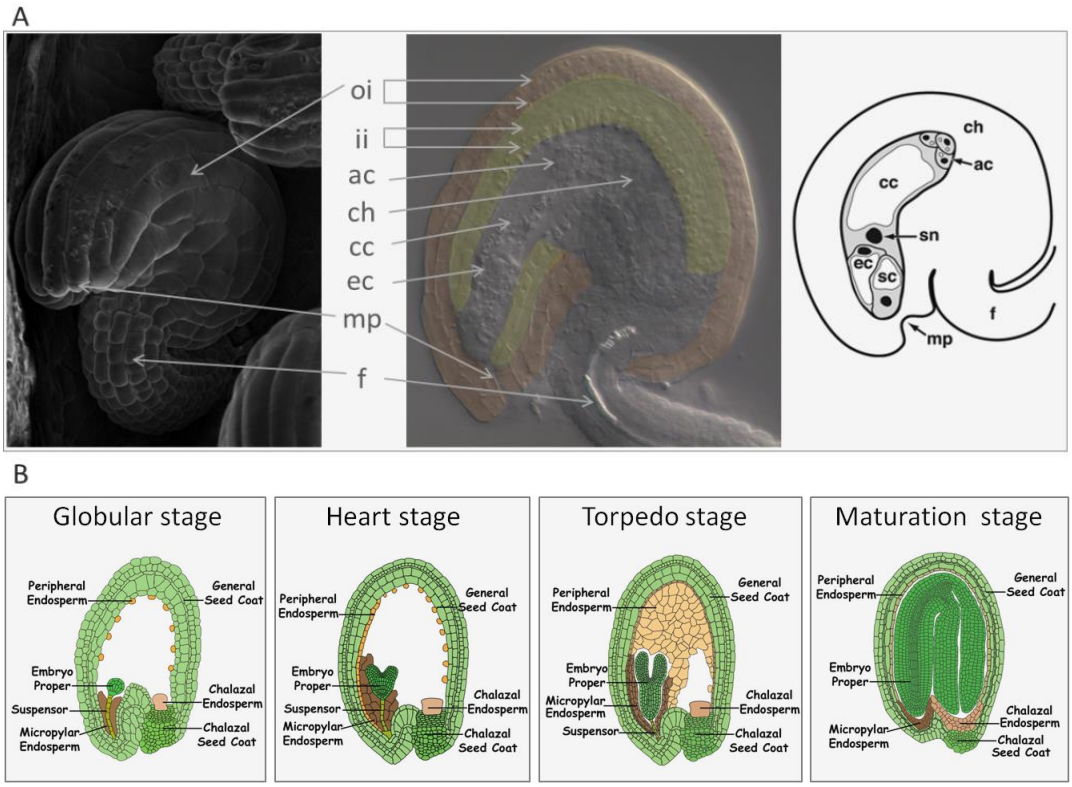


Figure 1.1 – Seed development in *Arabidopsis*. A) Ovule structure, the inner (ii) and outer (oi) integuments (yellow and orange respectively) cover the chalazal region (ch) and the female gametophyte composed of the egg cell (ec) and synergid cell (sc) localized next to the micropyle (mp), the central cell (cc) and antipodal cells (ac). B) Representation of the different stages of embryo development. (Adapted from Goldberg *et al* 2010 Gene Networks in Seed Development website (Le et al. 2010) and (Yadegari and Drews 2004))

The endosperm, which in *Arabidopsis* is formed by the fusion of a double haploid female gametophyte cell called the central cell, and a sperm nucleus, is a vital nutrient source for the embryo. In *Arabidopsis* it also plays a key role in creating the physical space within the seed coat that is necessary for subsequent embryo development. The endosperm has a highly modified cell cycle, and undergoes several rounds of nuclear division in the absence of cytokinesis, to form a syncitium of nuclei in a shared cytoplasm surrounding a central vacuole. It is the rapid expansion of this structure which drives most of the growth of the seed coat, well before the

embryo starts to grow significantly (Berger 2003; Olsen 2004). The micropylar/chalazal polarity of the endosperm results in the definition of three regions of the multinuclear single cell, the coenocyte (Berger 2003; Olsen 2004): peripheral endosperm, embryo surrounding region (ESR) or micropylar endosperm, and chalazal endosperm (Figure 1.1). The syncytial stage of endosperm development is followed by endosperm cellularization, after which, in *Arabidopsis*, the endosperm is consumed as the embryo grows to occupy most of the seed. Ultimately, the endosperm is reduced to one cell layer, the aleurone layer, which plays a critical role in regulating seed dormancy (Bethke et al. 2007). The coordination of endosperm and embryo growth requires developmental signals between the different seed tissues (reviewed in (Berger et al. 2006; Ingram 2010)). The signalling between the embryo and the ESR, and its impact on embryo patterning is very poorly understood.

1.2 Embryo patterning

In *Arabidopsis*, embryo development establishes the main features of the plant, namely root and shoot identity, and radial patterning. Embryo development requires a complex network of cell-cell communication, hormone signalling and specific transcriptional regulation (reviewed in (Jenik et al. 2007; Lau et al. 2010; Lau et al. 2012)). In a bottom-up approach several genes involved in this network have already been described, but the integration of different types of regulation and their impact on cell differentiation remains poorly understood. One of the critical events in embryo development, and the major focus of this thesis, is the formation of a functional protodermal cell layer, which will subsequently give rise to the shoot epidermis.

1.2.1 Description

The egg cell of *Arabidopsis* has a highly polar structure, with the nucleus and most of the cytoplasm located apically (Figure 1.1A). After zygote fertilization, this polar organization becomes even more marked, and is fixed by the first asymmetric cell division which occurs horizontally (i.e. perpendicular to the apical-basal axis). After the first division of the zygote, each daughter cell undergoes an independent cell division program which will shape the embryo and the suspensor.

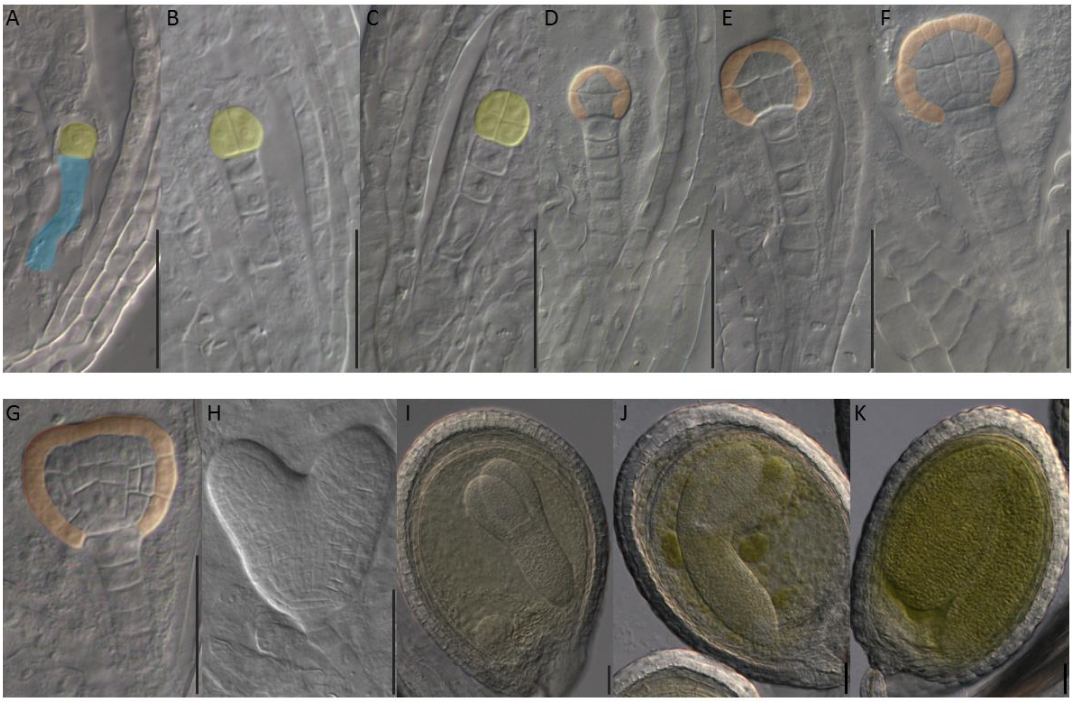


Figure 1.2 – Embryo development in *Arabidopsis*. The first division of the zygote (A) delimits the suspensor (blue) and embryo proper (yellow) cell lineage. The early stages of development encompass 2/4-cell embryo (B), 8 cell embryo (C), dermatogen (D), and globular stage (E-G). The epidermis (orange) is defined in the dermatogen stage and the complete radial patterning is defined at the globular stage (F). Cotyledon initiation defines the heart stage (H) followed by cotyledon elongation (I) and bending (J) until the maturation stage (K).

The larger basal cell will give rise to the lineage responsible for the formation of the suspensor (Figure 1.2 A in blue). The suspensor is a structure made of a file of seven to nine cells. The uppermost cell of the suspensor, the hypophysis, is incorporated into the embryo at the globular stage and differentiates to form the quiescent centre (QC) precursors in the nascent root meristem (Figure 1.2 F). The smaller, apical cell, gives rise to most of the embryo proper (Figure 1.2 A, B in yellow). This cell first divides vertically (parallel to the apical basal axis) twice, and then once horizontally, forming the 8-cell embryo proper (Figure 1.2C). The basal and apical cell populations of the embryo proper are separated at this stage. The subsequent oblique and asymmetrical periclinal cell division of the 8 cells of the embryo proper separates the protoderm, precursor of the epidermal layer, from a population of internal cells (Figure 1.2 D). The 16-cell embryo is also known as the dermatogen stage and comprises eight inner cells and eight outer cells organized in two cell files in the apical and basal regions of the embryo proper. From this stage onwards the

cell divisions of the protoderm are anticlinal (Figure 1.2 D-G in orange). During the passage between dermatogen and the 32-cell embryo the inner cells are divided vertically and the outer cells divide anticlinally (Figure 1.2E). Up until this point the division cycles occur almost in synchrony.

At the globular stage (Figure 1.2 F) two types of asymmetric division occur in the basal region of the embryo. The four inner cells divide asymmetrically giving rise to a layer of short upper cells and long lower cells defining the vascular tissue initials. The hypophysis (upper cell of the suspensor) is divided into a smaller disk shape cell on the top of a larger cell. These correspond to the precursors of the root quiescent centre and stem cells respectively, thus initiating the root meristem.

In summary, by the globular stage of embryogenesis the apical/basal cell fate and radial patterning are already established. With the definition of the shoot meristem, two opposite zones in the apical part of the embryo increase their growth rate to form the cotyledon initials at the heart stage (Figure 1.2 H). The final stages of embryogenesis involve cell growth and cell elongation, in particular the elongation and bending of the cotyledons (Figure 1.2 I, J). At the maturation stage the embryo occupies most of the seed and is prepared for dormancy or germination.

This introduction is dedicated to early stages of development since the most important events of embryo cell patterning are initiated at these stages, in particular the specification of epidermal cell fate.

1.2.2 Auxin, the master of embryo patterning

Plant hormones regulate plant growth, induce developmental programs and act as signal transducers. The instructive signals provided by plant hormones result in the establishment of developmental domains, specification of cell fates and diverse types of growth responses. It is therefore unsurprising that plant hormones play a key role in embryogenesis. The hormone most closely associated with embryo patterning is auxin, although roles for other hormones cannot be excluded. For example, the synthetic reporter for the output of cytokinin signalling is first expressed in the hypophysis at early stages of embryogenesis, and after the asymmetric cell division of this cell the expression of this reporter is repressed in the larger basal cell where

an auxin peak is located, suggesting that auxin and cytokinin could play antagonistic roles in root pole specification (Muller and Sheen 2008). Consistent with this, auxin has been shown to activate the expression of two type-A ARRs (ARABIDOPSIS RESPONSE REGULATORs) that in turn repress cytokinin signalling (Muller and Sheen 2008). However a role for cytokinin in specification of embryonic root has yet to be shown. Despite this and other examples, auxin is the undeniable master hormone of early embryogenesis. The starring role of auxin in embryo cell patterning is orchestrated by a complex and dynamic spatial and temporal distribution that, when disrupted, leads to several phenotypes related to defects in cell patterning (Figure 1.3A).

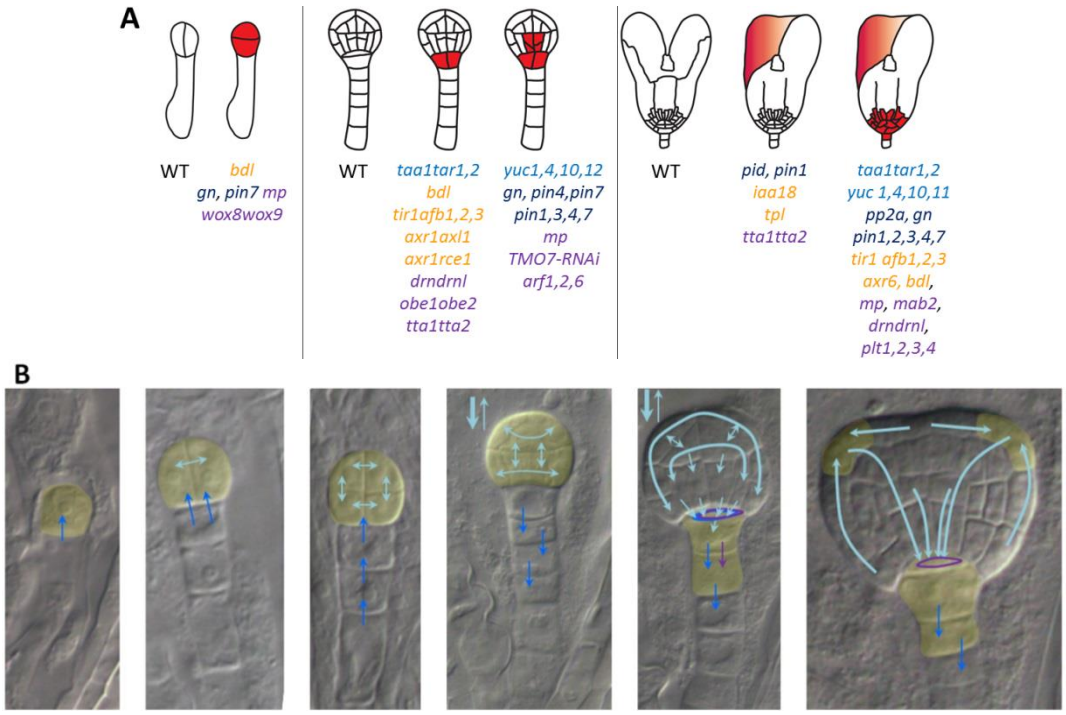


Figure 1.3 – Auxin in embryogenesis. A) Mutant embryo phenotypes (red) associated with disruption of auxin synthesis (light blue), transport (dark blue), response (orange) and regulation (purple). B) Distribution of PIN transporters (PIN1, PIN4 and PIN7 mediated auxin flux are indicated by light blue, purple and blue arrows respectively) and auxin peaks (DR5 promoter activity shown in yellow) during embryogenesis. (Adapted from (Benkova et al. 2003; Friml et al. 2002; Friml et al. 2003; Moller and Weijers 2009; Steinmann et al. 1999; Tanaka et al. 2006; Vieten et al. 2005))

The specification of this lively auxin pattern requires several layers of regulation, from biosynthesis, through transport, to perception and regulation of downstream transcriptional targets. The disturbance of auxin signalling generally leads to four main phenotypes with significant impact on embryo patterning (Figure 1.3 A). 1) In

some auxin-related mutants the first division of the embryo proper is transversal instead of longitudinal. 2) Typically auxin mutant embryos are unable to perform the asymmetrical divisions of the hypophysis and vascular precursors. 3) At later stages of development, auxin mutants display defects in the root pole that can lead to rootless seedling phenotypes. 4) Finally, cotyledon initiation and positioning tends to be impaired in these mutants resulting in phenotypes from asymmetric cotyledon initiation, to fused or abnormal numbers of cotyledons. In severe cases, auxin disruption leads to the lack of cotyledons.

1.2.2.1 Auxin distribution in early embryogenesis

Auxin signalling in plants is in some ways similar to a series of electric circuits that require a positive and negative pole to direct electric charge. The directionality of auxin flux, mediated by PIN (PIN-FORMED) efflux transporters, results in the establishment of domains of high and low auxin levels that induce specific developmental programs depending on their context. The polar localization of PIN transporters in the cell defines the direction of auxin flux and is critical in the regulation of auxin flux (reviewed in (Dettmer and Friml 2011)). Importantly, auxin flux is thought to positively reinforce the polarity of PIN transporters, and thus the positions of auxin maxima (Dettmer and Friml 2011; Tanaka et al. 2006; Vieten et al. 2005).

As mentioned above, during embryogenesis, auxin is involved in root meristem formation and cotyledon initiation. Consistent with this, at the globular stage of embryogenesis, auxin is synthesized in the apical pole of the embryo and directed to the basal pole by polarized PIN transporters to activate the auxin responsive genes that determine root pole formation. Later the auxin flux is redirect to the region where cotyledons will be initiated forming two new auxin peaks (Figure 1.3B) activating the necessary programs to define cotyledon development.

Perturbations of auxin synthesis or auxin transport disrupt auxin distribution and consequently auxin response (Figure 1.3A). In general, the detection of auxin distribution is carried out indirectly by monitoring the expression of auxin responsive promoters. The most common auxin marker is the DR5 synthetic element comprising

seven repeats of the ARE (Auxin Response Element) (Sabatini et al. 1999; Ulmasov et al. 1997). The ARE sequence is the cognate target of a family of transcription factors called ARF (AUXIN RESPONSE FACTOR), which act downstream of auxin perception (Ulmasov et al. 1999). The DR5 reporter is therefore a purely qualitative marker, the expression of which depends on the presence of an intact and active auxin signalling cascade in the cells studied.

At early stages of embryogenesis (Figure 1.3B yellow) the DR5 marker is detected in the embryo proper from the one cell stage until the dermatogen stage, when DR5 activity switches to the basal pole of the embryo, comprising the hypophysis and a few neighbouring cells in the suspensor (Benkova et al. 2003; Friml et al. 2003). At the transition stage (late globular stage), DR5 is also detected in the region of cotyledon primordia, and at heart/torpedo stage in the provasculature (Friml et al. 2003). The activity of DR5 has been shown to be similar to that of IAA (indole-3-acetic acid) in immunolocalization studies during embryogenesis (Benkova et al. 2003; Friml et al. 2003) supporting the fact that DR5 is a reliable marker of auxin accumulation in these tissues.

Several auxin IAA biosynthetic pathways have been proposed to exist in plants, of which the best known, and seemingly the most important for embryogenesis in *Arabidopsis*, involve synthesis from indole via tryptophan (reviewed in (Mano and Nemoto 2012; Zhao 2010)). Several protein families are involved in auxin synthesis via these pathways, of which the YUC (YUCCA) family and TAA1/TAR1/TAR2 (TRYPTOPHAN AMINOTRANSFERASE RELATED) family are perhaps the best characterized (Cheng et al. 2006; 2007; Stepanova et al. 2008; Tao et al. 2008). In particular *TAA1*, *YUC1* and *YUC4* are expressed in the apical part of the embryo from the globular stage onwards and have overlapping expression patterns with *YUC10* and *YUC11*. Consistent with the important role of these proteins in embryogenesis, *taa1tar1,2* and *yuc1,4,10,11* mutants display phenotypes associated with impaired auxin responses. In addition the DR5 auxin marker is undetectable at later stages of *taa1tar1,2* embryogenesis. (Cheng et al. 2007; Stepanova et al. 2008).

During the earliest stages of embryogenesis, auxin does not appear to be synthesized by the embryo, but is imported from supporting tissues. As mentioned above, the

correct localization of PIN transporters (Figure 1.3B) is required to direct auxin distribution, and provide spatial and temporal changes in auxin maxima. In very early embryos (1-cell to 8 cell embryo), the auxin flux is apically oriented by PIN7 localization in the apical membranes of the cells of the basal lineage to create the auxin peak in the apical zone of the embryo. At the 16-32-cell stage, PIN7 localization switches to the basal pole (hypophysis and suspensor cells) to induce the auxin peak necessary for the root pole cell patterning (Friml et al. 2003). Meanwhile, from the 2 cell stage until the dermatogen stage, PIN1 is localized in all the plasma membranes of the whole embryo proper, with the exception of membranes facing the outside or the suspensor. At the dermatogen stage, PIN1 polarity switches, driving auxin transport in an apical to basal direction, in particular to the hypophysis, and at globular and late globular stages this directionality is reinforced (Benkova et al. 2003; Friml et al. 2003; Steinmann et al. 1999). At the late globular stage and heart stage, during cotyledon primordium initiation, PIN1 distribution in the epidermis reorients to give two maxima on either side of the apical zone of the embryo (cotyledon primordia), while PIN1 in internal tissues continues to drain auxin towards the root pole. Auxin flow away from the cotyledons in internal tissues triggers patterning of the vasculature of these organs (Benkova et al. 2003; Friml et al. 2003; Grunewald and Friml 2010; Tanaka et al. 2006).

Other PIN proteins help to reinforce the PIN7/PIN1 mediated auxin distribution in the developing embryo. PIN4 is thought to reinforce the auxin gradient in the basal pole of the embryo. In particular, PIN4 is localized at the globular stage to the hypophysis and basal membrane of the neighbouring suspensor cell. Once the hypophysis has divided, PIN4 is localized particularly in the membranes of the apical daughter cell, the precursor of the quiescent centre, and in the basal poles of apical and basal neighbouring cells. At later stages PIN4 is restricted to the quiescent centre and the basal pole of provascular cells (Friml et al. 2002; Friml et al. 2003; Vieten et al. 2005). Additionally, PIN3 has been reported to be expressed in the root pole from the heart stage onwards (Friml et al. 2003). In conclusion, the distribution of PIN transporters correlates with the auxin distribution determined by DR5 expression.

The perturbation of PIN distribution in embryos causes changes in auxin distribution. DR5 expression in *pin4* and *pin7* mutants is altered compared to the wild-type. The *pin7* mutant is unable to establish an apical/basal auxin gradient during the first stages of embryogenesis, and DR5 is actively expressed in the centre of the embryo from globular to heart stage and only localizes in the basal pole at much later stages (Friml et al. 2002; Friml et al. 2003). Such disruptions in auxin distribution can cause disturbances in the localization of other PIN transporters (Friml et al. 2003; Vieten et al. 2005). Changes in PIN distribution and auxin pattern are correlated with defects in embryo patterning, especially in the basal pole. The increasing embryo phenotype severity of single to multiple *pin* mutants supports functional redundancy of PIN transporters in the embryo. A proportion of *pin1* and *pin4* embryos show defects in the basal pole of the embryo (Friml et al. 2002; Friml et al. 2003). The disruption of *PIN7* function entails problems in the definition of apical/basal cell fates and boundaries caused by defects in auxin distribution and again these are transduced into basal or apical defects that range from miss-divisions in the basal part of the embryo to the growth of two embryos one on top of the other. In severe cases, embryo arrest occurs at early stages of development due to a failure to develop the embryo proper (Friml et al. 2003; Vieten et al. 2005). However, in most *pin7* mutant embryos, embryo development is restored to near normality at the globular stage as PIN1 and PIN4 are able to induce the auxin gradient switch, supporting PIN functional redundancy (Friml et al. 2003). Although *pin1pin7* double mutant phenotypes did not differ significantly from that of *pin7*, the double mutant *pin4pin7* displays abnormal numbers of cotyledons, and *pin1,3,4* triple mutants additionally display fused cotyledons. The quadruple mutant *pin1,3,4,7* generally fails to establish apical/basal polarity and displays pleiotropic defects including severe root pole defects, abnormal suspensor divisions and compressed embryos. Depending on the ecotype studied, this mutant combination is either embryo lethal, or gives rise to damaged seedlings with compromised roots and severe apical defects (Friml et al. 2003; Vieten et al. 2005). Some of phenotypes of *pin7* or *pin* multiple mutants are similar to other mutants where either the auxin response or transport was disrupted.

In the context of the embryo, several proteins have been described to be involved in regulating PIN localization. PINs are transported to the plasma membrane via the ER

(Endoplasmatic Reticulum) and their polar distribution is established by endocytic recycling. The internalization of PINs via endocytosis is dependent on auxin, clathrin and the Rab5 GTPase pathway. The impairment of this mechanism influences auxin response and consequently embryo patterning (Dhonukshe et al. 2007; Dhonukshe et al. 2008). To complete PIN recycling, transport from the endosomes to the plasma membrane involves vesicle transport orchestrated by ARF-GEFs (ADP-ribosylation factors and guanine nucleotide exchange factors). In particular GN (GNOM) is responsible for the transport of PINs to the basal part of the membrane (Kleine-Vehn et al. 2008). Interference with GN function, either genetically, or using chemical treatments (like the ARF-GEF inhibitor Brefeldin A (BFA)), results in changes in PIN polarity, auxin distribution and subsequently embryo patterning (Benkova et al. 2003; Geldner et al. 2003; Kleine-Vehn et al. 2008; Steinmann et al. 1999). The phosphorylation status of PINs is also critical to their subcellular localization. The overexpression of the PID (PINOID) serine-threonine kinase causes a switch from basal to apical localization of the PIN1 protein and conversely the loss of PID causes an apical to basal switch (Friml et al. 2004). The role of PID is opposed by PP2A (PROTEIN PHOSPHATASE 2A) which is involved in PIN dephosphorylation, so that phosphorylated PINs are directed to the apical pole whereas dephosphorylated PINs are directed to the basal pole. The perturbation of this mechanism of regulation again changes auxin distribution and embryo patterning (Friml et al. 2004; Michniewicz et al. 2007). The polarization of PIN transporters, mediated by regulated trafficking, is likely to be controlled by multiple internal and external signals, many of which remain to be pinpointed (Grunewald and Friml 2010; Lofke et al. 2013).

1.2.2.2 Auxin response

Responses to auxin are mediated by two different mechanisms. The first, perceiving cytoplasmic auxin, is initiated by auxin binding to the F-box TIR1 (TRANSPORT INHIBITOR RESPONSE 1) receptor, a component of the SCF complex. Auxin binding increases the affinity of TIR1 for a family of small Aux/IAA (INDOLE-3-ACETIC ACID INDUCIBLE) proteins which are consequently ubiquitinated and degraded. Aux/IAAs bind to, and inhibit, ARFs, a family of B3-related transcription

factors. As a result of TIR1 mediated degradation of Aux/IAAs, at higher concentrations of auxin ARFs are therefore released and free to induce or repress the transcription of their target genes.

The central core of auxin response in embryo patterning is regulation of the activity of MP/ARF5 (MONOPTEROS) which is repressed by BDL/IAA12 (BODENLOS) which in turn requires the corepressor TPL (TOPLESS) to inhibit auxin responsive genes (Hamann et al. 2002; Szemenyei et al. 2008). The disruption of the auxin signalling pathway upstream of MP, including auxin synthesis, transport, and sensing by the SCF complex leads to altered activity of MP and hence similar embryo patterning phenotypes to the *mp* mutant (Figure 1.3A). (Reviewed in (Benjamins and Scheres 2008; Mockaitis and Estelle 2008; Oliva et al. 2013; Vanneste and Friml 2009; Weijers and Jurgens 2004; Yoshida et al. 2013))

Mutant alleles of *MP* and gain of function alleles of *BDL* (stabilized Aux/IAA, therefore unresponsive to auxin) have similar phenotypes which include abnormal hypophysis cell divisions and root pole defects that can lead to rootless seedling phenotypes. Cotyledon defects (fused cotyledons) are also observed in these backgrounds, (Berleth and Jurgens 1993; Hamann et al. 1999). PIN transport and auxin distribution (DR5 activity) are altered in *mp* and *bdl* mutant backgrounds, although since MP is an ARF and probably activates the expression of DR5, this disturbance is unsurprising (Friml et al. 2003; Steinmann et al. 1999; Weijers et al. 2006). IAA13 has a similar function to BDL in repression of MP activity (Weijers et al. 2005a). Interestingly MP and BDL are not expressed in the domains which they specify. At early stages of embryogenesis these proteins are expressed in the embryo inner cells adjacent to the uppermost cell of the suspensor or hypophysis and after hypophysis division they are expressed in the apical daughter cell, suggesting that hypophysis specification depends on cell-cell signalling dependent on BDL/MP auxin response (Hamann et al. 2002; Weijers et al. 2005a; Weijers et al. 2006). *PIN1* has a similar expression pattern to *MP* and is misregulated in *mp* mutants suggesting that it might be a target of MP.

TIR1 is functionally redundant with at least three AFB (AUXIN SIGNALING F BOX PROTEIN) proteins expressed in the embryo that have been proved to bind to

auxin and AUX/IAAs (Dharmasiri et al. 2005a; Dharmasiri et al. 2005b; Kepinski and Leyser 2005). In the multiple mutant *tir1afb1,2,3*, the levels of BDL/IAA12 are high and mutants therefore display similar phenotypes to *mp* and *bdl* gain of function mutants with impaired auxin response (Dharmasiri et al. 2005b). Furthermore, loss of CUL1/AXR6 (CULLIN1 or AUXIN RESISTANT 6) function, or interference with related-to-ubiquitin (RUB) posttranslational modification of CUL1 by loss of function of both *RCE1 (RUB1 CONJUGATING ENZYME 1)* and *AXR1*, leads to phenotypes associated with auxin response disruption during embryogenesis (Dharmasiri et al. 2003; Hobbie et al. 2000). AXR1 and the related AXL (AXR1-LIKE) protein act redundantly in the RUB modification of CUL1, and *axr1axl* double mutants show embryo or seedling lethal phenotypes due to severely reduced auxin response and resulting embryo patterning defects (Dharmasiri et al. 2007). The RUB conjugation pathway involves the activity of the heterodimer ECR1 (E1 C-TERMINAL RELATED 1)/AXR1 or ECR1/AXL, the RCE1 and the RUB ligase RBX1 (RING-BOX 1) of the SCF complex (Dharmasiri et al. 2007; Dharmasiri et al. 2003; Gray et al. 2002).

Several genes have been suggested to act downstream of *MP*. Multiple mutant combinations in *PLT (PLETHORA)* transcription factor-encoding genes result in root patterning defects in embryos and *PLT1* and *PLT2* gene expression is lost in *mp* mutants during embryogenesis (Aida et al. 2004). Two pairs of closely related and functionally redundant plant homeodomain finger proteins OBE1/OBE2 (OBERON 1 and 2) and TTA1/TTA2 (TITANIA 1 and 2) involved in chromatin transcriptional regulation, act in the same pathway and are implicated in the MP mediated auxin response required for root meristem initiation. The respective double mutants have similar rootless phenotype to the *mp* mutant (Saiga et al. 2008; Saiga et al. 2012). Two bHLH (basic helix-loop-helix) transcription factor encoding genes, *TMO5* and *TMO7 (TARGET OF MONOPTEROS5* and *7)*, have also been described to be direct targets of MP in embryogenesis. Interestingly TMO7 is able to move to the hypophysis from vascular cells and both *TMO5* and *TMO7* partially restore the rootless phenotype of *mp* (Schlereth et al. 2010).

As mentioned above, cotyledon initiation in embryogenesis is accompanied by the accumulation of auxin maxima in the tips of cotyledon primordia. Consequently, the disruption of auxin signalling in embryos often leads to defects in cotyledon development (Figure 1.3A). The genes *CUC1* and *CUC2* (*CUP-SHAPED COTYLEDON*) are involved in cotyledon development and shoot meristem initiation during embryogenesis. It has been shown that the expression of *CUC1* and *CUC2* in embryos is dependent on auxin, since MP, PIN1 and PID are required for the activity of these genes, although the mechanism is unclear (Aida et al. 2002; Furutani et al. 2004). The redundant transcription factors DRN (DORNOSCHEN) and DRNL (DRN-like) are also involved in cotyledon development. *drn* and *drn drnl* double mutants display cell patterning defects in embryogenesis and changes in auxin (DR5 marker) and PIN distribution. Furthermore *DRN* is believed to be a target of MP, suggesting that DRN acts both upstream and downstream of auxin transport (Chandler et al. 2007; Cole et al. 2009). *DRN* and *DRNL* are expressed in the apical part of the embryo but the double mutants also display basal defects, presumably due to a disruption in auxin transport (Chandler et al. 2007; Cole et al. 2009). A gain of function allele of the apically expressed *IAA18* gene, which stabilizes the IAA18 protein, also displays defects in cotyledon development, again probably due to changes in PIN1 distribution. This mutation increases the occurrence of rootless phenotypes in several auxin mutants (Ploense et al. 2009). Finally MAB2 (MACCHI BOU 2), a component of RNA polymerase transcriptional complexes has been shown to be necessary for the regulation of auxin responsive genes specifically in the apical part of the embryo. The auxin response in *mab2* mutants is altered as the DR5 reporter is expressed only at basal pole of the embryo and not in the cotyledon primordia. *mab2* embryos display abnormal cell divisions at early stages of embryogenesis and defects in cotyledon development that can lead to embryo development arrest (Ito et al. 2011).

Although much of the information regarding auxin responses is connected to MP, several other *ARF* genes have been shown to be strongly expressed during embryogenesis (Rademacher et al. 2011), suggesting that auxin response regulation is more complex. Consistent with this view several genetic interactions between ARFs imply gene redundancy. The percentage of rootless phenotypes in *mp arf1*

mutants is lower than *mp* single mutant, but the same phenotype in *mp arf6* embryos is more frequent than for *mp*. These results are consistent with the fact that ARF5 and ARF6 act redundantly. The *arf1,2,6* triple mutant is male sterile and displays abnormal cell divisions at the embryo/suspensor boundary (Rademacher et al. 2011).

In summary, auxin signalling is important for root meristem and cotyledon development in embryogenesis. Embryo patterning involves the determination of specific domains, precursors to different tissues. Auxin is involved in the establishment of these domains together with transcription regulation programs (discussed below) and several signalling pathways (again discussed below). These three keystones are intricately connected. The disruption of any one process disturbs the others and consequently it is often unclear what the origin of embryo patterning defects is. Probably the key genes involved in embryo patterning have already been identified, but the study of downstream and upstream interactions is required to understand the mechanisms behind embryo patterning. In the next sections I will describe separately our current understanding of the transcriptional regulation of embryo development and of signalling pathways involved in embryo patterning.

1.2.3 The WOX family defines apical/basal cell fate

In *Arabidopsis* the WOX (WUS RELATED HOMEOBOX) family of homeodomain transcription factors is composed of 15 proteins related to WUS (WUSCHEL), that are expressed at different stages of plant development and which are considered to have biological activities related to cell patterning and cell identity (Haecker et al. 2004; van der Graaff et al. 2009). The expression patterns of several *WOX* genes in early embryogenesis are spatially and temporally dynamic and as a result these genes are thought to be good candidates for roles in the definition of different embryonic cell fates. It has been shown that WUS binds to two different regulatory sequences; WUS binding site (TTAATSS) and a G box sequence (CACGTG) supported by two independent studies and using different methods (Busch et al. 2010; Lohmann et al. 2001). It is possible that other WOX proteins share similar binding sites, as it has been shown for the members of other homeodomain families.

The dynamic expression patterns of *WOX* genes during early steps of embryogenesis are important for the establishment of the apical/basal axis of the embryo and subsequently shoot/root cell fate (Figure 1.4A). Initially, *WOX2* and *WOX8* are expressed in the egg cell and zygote. After the first division the expression of *WOX2* and *WOX8* resolves into apical and basal expression domains respectively. *WOX9/STIP* is also expressed in the basal cell after this first division. From this point onwards *WOX2* is restricted to the apical cell lineage and *WOX8* to the basal cell lineage. *WOX9* has a more restricted expression domain than *WOX8* and remains restricted to the uppermost suspensor cell, the hypophysis until the third division of the embryo proper, when its expression spreads into the lower of the two tiers of cells generated by this division. This spreading of *WOX9* expression correlates with the restriction of *WOX2* expression to the upper tier of cells in the embryo proper. Thus in the 8 cell embryo proper the apical and basal domains of the embryo are characterized by *WOX2* and *WOX9* expression respectively (Haecker et al. 2004).

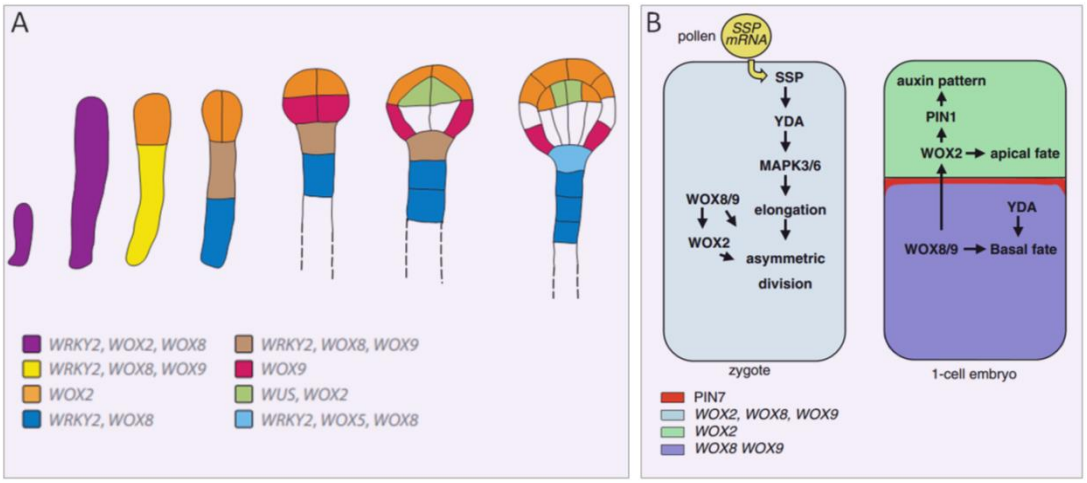


Figure 1.4 –The WOX protein family in embryogenesis. A) Expression pattern of different *WOX* genes and *WRKY2*, B) Predicted mechanisms required for asymmetric division of the zygote. After fertilization SSP is delivered to the zygote and activates the YDA signalling cascade that, together with WOX2, WOX8 and WOX9, regulates zygote elongation and asymmetric division. After the first division, YDA together with WOX8 and WOX9 regulate basal cell fate. Additionally WOX8 and WOX9 are required for the expression of *WOX2* in the apical cell, which in turn is necessary for *PIN1* expression and consequently auxin response and apical cell fate. (Adapted from (Lau et al. 2012; Zhang and Laux 2011))

Radial patterning is initiated at the dermatogen stage and *WUS* is activated in the four apical inner cells, *WOX9* expression is restricted to the basal cells of the

epidermal layer and hypophysis whereas the expression patterns of *WOX8* and *WOX2* are maintained. Interestingly, at the early globular stage *WOX5* is activated in the hypophysis and after the asymmetric division *WOX5* is expressed in the small apical daughter cell, precursor of the root quiescent centre. In parallel the expression of *WUS* is progressively restricted to the apical centre cells adjacent to the epidermis, which are the precursors of the shoot apical meristem. At the globular stage *WOX8* is expressed in the suspensor cells, *WOX9* is expressed in the basal epidermal cells, and *WOX5* is only expressed in the hypophysis. At later stages of development *WOX1*, *WOX3* and *WOX5* are expressed in different domains in the cotyledon primordia, *WUS* is expressed in the shoot apical meristem, *WOX5* in the root quiescent cell lineage, *WOX8* in the suspensor cells, and *WOX9* is restricted to the epidermis in central/basal domain. Interestingly *WOX2* expression in the apical region of the embryo is lost at the globular stage and heart stage but is detected in the region of the epidermis corresponding to the boundary between the hypocotyl and root at later stages (Haecker et al. 2004; Mayer et al. 1998; Sarkar et al. 2007). It should be noted that an independent study of *WOX9* indicates that this gene as a broader expression pattern in embryogenesis (Wu et al. 2007). In this study *WOX9* transcript was detected by *in situ* hybridization after the first asymmetric division of the zygote in both the apical and basal cell, and subsequently expressed in the whole embryo proper and hypophysis. With cotyledon initiation, *WOX9* expression became focused in cotyledon primordia and epidermal cells of the basal zone of the embryo (Wu et al. 2007). These results are consistent with the expression of a *WOX9-GFP* construct that complemented the *stip-2* mutant (Wu et al. 2007).

The expression patterns of *WOXs* are correlated to their functions in embryo patterning. *WOX2* is required for apical cell patterning and a proportion of *wox2* embryos show abnormal cell divisions in this region at early stages of plant development. In particular horizontal rather than vertical division planes have been observed during the first or second embryonic division. In addition, at the dermatogen stage, the division planes of the cells in the apical part of the embryo are also abnormally placed, retarding the definition of an epidermal cell population. Finally, a proportion of *wox2* seedlings displays cotyledon defects including the production of only one cotyledon (Breuninger et al. 2008; Haecker et al. 2004).

Interestingly these phenotypes are more frequent in multiple mutant combinations between *wox2* and *wox1*, *wox5*, *wox8* and/or *prs (wox3)* suggesting functional redundancy especially at the level of seedling phenotypes which can be much more severe than in *wox2* single mutants (Breuninger et al. 2008). Nevertheless, these phenotypes are not found in *wox1 wox5 prs* triple mutants indicating that *WOX2* has a central role in apical embryo patterning. Somewhat counterintuitively, the seedling phenotypes of the *wox2wox8* double mutant suggest that *WOX8* acts redundantly with *WOX2* in the promotion of apical cell fate, and that *WOX8* expression in the basal lineage is therefore required for *WOX2* function in embryo patterning (Breuninger et al. 2008; Wu et al. 2007). The role of *WOX8/WOX2* in apical cell fate needs to be clarified since in theory *WOX8,* expressed in basal cells, should be involved in basal cell fate. Most likely, the phenotypes of *wox2 wox8* mutants are due to problems in cell fate determination during the zygote asymmetric division rather than *WOX8* influencing *WOX2* apical expression at later stages.

WOX9 loss of function mutants display inconsistent phenotypes. While the *wox9-1* phenotype is similar to the wild-type, the *stimpy (stip)* alleles of *WOX9* (where STIMPY is WOX9) display embryo phenotypes with different degrees of severity (Breuninger et al. 2008; Wu et al. 2007). The weakest stip phenotype corresponds to elongation defects in *stip-1* embryos which expand horizontally at late stages of embryogenesis. *stip-2* and *stip-3* embryos arrest at early stages of embryogenesis displaying abnormal cell divisions in the embryo or suspensor. While a few *stip-2* seeds are able to germinate, no homozygous *stip-3* seedlings were identified (Wu et al. 2007; Wu et al. 2005). The differences of phenotypes could be attributed to the different transcripts produced in the different mutant alleles. For instance *wox9-1* expresses a small N-terminal transcript, whilst *stip-2* and *stip-3* result in a stop codon in the Homeodomain (HD) encoding region and nonsense mutation respectively that could interfere with protein function. *WOX8* loss of function in *stipl-1*, *wox8-1* or *wox8-2* mutants, gives no visible phenotype (Breuninger et al. 2008; Wu et al. 2007). *wox8-1wox9-1* or *stip-1stipl-1* double mutant embryos show abnormalities in early cell divisions that can lead to embryo lethality suggesting that these genes act partially redundantly in basal and apical cell lineage determination (Breuninger et al. 2008; Wu et al. 2007). Firstly the apical cell of *wox8-1wox9-1* one-cell stage

embryos divides horizontally rather than vertically. In addition, a proportion of suspensors show abnormal divisions or large misshapen cells, and neither the embryo nor suspensor of *wox8wox9* embryos develops normally at later stages (Breuninger et al. 2008; Wu et al. 2007). The distribution of auxin and PIN proteins is disturbed in *wox9wox8* double mutants. PIN1 is absent in globular stage embryos and the DR5 marker is expressed throughout the whole embryo and suspensor. Interestingly, the expression of *WOX2* under the *WOX9* promoter in this double mutant partially rescues some apical phenotypes, auxin patterning and PIN1 distribution (Breuninger et al. 2008; Wu et al. 2007). In addition, the phenotypes of *wox8 mp* and *wox2 wox8 mp* mutants are more severe than the single *wox8* mutants or the *wox2 wox8* double mutant, and consequently the expression of *PIN1* and activity of DR5 is compromised (Breuninger et al. 2008; Haecker et al. 2004). However, *WOX8* expression in early embryogenesis was unaffected by auxin treatment (Ueda et al. 2011). Taken together these results suggest that *WOX* regulation and auxin responses are connected to promote embryo patterning, although the mechanism is unclear.

After fertilization, the zygote elongates following the apical/basal axis already defined in the egg cell. The elongation of the zygote involves different signalling inputs and regulation events that are important for cell polarity and consequently the first asymmetric division of the zygote that separates the apical and basal lineages (reviewed in (Zhang and Laux 2011)). The apical/basal identity fixed in the first division of the zygote (Figure 1.4B) is dependent on signalling by the MAP kinase YDA (YODA) and *WOX (WOX2, WOX8, WOX9)* regulation (Breuninger et al. 2008; Lukowitz et al. 2004; Zhang and Laux 2011). The GRD (GROUNDED) transcription factor and SSP (SHORT SUSPENSOR) kinase are required upstream of YDA (Bayer et al. 2009; Jeong et al. 2011). SSP is produced in mature pollen and delivered to the zygote during fertilization. The signals transduced by SSP kinase then trigger the YDA pathway (Bayer et al. 2009). In parallel, GRD acts together with WOX proteins, upstream of YDA signalling, to establish the zygote asymmetric division. Consequently the triple mutant *grd wox8 wox9* arrests at the one cell or zygote stage due to extreme polarity problems (Jeong et al. 2011). The transcription factor WRKY2 (WRK DNA-BINDING PROTEIN2) also plays a role in zygote polarization, probably by activating the transcription of *WOX8* and *WOX9* (Ueda et

al. 2011). The loss of function of *WRKY2* disrupts zygote polarity but not egg polarity, and consequently the uppermost cells of the basal lineage show apical cell fate behaviour, displaying embryo-like division patterns, although later, at the heart stage, *wrky2* embryos recover and germinate normally. Interestingly the abnormal group of cells in the suspensor/embryo boundary of *wrky2* embryos lose the expression of the *AGO10* (*ARGONAUTE 10*) apical marker and *WOX8* basal marker but keep some basal identity, namely the expression of the *ARR5* basal lineage marker (Ueda et al. 2011).

As discussed above, auxin responses are also important for cell polarity and asymmetric division. Auxin is accumulated in the apical daughter cell and this process is dependent on specific localization of PIN7 in the basal cell (Friml et al. 2003). Together, WOX2, WOX8 and WOX9 specify both apical and basal cell identity, and subsequently embryo cell patterning (Figure 1.4). WOX regulation and auxin response are clearly interrelated. It will be interesting to address the relationships between PIN1/PIN7, WOX proteins and the YDA pathway in the first division of the zygote in order to understand fully how cell polarity is established and consequently how the asymmetric division of the zygote is regulated. Only a few genetic relationships have been established, but the mechanism remains poorly understood. The identification of WOX targets, novel components of the YDA pathway and components of auxin responses in the zygote will help in understanding the mechanisms of establishment of embryo polarity.

1.2.4 HD-ZIP IV proteins play a vital role in epidermal layer design

As discussed above, embryo patterning requires a specific sequence of cell divisions that is controlled by multiple regulatory programs. A complex network of regulatory gene expression patterns generates a robust program that defines the main structures of the plant. The epidermal layer is defined at early stages of embryogenesis, being distinct from the dermatogen stage onwards (Figure 1.2 D). From this point it is generally considered that basal epidermal cell fate simply needs to be maintained in outer cells throughout the rest of plant development (reviewed in (Ingram 2007; Javelle et al. 2011b))

Although several genes have been shown to play an important role in epidermal regulation and signalling, the pathways responsible for epidermal cell fate specification remain unclear. The HD-ZIP IV (Homeodomain leucine zipper IV) family of transcription factors are generally expressed in the epidermal cell layer and have been described in several different species (Javelle et al. 2011a; Nakamura et al. 2006). In particular two redundant members of this family, PDF2 (PROTODERMAL FACTOR2) and ATML1 (ARABIDOPSIS THALIANA MERISTEM LAYER 1) have a vital role in epidermis determination.

In the literature, the expression patterns of *ATML1* and *PDF2* (Figure 1.2 and 1.5) are described by *in situ* hybridization and by studies of the *ATML1* promoter. Both transcription factors are expressed in the embryo at early stages of embryogenesis and at the dermatogen stage their expression becomes restricted to the outermost cell layer. This epidermis specific pattern is maintained until late stages of embryo development, and subsequently in the outer cell layer of meristems and organs (Abe et al. 2003; Lu et al. 1996; Sessions et al. 1999; Takada and Jurgens 2007). *PDF2* expression was detected from the 4-cell stage onwards (Abe et al. 2003), but may be expressed earlier, whereas *ATML1* transcripts were detected continuously in the apical cell and its progeny after zygote division until the 8-cell stage being absent in the suspensor and hypophysis during embryogenesis (Lu et al. 1996). In contrast the *ATML1* promoter was reported to be additionally expressed in the basal cell and basal lineage including hypophysis and suspensor (Tanaka et al. 2007). It is possible that the RNAs of these genes are regulated post transcriptionally, explaining this discrepancy.

The expression of *PDF2* driven by a 35S promoter resulted in late flowering phenotypes and flower morphology phenotypes in lines with reduced levels of *PDF2* transcript (silenced lines), but no embryo phenotypes were described (Abe et al. 2003). Recently, the inducible overexpression of *ATML1* under a ribosomal promoter was shown to lead to different seedling phenotypes including seedling lethality depending on the *ATML1* expression levels and duration of the induction (Takada et al. 2013). Unfortunately, the authors focused largely on post germination phenotypes

(described in Section 1.4) and therefore the consequences of overexpression of ATML1 in embryo patterning are unclear.

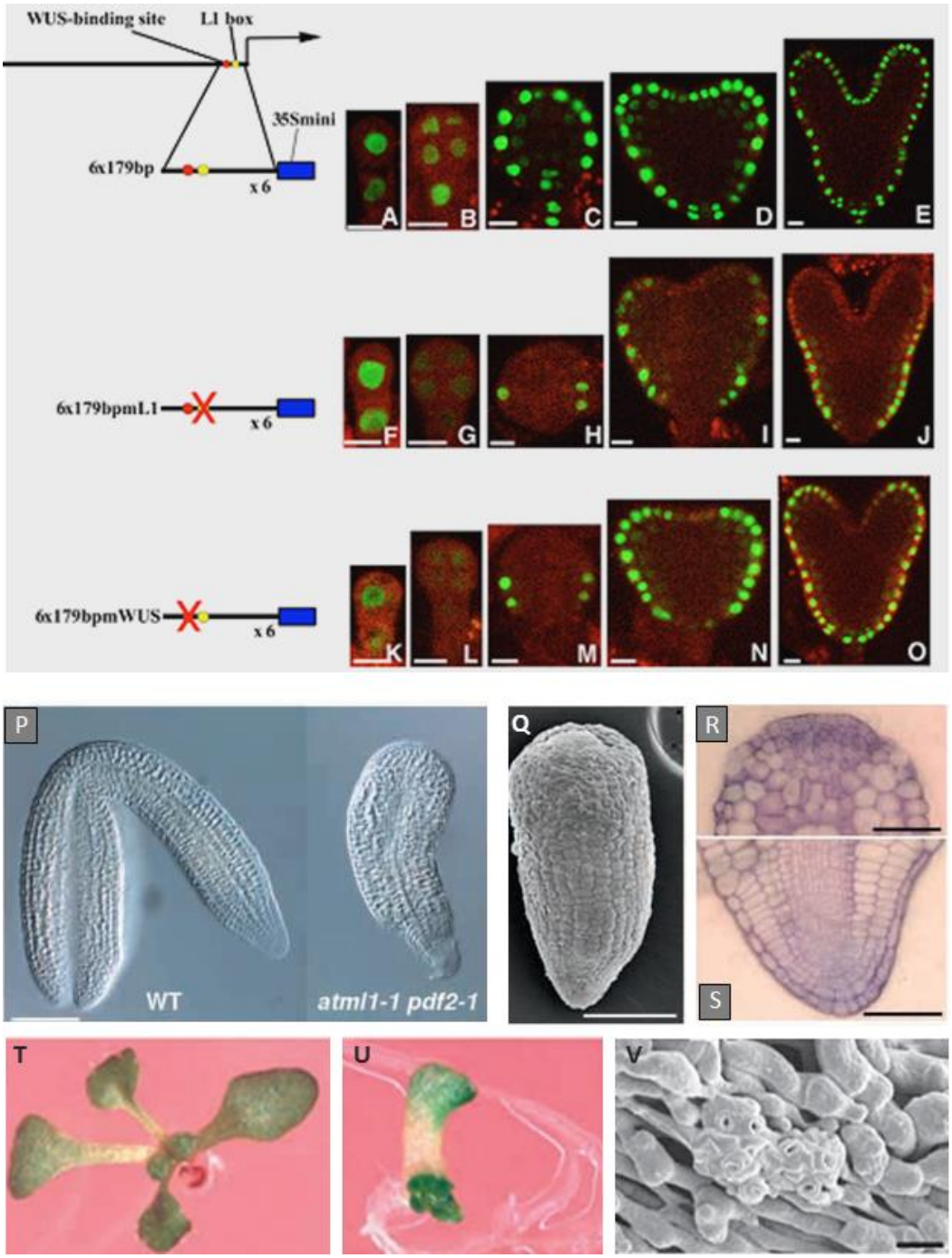


Figure 1.5 – ATML1 and PDF2 are involved in specifying epidermal cell fate. (A-0) Show GFP expression in embryos at different stages of embryogenesis driven by three versions of the *ATML1* promoter. Expression profile of the *ATML1* promoter (A-E), L1 box mutated version of the *ATML1* promoter (F-J) and WUS box mutated version of the *ATML1* promoter (K-0) during embryogenesis; (P-V) show the embryo and seedling phenotypes of wild-type (P) and *atml1-1pdf2-1* double mutants (Q-S, U, V). (Taken from (Abe et al. 2003; Takada and Jurgens 2007))

The study of independent *atml1* and *pdf2* mutant lines has shown that the plant development is comparable to that of wild-type suggesting functional redundancy, although weak sepal phenotypes have been attributed to these single mutants, which lack giant cells in sepals in the L*er* (Landsberg *erecta*) background (Abe et al. 2003; Roeder et al. 2012). The double mutant *atml1-1pdf2-1* has been reported to display apical embryo patterning defects and strong seedling phenotypes (Figure 1.5), when germinated the double mutant seedling lacked the epidermal layer with only local clusters of differentiated epidermal cells. These defects lead to seedling lethality (Abe et al. 2003). The analysis of transcript levels of *atml1-1pdf2-1* double mutant seedlings suggested that these transcription factors positively regulate *ACR4* (*ARABIDOPSIS CRINKLY 4*) which encodes a receptor kinase and *PDF1* (*PROTODERMAL FACTOR1*), which encodes a secreted protein, both of which are expressed specifically in the L1 layer of the embryo (Abe et al. 2003; Abe et al. 1999; Gifford et al. 2003; Tanaka et al. 2002). This regulation is supported by the fact that both ATML1 and PDF2 have been proven to bind *in vitro* to the L1 box motif (TAAATGYA) which has been found in the promoter of several L1 specific genes including *ACR4* and *PDF1* (Abe et al. 2003; Abe et al. 2001). Together, these results indicate that *ATML1* and *PDF2* are redundantly required for epidermal specification during embryogenesis. The transcriptional activity of these two genes suggest that they act upstream of the genes required for epidermal cell identity, possibly by binding to the L1 box promoter element.

Besides *ATML1* and *PDF2* very few genes have been shown to play a significant role in epidermis specification during embryogenesis and their interactions with *ATML1* and *PDF2* are unclear. The genes upstream and downstream of *ATML1* and *PDF2* are unknown and consequently the mechanism of specification of the L1 layer is yet to be determined. One hypothesis regarding the regulation of *ATML1* and *PDF2* includes positive autoregulation, with their epidermis specific expression pattern attributed to their default expression at early stages of embryogenesis and a subsequent repression by an unidentified transcription factor in inner cell layers (Abe et al. 2003). Alternatively these two genes could be activated by a signalling pathway involving signals originated either within or from outside the embryo, possibly involving the ESR of the endosperm, or the embryonic cuticle layer, which are only

perceptible in the outermost cells of the embryo. In this model the inner cells would lose the signal and stop expressing these two genes at the dermatogen stage. This second hypothesis is included in the current accepted model of a positive feedback loop consisting of the positive autoregulation of *ATML1* and *PDF2* which also activate genes involved in protodermal differentiation, and an additional positive regulation of *ATML1* and *PDF2* via signalling pathways (Section 1.2.6) that include the ACR4 receptor kinase which is also a potential target of ATML1/PDF2 (Abe et al. 2003; Javelle et al. 2011b; Tanaka et al. 2007).

Expression analysis of the *ATML1* promoter during embryogenesis is consistent with the positive feedback loop model to the detriment of the repressor model (Takada and Jurgens 2007). This study includes the analysis of different regions of the *ATML1* promoter fused to a nuclear multi GFP, focusing in on a minimal region containing an L1 box and a potential WUS binding site (TTAATSS). Important conclusions arise from this study. Firstly, the authors identified a sequence of 101 bp (Figure 1.4 A-E), containing both the above binding sites, which was capable of driving marker expression in a manner comparable to the *ATML1* full length promoter, and was required for the expression in the one cell embryo and apical region of the embryo. The deletion or mutation of either of these two binding sites (Figure 1.5 F-O), mimicking the role of a potential repressor, is not sufficient to abolish the full L1 expression pattern. However, mutated versions of this fragment can no longer mimic the entire expression pattern of the full length promoter, with the exception of early stages of embryo development before the epidermis is specified. This result indicates that both the L1 box and potential WUS binding site are necessary for *ATML1* expression, and more importantly, that transcription factors other than ATML1 and PDF2 are likely required for the *ATML1* transcription. Analysis of different regions of the promoter allowed the generation of different spatial and temporal expression patterns attributable to different combinations of *ATML1* promoter regions, and provided evidence for at least three regulatory mechanisms corresponding to initial activation, maintenance of expression and reactivation at later stages of embryogenesis (Takada and Jurgens 2007). The hypothesis of an *ATML1/PDF2* repressor is inconsistent with these results, not least because no promoter variants gave significant ectopic gene expression. Thus the

regulation of *ATML1* is complex at the level of transcription, and other mechanisms involving positional signalling are likely required to ensure the L1 specific expression pattern observed in embryos. Moreover the intricate regulation of *ATML1* suggests that simple autoregulation is likely too simplistic and is not sufficient to activate these genes. Amongst potential regulators of *ATML1* are members of the WOX family, which could potentially bind to the WUS binding sites and clearly cannot be ignored (Section 1.2.3).

A number of genes from the *HD-ZIP IV* family in both *Arabidopsis* and maize (*Zea mays*) are not only expressed specifically in epidermis, but have important roles in regulating epidermal cell fate post germination (Javelle et al. 2011a; Nakamura et al. 2006). In the context of embryogenesis, in *Arabidopsis* the genes *HDG2*, *5*, *7*, *8*, *11*, and *12* (*HOMEODOMAIN GLABROUS*) have, like *ATML1* and *PDF2* been shown to be expressed in the embryo and most of these genes have been shown to be expressed specifically in the epidermis of inflorescences and seedlings (Nakamura et al. 2006). However, their roles in embryogenesis are unknown. The likely genetic redundancy in this family of transcription factors may explain this lack of functional information. In the context of embryo patterning only *ATML1* and *PDF2* have been proven to be involved in epidermal specification, nevertheless the study of different mutant combinations and overexpression lines of *HD-ZIP IV* genes supports the theory other HD-ZIP IV transcription factors work together with ATML1 and PDF2 in regulating epidermal cell fate decisions post germination, and the same is likely to be true during embryogenesis. The description of HD-ZIP IV family functions post germination can be found in more detail in Section 1.4.

Similar to *ATML1* or *PDF2*, the maize *HD-ZIP IV* gene *OCL1* (*OUTER CELL LAYER1*), is expressed at early stages of embryogenesis and its expression is progressively restricted to the L1 layer as the embryo develops. Additionally OCL1 is also expressed in the epidermis of root and shoot meristems, developing leaves and floral organs (Ingram et al. 1999). The analysis of transgenic lines overexpressing *OCL1* identified several potential targets of this transcription factor related to plant defence, cuticle biosynthesis and lipid metabolism (Javelle et al. 2010). In contrast to the uniform epidermal expression of *OCL1*, the transcript expression patterns of the

HD-ZIP IVs OCL3, *OCL4* and *OCL5* in embryogenesis are restricted to different regions of the embryo epidermis. After germination these genes are expressed in the L1 layer of shoot meristems and young flowers, again in specific domains (Ingram et al. 2000). The analysis of the transcript levels of the 17 *HD-ZIP IV* maize genes indicates that these genes expressed more strongly in reproductive organs rather than vegetative tissues, and that most genes are expressed in the kernel with different expression levels. The most strongly expressed in kernels are *ZmHDZIV9*, *11* and 16. The analysis of the expression levels of this family in the L1 layer and inner layers of young leaves and meristems show that, as in *Arabidopsis*, most of these genes are strongly expressed in epidermis compared to inner layers, and several are specifically expressed in the L1 layer (Javelle et al. 2011a).

It has been proposed that *OCL1* has a similar role to *ATML1* and *PDF2* and that possibly other *HD-ZIP IV* genes act redundantly with *OCL1* in epidermis specification in maize.

1.2.5 Embryo patterning regulatory network

The definition of different regions that are precursors to different tissues is important for embryo patterning and consequently plant development. The determination of root and shoot apical meristems in embryo in particular, is a key step in the definition of the plant body. Interestingly, some of the genes involved in the definition of the root pole and shoot meristem in embryogenesis maintain similar roles throughout plant development.

Embryogenesis is characterized by the sequential determination of specific patterns. The first is the determination of apical/basal axis followed by the radial patterning and the adaxial/abaxial cell fate defined along with cotyledon development. In this field the knowledge of transcriptional regulation is significantly greater than knowledge of signalling pathways, which are poorly understood. Nevertheless positional signalling is essential for the establishment of different cell fates. The study of the expression patterns of receptor-like kinases (RLKs) in the embryo suggests that RLKs could play an important role in intercellular signalling during embryogenesis (reviewed in (Nodine et al. 2011)). In this section I will try to

describe the main regulators of embryo patterning focusing on shoot and root meristem patterning and apical/basal cell fate summarized in Figure 1.6 (reviewed in (Jenik et al. 2007; Lau et al. 2010; Lau et al. 2012)).

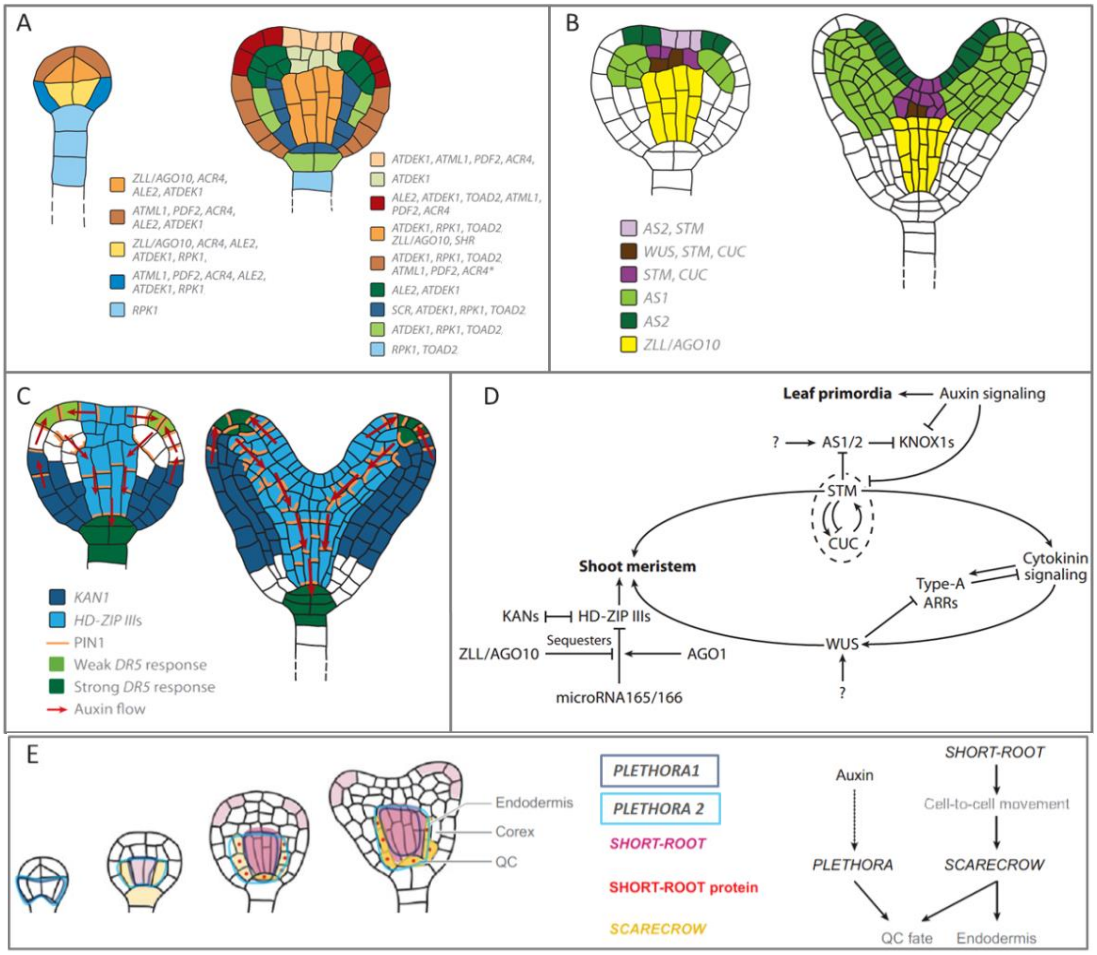


Figure 1.6 – Embryo patterning regulation. A) Regulation of radial patterning from dermatogen to late globular stage, B) Formation of the shoot meristem and cotyledon initiation from globular to heart stage, C) *KAN1* and *HD-ZIP III* expression patterns (example for *REV*) in peripheral and central domains respectively, correlated with auxin response in the embryo, D) Predicted mechanism for SAM and organ initiation regulation, and E) Regulation of root pole patterning. (Adapted from (Jenik et al. 2007; Lau et al. 2012))

1.2.6.1 The importance of radial patterning

As mentioned before, the radial patterning is initiated at the dermatogen stage when the protoderm is differentiated from the inner cell layers. At this stage *ATML1/PDF2* are expressed in the L1 layer and *AGO10* in the inner cells. Later, at the globular stage, *AGO10* is expressed in the provascular region, overlapping with *SHR* (*SHORT ROOT*), *PLT1* and *PLT2* (Aida et al. 2004; Helariutta et al. 2000; Lynn et al. 1999);

two transcription factors involved in basal cell fate (Figure 1.6 A, B, and E). Epidermal specification and the role of auxin in embryo patterning have already been described in previous sections. Nevertheless it is important to highlight the interactions between auxin signalling and several key regulators of embryo patterning, in particular during root pole formation and cotyledon initiation (Figure 1.6 C-E).

The organization of the embryonic root (Figure 1.6 E) involves the activation of *SHR* in the provascular cells and is probably followed by the transport of SHR to the cells of the ground tissue adjacent to the provascular domain, where it activates the expression of *SCR (SCARECROW)* which is involved in specifying endodermal cell fate (Nakajima et al. 2001; Wysocka-Diller et al. 2000). In the lens shaped cell of the hypophysis, SCR together with PLT transcription factors and ARFs promote quiescent centre formation and, in doing so, merge radial and apical/basal patterning (Aida et al. 2004; Yoshida et al. 2013). *WOX5*, expressed in the root quiescent centre and its embryonic precursors, is involved in stem cell fate maintenance in the root stem cell niche (Haecker et al. 2004; Sarkar et al. 2007). In addition, WOX5 is involved in QC specification together with auxin and SHR/SCR regulation but almost independently of PLT1/PLT2. Interestingly, *WUS* and *WOX5* have similar functions in shoot and root pole stem cell niche respectively, and the expression of each of these genes under *WOX5* or *WUS* promoters rescue *wus* and *wox5* mutant phenotypes correspondingly (Sarkar et al. 2007).

The loss of function of the co-repressor TPL causes the transformation of the shoot pole into a root pole. Several shoot markers are not expressed in the *tpl* mutant and, in contrast, the expression of root markers can be duplicated (Long et al. 2002). Both *PLT1* and *PLT2* have been identified as targets of TPL, and have been proven to be necessary for the conversion of apical to basal cell fate (Smith and Long 2010). The same authors showed that HD-ZIP III transcription factors (notably REVOLUTA (REV) and PHABULOSA (PHB)) are sufficient to drive the conversion of roots to shoots in embryos in an antagonistic relationship with PLT function, since *rev plt1 plt2* mutants display a double seedling phenotype and *phb plt1 plt2* seedlings lack roots. Therefore HD-ZIP IIIs not only function in the promotion of apical cell fate

but also by antagonising basal cell date (Smith and Long 2010). Interestingly, the loss of function of *HAN* (*HANABA TARANU*), a gene encoding a GATA transcription factor, causes defects in embryo basal patterning, and consequently a coordinated apical shift of *WOX5* and *SHR* expression patterns and auxin distribution, suggesting that this gene is important to define the basal region of the embryo and in particular the embryo/suspensor boundary (Nawy et al. 2010).

In summary, radial patterning is important for the establishment of regions corresponding to specific gene expression zones that are required for the regulation of apical/basal identity and meristem specification. The formation of the root pole is dependent on the action of auxin signalling, basal cell fate and specific regulatory patterns.

1.2.6.2 Regulation of SAM and organ initiation in embryogenesis

The HD-ZIP III protein family is implicated in several aspects of plant development, playing roles in apical cell fate regulation, meristem regulation, vascular development and organ polarity (Prigge et al. 2005). Different combinations *of HD-ZIP III* mutants show different embryonic and post germination developmental phenotypes (Prigge et al. 2005). With the exception of *ATHB8*, the other four *HD-ZIP III* genes *REV*, *PHB*, *PHV* (*PHAVOLUTA)* and *CNA* (*CORONA*) are expressed in the apical region of globular embryos, and later *PHV* expression is restricted to the adaxial region of cotyledons while *PHB* and *REV* are expressed in the SAM region (Emery et al. 2003; McConnell et al. 2001; Otsuga et al. 2001; Prigge et al. 2005). Interestingly, the central apical region of *rev phb phv* and *rev phb cna* embryos structurally resembles the cotyledon primordium although hypocotyl and root regions display a regular development (Emery et al. 2003; Prigge et al. 2005). This phenotype is distinct from auxin related monocotyledon phenotypes or disruption of polar auxin transport and the CUC/STM (CUP-SHAPED COTYLEDON/ SHOOTMERISTEMLESS) pathway, namely *pin1cuc1cuc2* or *pin1stm* mutants (Aida et al. 2002; Moller and Weijers 2009). The comparison of phenotypes suggests that *HD-ZIP III* genes are involved in the differentiation of central and peripheral regions of the meristem. In the absence of HD-ZIP III function, all cells in the apical part of the embryo display peripheral identity and consequently the structure of

abaxialized cotyledons (Prigge et al. 2005). Auxin related mutants rarely lose peripheral and central identity.

There may be an antagonistic relationship (Figure 1.6) between HD-ZIP III regulation in the central domain and regulation via *KAN* (*KANADI*) and *YABBY* genes expressed in the peripheral zone and involved in abaxial cell fate. This relationship, which determines the adaxial/abaxial differentiation of organs, is maintained post embryonically in lateral organ formation (Emery et al. 2003; Prigge et al. 2005; Siegfried et al. 1999). Additionally, AGO1 (ARGONAUTE 1) and AGO10 act antagonistically in the regulation of *HD-ZIP III* (Figure 1.6). These proteins bind regulatory microRNAs, and AGO10 has been shown to positively regulate HD-ZIP IIIs by sequestering miRNA165/166 which would otherwise repress HD-ZIP III production, thereby promoting shoot meristem specification (Zhu et al. 2011). It has also been suggested that HD-ZIP IIIs are involved in SAM specification and organ formation in parallel to the CLV/WUS (CLAVATA/WUSCHEL) pathway. In particular *CNA* may act in parallel with the CLV pathway, since in *clv* mutant backgrounds the loss of *CNA* function leads to defects in organogenesis, meristem size and stem cell maintenance with impaired *WUS* and *CLV3* expression patterns (Green et al. 2005).

The *CUC* genes together with *STM, AS1* (*ASYMMETRIC LEAVES1*) and auxin response, are involved in the specification of the SAM and cotyledon separation (Figure 1.6). *CUC* genes are required for cotyledon separation and promote the expression of *STM* in the shoot meristem. Conversely STM negatively regulates *CUC* and *AS1*. AS1 promotes cell differentiation while STM promotes stem cell fate (Aida et al. 1999; Aida et al. 2002; Byrne et al. 2002; Hibara et al. 2006). STM also induces cytokinin biosynthesis, and in turn cytokinin signalling induces *WUS*, the master regulator of the SAM organizing centre, which plays a key role in SAM maintenance via the action of CLV1 and CLV3, and together with *STM* acts as a repressor of cell differentiation (Gordon et al. 2009; Lenhard et al. 2002; Yanai et al. 2005).

The *HD-ZIP II* family has also been implicated in embryo patterning, in particular *ATHB2, HAT3* and *ATHB4* (*HOMEOBOX-LEUCINE ZIPPER PROTEIN 2*, *3* and *4*

respectively). These three genes act redundantly in the regulation of SAM formation and the specification of adaxial identity in lateral organs (Turchi et al. 2013). The loss of function of both ATBH4 and HAT3, which are coexpressed during embryogenesis, results in defects in cotyledon development including vascular patterning and altered auxin distribution and response. The phenotype of the *hat3athb4* double mutant is enhanced by loss of *ATHB2* function or loss of either *PHV* or *REV*. Moreover *ATHB2* is a direct target of *REV* (Turchi et al. 2013).

It can be seen from the summaries above, that HD-ZIP transcription factors have a strategic role in embryo patterning. *HD-ZIP IV* genes are involved in epidermal determination, *HD-ZIP III* genes control the central domains of embryo patterning and both gene families are required for shoot apical meristem determination and maintenance. Additionally, *HD-ZIP II* and *HD-ZIP III* genes act together in the establishment of bilateral symmetry in embryogenesis and regulation of the SAM. It is possible that other interactions between *HD-ZIP* genes of different classes occur in embryo patterning suggesting that the network of HD-ZIP-mediated transcriptional regulation in embryogenesis is complex and robust due to functional redundancy between members of these families. Finally, genetic analysis suggests that there is an important link between HD-ZIP and WOX homeodomain protein-mediated transcriptional regulation that should be explored further.

1.2.6 Signalling pathways involved in epidermal specification

Cell-cell communication is essential in embryo patterning and in cell fate decisions more generally in plants. In this section, I will describe the most relevant genes involved in signalling during embryogenesis with a focus on epidermis specification and differentiation. In particular three signalling cascades have been proposed to be potentially situated upstream of the transcriptional activation of *ATML1* and *PDF2*. These involve 1) the protease AtDEK1 (DEFECTIVE KERNEL1), 2) the receptor kinases ACR4 (Arabidopsis CRINKLY4) and ALE2 (ABNORMAL LEAF SHAPE 2) and 3) the receptor kinases GSO1 and GSO2 (GASSHO1/GASSHO2) (Javelle et al. 2011b; Tanaka et al. 2007; Tsuwamoto et al. 2008).

1.2.6.1 Is DEK1 responsible for the first responses in epidermis specification?

The quest of finding the initial positional signalling events required for epidermal specification in embryogenesis has led to the identification of several candidates, linked by unattractive black boxes signalling the presence of unknown substrates, targets, ligands, and regulatory interactions. One of these potential candidates is the DEK1 protein, which plays a yet unidentified role in regulating epidermis identity. Intriguingly, this gene is expressed constitutively whereas many genes proven to be involved in epidermis identity are specifically expressed in the epidermis.

The DEK1 protein contains a calpain-like cysteine proteinase domain in its C terminal, attached to a membrane anchor that contains around 21 transmembrane domains broken by a large predicted extracellular loop. The membrane anchor and calpain domain are separated by a third cytoplasmic domain (Lid et al. 2002; Wang et al. 2003a). The mechanisms underlying DEK1 activity are still unclear. However, this protein is localized in membranes and can undergo autolytic cleavage followed by the release of the calpain domain into the cytoplasm (Johnson et al. 2008).

In both *Arabidopsis* and maize, the loss of *DEK1* function leads to endosperm defects and embryo lethal phenotype at the globular stage due to cell patterning problems in both embryo and suspensor. Both tissues display abnormal cell divisions, and consequently the epidermal cell layer is undefined (Becraft et al. 2001; Johnson et al. 2005; Lid et al. 2002; Lid et al. 2005). The promoters of *ATML1* and *ACR4* (used as L1 markers) are not expressed in *atdek1* mutant embryos (Johnson et al. 2005), supporting the observation that these embryos lack a defined protoderm. To circumvent the early embryo lethal phenotype of *atdek1* mutant embryos, Johnson and colleagues applied an RNAi knockdown approach to allow the embryos to partially complete embryogenesis before losing *AtDEK1* function. The resulting transgenic lines displayed seedling phenotypes characterized by loss of epidermal cell identity, leading to the presence of mesophyll-like cells instead of wild-type epidermal cells on cotyledon surfaces. This often led to cotyledon fusion (Johnson et al. 2005). *AtDEK1*RNAi phenotypes also included a reduction or loss of meristematic structure and function. In *AtDEK1*RNAi seedlings the transcript level of *AtDEK1*, epidermal genes *ACR4*, *ATML1* and meristem genes *CLV1* (*CLAVATA1*)

and *STM* was reduced compared to the wild-type (Johnson et al. 2005). *atdek1* null mutants complemented with *pRPS5A-AtDEK1* showed phenotypes ranging from wild-type to phenotypes similar to the *AtDEK1*RNAi lines depending on the level of *pRPS5A-AtDEK1:GFP* expression. Similar complementation was observed in mutant lines expressing a *pRPS5A* driven *DEK1* calpain domain alone, suggesting that this domain represents an active form of the protein. Lines overexpressing this construct displayed several post germination phenotypes that are related to cell growth regulation (Johnson et al. 2008).

Although the loss, or reduction of *AtDEK1* function consistently leads to phenotypes related to problems in the maintenance of epidermal identity or cell growth regulation (Becraft et al. 2002; Johnson et al. 2005; Johnson et al. 2008; Roeder et al. 2012), lines overexpressing the active form of AtDEK1 show neither the formation of multiple epidermal cell layers, nor the overexpression of epidermal markers (Johnson et al. 2008). As a result, it seems logical to conclude that although the activity of DEK1 is required for epidermal specification, epidermal specification is not the primary role of DEK1. DEK1 may therefore be involved in another fundamental process (such as cell-cell adhesion, cell division, etc.) which is fundamentally important for epidermal cell fate maintenance. Understanding the role of DEK1 will therefore shed light on the developmental requirements for the formation of a functional epidermal cell layer.

1.2.6.2 RLKs are involved in cell-cell communication in the epidermis

The study of *ACR4* and *ALE2* in *Arabidopsis* indicates that these two receptor kinases might work together to regulate epidermal cell fate although neither the ligands nor the substrates of these membrane proteins are known.

CR4 was first described in maize as a receptor-like kinase (Becraft et al. 1996). The structure of CR4 and the *Arabidopsis* homolog ACR4 (Figure 1.7A) is analogous and is characterized by an extracellular domain with seven repeats of a similar sequence (denominated CR repeats) and a cysteine-rich region similar to the ligand binding domain of the TNFR (tumor necrosis factor receptor), a transmembrane domain, a

serine/threonine kinase domain, and a highly conserved and unique C-terminal domain (Becraft et al. 1996; Tanaka et al. 2002).

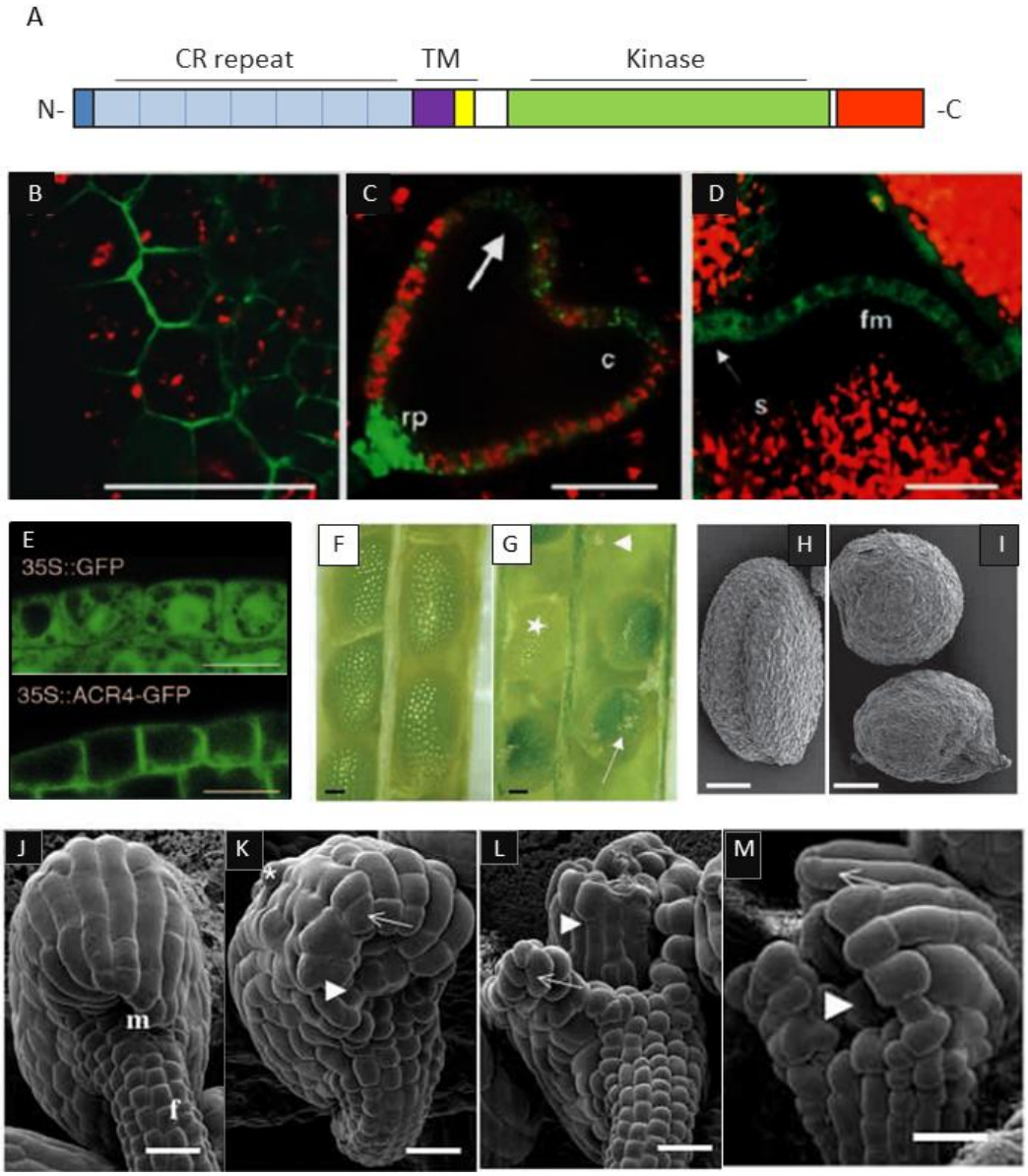


Figure 1.7 – ACR4 role in epidermal signalling. A) The ACR4 receptor kinase comprises 7 CR repeats, a transmembrane domain and the kinase domain; *pACR4:ACR4-GFP* is expressed specifically in the L1 layer of ovules (B), embryo (C) and SAM (D) and ACR4-GFP is localized in the lateral and basal membranes of epidermal cells of leaf primordia (E). Compared to wild-type (F, H, J), loss of *ACR4* function results in seed development problems (G) (namely seed abortion (arrow head), seed retardation (star) and epidermal outgrowths (arrow)), round shaped seeds (I) and defects in ovule integument development (K-M). (Adapted from (Gifford et al. 2003; Watanabe et al. 2004))

Meyer and colleagues have analysed the *in vitro* activity of ACR4 in detail and concluded that the protein is capable of autophosphorylation with conformational

changes. The same authors identified 16 autophosphorylation sites and demonstrated that after autophosphorylation occurs, the intracellular domain loses its dimerization ability (Meyer et al. 2011). The kinase domain is functional in both species, but this activity is not necessary for ACR4 function suggesting that ACR4 might interact with other receptor-like kinases (RLKs) during signalling events, as has recently been shown in roots where ACR4 interacts with CLV1 (Gifford et al. 2005; Jin et al. 2000; Meyer et al. 2011; Stahl et al. 2013).

While the expression pattern of *ACR4* is restricted to the epidermal layers of the embryo, meristems and organ primordia (Figure 1.7 C,D), the maize homolog CR4 is also expressed in the inner layers of the inflorescence (Becraft et al. 2001; Gifford et al. 2003; Tanaka et al. 2002). The *ACR4* transcript was first detected in the whole embryo at the 8-cell stage and afterwards the expression of this gene becomes restricted to the L1 layer. In inflorescence tissues the expression of *ACR4* promoter or ACR4-GFP fusion protein is clearly restricted to the L1 layer (Gifford et al. 2003; Tanaka et al. 2002). The expression pattern of *ACR4* is strikingly similar to the expression pattern reported for *ATML1/PDF2.*

The analysis of GFP fusions to ACR4 has been important in gaining an understanding of the function of this protein *in planta*. In leaf primordia ACR4 has been detected in the basal and lateral plasma membrane of epidermal cells (Watanabe et al. 2004) consistent with a role in communication between L1 cells and possibly positional signalling between L1 and L2 layers. At the cellular level the ACR4 fusion protein is localized in protein export bodies and internalized vesicles (Gifford et al. 2005), suggesting an active protein turnover. Functional analysis of different domains of ACR4 suggests that the extracellular domain (CR domain and TNRF) is required for ACR4 functionality and proper internalization, since ACR4 with mutations in the extracellular domain is in general stabilized, contrary to functional ACR4 that is rapidly internalized and turned over (Gifford et al. 2005). The results from functional analysis suggest that signal transduction by ACR4 requires ligand binding to the extracellular domain which may induce targeted interactions with partners, and subsequently trigger the internalization and degradation or recycling of this receptor.

The L1 expression pattern and localization of ACR4 are consistent with a role of ACR4 in regulating epidermis cell identity and indeed, the phenotypes of *acr4*, or *cr4* mutants in maize, demonstrate that this receptor kinase is involved in the specification of epidermis cell fate. Surprisingly, however the phenotypes of *acr4* are significantly weaker than those of *cr4*. The loss of function of *CR4* in maize is characterized by disruption of specification of the aleurone layer (the specialized outer layer of the endosperm), which in *cr4* mutants develops as starchy endosperm (Becraft et al. 1996). Additionally *cr4* mutants display several phenotypes throughout embryo and shoot development associated with epidermal cell fate disorders. These include misshapen cells with defects in cell wall structure and thickness, cuticle formation and vesicle trafficking, and the formation of clusters of cells with abnormal epidermal identities in leaves (Becraft et al. 1996; Jin et al. 2000).

Several *acr4* mutant alleles have been analysed in two independent studies presenting a variety of phenotypes. The *acr4-1* and *acr4-2* alleles displayed seed development defects (Figure 1.7) from malformed and unfertilized ovules, to fused, misshapen or aborted seeds (Gifford et al. 2003). The proportion of seed abortion and abnormal seed morphology varied among siliques and was traced to maternal sporophytic defects caused by absence of a functional ACR4. In particular, the disruption of integument outgrowth was very marked (Gifford et al. 2003). Interestingly *acr4* ovules that were successfully fertilized showed relatively normal embryo development, and the post germination phenotype of *acr4* plants was globally comparable to wild-type with the exception of irregular cell organization of the sepal margins and ovule integuments (Gifford et al. 2003). Additionally, *acr4-1* and *acr4-5* mutant seedlings and seeds were reported to show mild seedling cuticle defects detected by the toluidine blue staining test (Watanabe et al. 2004), once more suggesting that ACR4 as an important role in epidermal development.

The mild phenotypes of *acr4* mutant compared to *cr4* mutants in maize, or other mutants affecting embryo patterning in *Arabidopsis* could be indicative of functional redundancy with other receptor kinases, including other CRR (CRINKLY4 RELATED) proteins. However, quintuple mutants in which the function of ACR4

and the four CRR kinases present in *Arabidopsis* is eliminated, have mutant phenotypes which are not significantly stronger than those of *acr4* single mutants (unpublished work of Jannat-e-Zereen, a previous PhD student). It is also possible that other pathways act functionally parallel to the ACR4 signalling pathway, at least in the embryo. The requirement for ACR4 signalling is clearly particularly strong in sepals and integuments, tissues consisting of juxtaposed epidermal layers, and thus potentially more sensitive to disruptions in L1 cell fate signalling.

The loss of function of the *ALE2* gene which encodes a receptor-like kinase unrelated to ACR4, also results in defects in seed and post germination development indicative of a loss of epidermal cell fate. Embryo patterning and in particular epidermal cell morphology in *ale2* mutants is defective, and cell layer disorganization can be observed in the epidermal layer of the seedling SAM. As shown for *acr4* mutants, *ale2* seedlings are permeable to toluidine blue (Tanaka et al. 2007). Compared to *acr4* mutants the cuticle and ovule phenotypes are significantly more severe in *ale2* mutants. Contrary to *acr4*, the postgermination phenotypes of *ale2* mutants include defects in leaf and flower morphology, and fused organs including ovule fusion which causes total sterility. The *ale2* mutation is haploinsufficient in an *acr4* mutant background in terms of fertility; while *ale2acr4* plants are viable but sterile, plants homozygous for *acr4* but heterozygous for *ale2* are less fertile than *acr4* plants. This indicates that these proteins could act together in heterodimers, an idea supported by the fact that ACR4 and ALE2 can trans-phosphorylate one another (Tanaka et al. 2007).

Interestingly, Tian and colleagues observed the co-localization of DEK1 and CR4 in plasma membranes and endosomes in maize aleurone cells and suggested that CR4 was associated with the plasmodesmata between aleurone cells (Tian et al. 2007). The authors proposed that the mechanism of aleurone cell fate specification involves an unknown substrate of DEK1 that is cleaved and activated upon signalling events that occur in aleurone surface, and afterwards transported between aleurone cells via plasmodesmata, facilitated by CR4.

The relationship between DEK1 and CR4 has been discussed both in maize and *Arabidopsis*. In maize, although DEK1 is not required for the CR4 activity, in

general weak *dek1* mutations are epistatic to weak *cr4* alleles. Strong alleles of either gene result in lethality (Becraft et al. 2002). Interestingly at the cellular level the phenotypic analysis of double and single mutants suggests that for different traits the relationship between these two genes could be epistatic, additive or synergistic (Becraft et al. 2002). Considering the different phenotypes between loss of function of these genes in *Arabidopsis* and maize, the genetic analysis of *ACR4* and *AtDEK1* could help to define their relationship in epidermis specification. In order to analyse the interaction between *ACR4* and At*DEK1* in *Arabidopsis*, an *AtDEK1RNAi* construct was transformed into the *acr4* mutant background. A higher proportion of seedlings with stronger RNAi phenotypes was observed in the *acr4* background compared to the wild-type background consistent with ACR4 and DEK1 acting in parallel signalling pathways, possibly upstream the same targets involved in epidermal cell fate specification (Johnson et al. 2005).

Two further closely related receptor-like kinases have also been implicated in embryo morphogenesis: TOAD2 (TOADSTOOL 2) and RPK1 (RECEPTOR-LIKE PROTEIN KINASE 1). The phenotypes of *rpk1toad2* double mutants suggest that these two proteins are redundantly required in radial patterning and basal pole differentiation in embryogenesis (Nodine et al. 2007). In particular, *rpk1toad2* embryos display abnormal divisions in the basal epidermal cells and root pole. Consequently they lose epidermal cell identity (loss of *ATML1* expression), and display an expansion of *AGO10* and *SHR* expression patterns to the outermost cell layer. *WOX5* expression pattern was also generally disrupted in abnormal double mutant embryos (Nodine et al. 2007). These results suggest that TOAD2 and RPK1 are important for the communication between epidermis and inner layers (Figure 1.9) and that the disruption of this signalling has a significant impact in radial patterning and basal pole formation.

1.2.6.3 Seed specific pathways involved in cuticle formation

So far I have described three possible signalling events involved in epidermal cell fate specification and maintenance, namely via DEK1 or two pairs of receptor kinases ALE2 and ACR4 or TOAD2 and RPK1. Although these pathways are important for embryogenesis, they are also necessary during the whole post

embryonic growth. In this section I will describe a seed specific pathway specifically active during embryogenesis and important for embryo cuticle deposition/surface formation and epidermal cell fate specification (Figure 1.8 and 1.9). This pathway involves endosperm/embryo signalling.

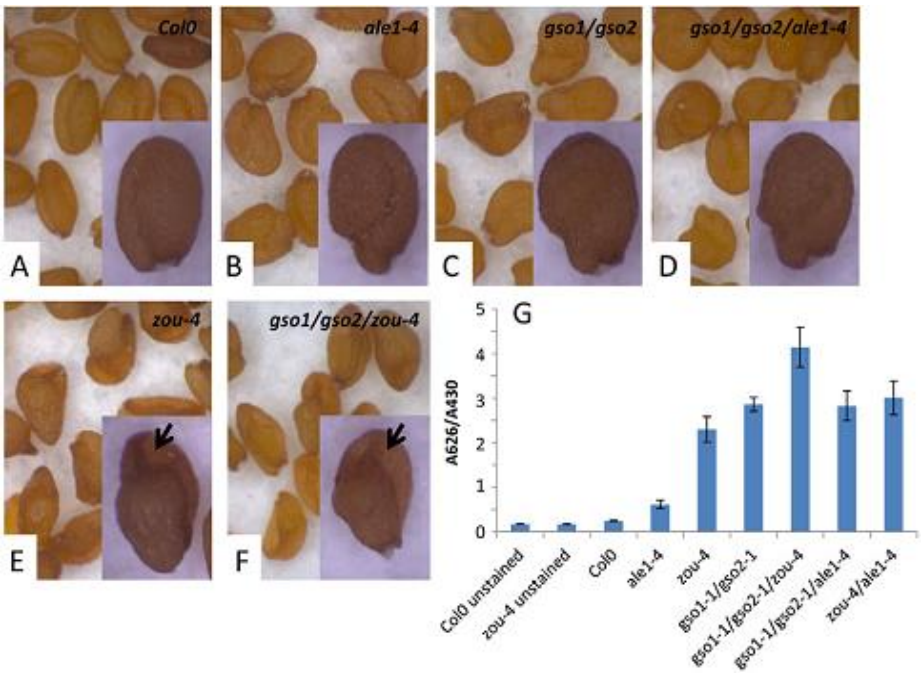


Figure 1.8 – Effects of the disruption of embryonic cuticle formation pathways. A-F) Seed phenotypes of *Col0*, *ale1-4*, *gso1-1gso2-1*, *gso1-1gso2-1ale1-4, zou-4* and *gso1-1gso2-1zou-4* homozygous mutants (arrows show the effects of endosperm persistence). G) Toluidine blue assay using seedlings of different mutant backgrounds. The A626/A430 ratios quantify the cuticle permeability and are consistent with the degree of seedling defects in different mutant backgrounds. (Adapted from (Xing et al. 2013))

Seed development in *Arabidopsis* requires constant communication between different tissues. The endosperm develops in three main stages: a syncytial stage where nuclei divide in a common cytoplasm surrounding a huge and rapidly expanding vacuole, the subsequent cellularization of the endosperm, and the degradation of the endosperm, which is coordinated with embryo growth and seed filling. Signalling between the endosperm and embryo appears to be important during late stages of embryogenesis to coordinate the growth of embryo and endosperm breakdown. In addition, communication between embryo and ESR could play a key role in regulating epidermal cell fate, potentially being involved in the perception of the "outside", the initial signalling event required to trigger

ATML1/PDF2 transcription and potentially other signalling pathways. However, the mechanisms involved in this communication, if it exists, are poorly understood.

The bHLH transcription factor ZOU, specifically expressed during seed development, has been shown to be involved in endosperm breakdown and embryo surface formation (Xing et al. 2013; Yang et al. 2008). After fertilization, *ZOU* expression is restricted to the ESR of the endosperm (Yang et al. 2008). In *zou* seeds, the boundary between endosperm and embryo is abnormal due to defects in cuticle deposition which leads to epidermal cell adhesion to the endosperm, which is persistent (Figure 1.8). The endosperm persistence results in the production of a very small embryo and the shrivelling of mature seeds (Yang et al. 2008). However, *zou* seeds are able to germinate, and although the seedlings display severe cuticle defects they are able to recover and produce adult plants which are similar to the wild-type (Yang et al. 2008). The epidermal defects of *zou* seedlings are restricted to the embryo related tissues, since leaves have a wild-type phenotype. Interestingly the expression of *ATML1* and *ACR4* promoter in *zou* embryos is also similar to the wild-type (Yang et al. 2008) suggesting that the epidermal defects are a consequence of cuticle malformation rather than epidermal cell fate specification.

ALE1 encodes a subtilisin-like serine protease required for embryonic surface formation (Tanaka et al. 2001). The ALE1 expression pattern is similar to that of *ZOU* since it is restricted to the ESR region during seed development. *zou* mutants are epistatic to *ale1* mutants and *ALE1* is not expressed in *zou* mutants. Taken together these results suggest that ZOU and ALE1 act in the same pathway (Figure 1.9) and that *ALE1* may be a direct target of ZOU (Tanaka et al. 2001; Yang et al. 2008). As in *zou* mutants, the expression patterns of epidermis specific genes such as *ATML1*, *FDH* (*FIDDLEHEAD*) or *PDF1* (*PROTODERM FACTOR 1*) in *ale1* mutants is similar to the wild-type (Tanaka et al. 2001; Tanaka et al. 2007) suggesting that the epidermal defects are likely caused by lack of intact cuticle layer and consequently epidermal differentiation, but that epidermal cell identity itself is well defined. The identification of *ALE1* as a possible direct target of ZOU, and the fact that *ale1* embryos display adhesion to the endosperm but *ale1* endosperm is not persistent, suggest that ALE1 is involved in cuticle formation but not in endosperm

breakdown. Consistent with this, the expression of *ALE1* under the *pSUC5* promoter (which is expressed in ESR but is not regulated by ZOU) is able to partially rescue the cuticle phenotype of *zou* mutant embryos, but is unable to restore endosperm breakdown (Xing et al. 2013; Yang et al. 2008). Therefore ZOU controls two separable pathways regulating embryo cuticle formation and endosperm breakdown.

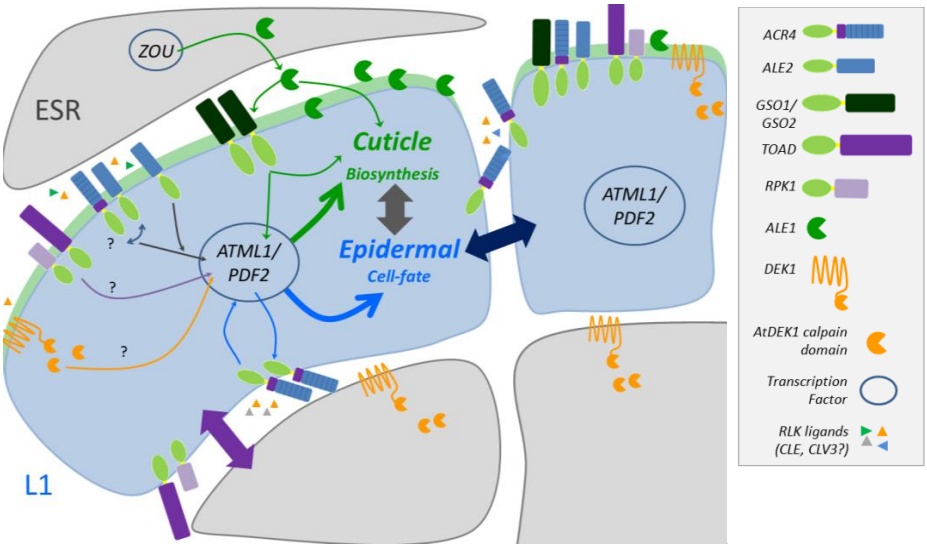


Figure 1.9 – Epidermal signalling pathways in embryogenesis. One model of epidermal signalling proposes that different signalling pathways contribute directly to activation of the master regulators *ATML1/PDF2* which in turn regulate genes involved in epidermal cell fate and cuticle biosynthesis. A seed specific cuticle reinforcement pathway (green) involves ZOU activation of *ALE1* transcription. The signal processed by ALE1 is transduced by the GSO1/GSO2 receptor kinases which promote cuticle formation, possibly via *ATML1/PDF2* regulation. ACR4 and ALE2 probably act together in a signalling pathway that activates *ATML1/PDF2* transcription. DEK1 is involved in the perception of unknown signal that activates the cleavage of DEK1 calpain which is also involved in epidermal cell fate determination/maintenance possibly by activating *ATML1/PDF2* transcription (orange). Finally TOAD2 and RPK1 have redundant roles in L1/L2 signalling (Purple double arrow) important for epidermal L1 identity. The TOAD2/RPK1 signalling pathways may also have an impact on *ATML1/PDF2* regulation. ACR4 has an important role in cell-cell communication between epidermal cells (blue double arrow) and possibly participates in a positive feedback loop with the ATML1/PDF2 transcription factors (blue arrows). (Adapted from (Javelle et al. 2011b; Nodine et al. 2007).

The GSO1 and GSO2 receptor-like kinases act redundantly in embryo cuticle formation (Tsuwamoto et al. 2008). These genes are expressed in the embryo from the globular stage onwards, seedlings and flower buds. While the single mutants have no visible defects, the *gso1gso2* double mutant displays pleiotropic embryo and seedlings phenotypes that can be related to cuticle defects in the embryo. The complementation of the double mutant with either gene is enough to rescue the

gso1gso2 phenotype (Tsuwamoto et al. 2008). In embryogenesis the lack of these two genes results in abnormal bending of embryos due to adhesion between the embryo and endosperm. Additionally, *gso1gso2* seedlings display epidermal morphology defects that include adaxial adhesions of cotyledons, misshapen cotyledons, and increased permeability to toluidine blue of cotyledons but not true leaves (Tsuwamoto et al. 2008). The role of GSO1/GSO2 after germination is open to question. Weak phenotypes, but no cuticle defects were identified in the vegetative tissues of *gso1gso2* plants (Xing et al. 2013). Possibly GSO1 and GSO2 are functionally redundant with other RLKs in vegetative tissues whereas in embryo development they play an important role in cuticle formation. Like *ale1* mutants, *gso1gso2* double mutants show no defects in endosperm breakdown.

The study of the *zouale1*, *gso1gso2zou* and *gso1gso2ale1* mutant combinations (Figure 1.8) shows that the *gso1gso2* double mutation is epistatic to *ale1* and *zou* seedling phenotypes, whereas in endosperm development *zou* mutation is epistatic to the *gso1gso2* or *ale1* mutation (Xing et al. 2013; Yang et al. 2008). Therefore, ALE1, GSO1/GSO2 and ZOU all appear to act in the same pathway with respect to embryonic cuticle formation (Figure 1.9). Interestingly, seedling permeability to toluidine blue (Figure 1.8G) shows that the *ale1* cuticle defects are significantly weaker than those of *gso1gso2* or *zou* (Xing et al. 2013), suggesting that other proteins may act redundantly with ALE1.

The genetic relationship between ACR4/ALE2 and ZOU/ALE1/GSO1/GSO2 suggests that they act in independent pathways. The double mutants *acr4 ale1*, *ale1 ale2*, *zou ale2* or *zou acr4* show synergistic genetic interactions, since the seedling and embryo phenotypes of double mutants are significantly more severe than the respective phenotypes in single mutants, leading to seedling lethality (Tanaka et al. 2007; Yang et al. 2008). These observations suggest that ZOU, ALE1 and GSO1/GSO2 regulate cuticle formation independently of ACR4/ALE2, despite the fact that the disruption of the ACR4/ALE2 signalling pathway also has consequences for cuticle integrity. Importantly, the expression of epidermal markers *PDF1*, *FDH* and *ATML1* is disrupted in the *ale1ale2* and *ale1acr4* mutants (Tanaka et al. 2007)

suggesting that correct epidermal cell fate specification requires both pathways (Figure 1.9).

In order to understand the relationships between the acquisition of epidermal cell fate and cuticle formation during embryogenesis it will be necessary to complete the genetic interaction analysis between genes in both pathways, including the genes encoding the ATML1/PDF2 transcription factors which regulate epidermal cell fate. One possibility is that ATML1 and PDF2 are the central regulators of both epidermal identity and cuticle deposition and that both the ACR4/ALE2 and GSO1/GSO2 signalling pathways are required for the activation of these genes. In this scenario signalling via the ZOU/ALE1/GSO1/GSO2 pathway could also play an important additional role in reinforcing cuticle formation (Javelle et al. 2011b). Alternatively there is a feedback between cuticle formation and epidermal identity that is independent of ATML1/PDF2 regulation. One hypothesis is that an intact cuticle layer promotes ACR4/ALE2 signalling activity via a mechanism yet to be defined, leading to synergistic effects when both pathways are disrupted (Figure 1.9).

1.2.6.3 Finding the unknown trigger of RLK mediated signalling

So far the ligands of the different receptor kinases involved in epidermal specification in embryogenesis remain unknown. Nevertheless, it is likely that at least some of these RLKs recognize CLE (CLV3/ESR-RELATED) peptides. In general, *CLE* genes encode secreted proteins that promote or inhibit cell differentiation in plants. It should be noted that the "ESR-RELATED" peptides referred to in the acronym "CLE" are secreted peptides expressed in the maize ESR region, and identified by the group of Peter Rogowsky in my host laboratory in Lyon.

Interestingly, four CLE signalling pathways with similar structures have been identified in *Arabidopsis*. In each pathway CLE peptides activate receptor-like kinases that in turn regulate WOX genes. These pathways are involved in SAM regulation (CLV3 (CLE peptide), CLV1/TOAD2/CLV2 (RLKs), and WUS (WOX)), vascular development (TDIF (CLE peptide), TDR (RLK) and WOX4), root columella (CLE40, ACR4, WOX5) and seed development (CLE8, unknown RLK,

WOX8)(Fiume and Fletcher 2012; Ha et al. 2010; Kinoshita et al. 2010; Sarkar et al. 2007; Stahl et al. 2013; Stahl et al. 2009; Yadav et al. 2011). CLE peptides are expressed throughout plant development and in particular several CLE peptides are expressed in specific regions of root tip and SAM suggesting that CLE signalling is organized in complex networks associated with cell fate specification (reviewed in (Betsuyaku et al. 2011; Kiyohara and Sawa 2012; Wang and Fiers 2010)).

Briefly, the core of SAM organizing centre is a feedback loop where WUS induces stem cell fate by repressing *CLV1* expression and activates CLV3 production which in turn represses *WUS* (Busch et al. 2010; Lenhard and Laux 2003; Yadav et al. 2011). In roots, ACR4 is implicating in root patterning, regulating the restriction of formative divisions in roots important for specification and maintenance of stem cell pools (De Smet et al. 2008). The investigation of *ACR4* loss of function indicates that ACR4 represses irregular divisions in columella cells of root tip, and together with other CR4 receptor kinases, participates in the repression of pericycle cell divisions (De Smet et al. 2008). The combination of genetic approaches, expression profiling and peptide treatments has permitted the definition of a regulatory network controlling stem cell fate in the root apex, where CLE40 peptide activates ACR4 signalling which in turn represses *WOX5* expression and consequently stem cell fate (Stahl et al. 2009). It has recently been shown that the CLV1 receptor kinase is also activated by CLE40 and interacts physically with ACR4 in the repression of stem cell fate in roots (Stahl et al. 2013). During seed development, *CLE8* is exclusively expressed at early stages of embryogenesis in embryos and endosperm. The study of *cle8* and *CLE8i* lines shows that CLE8 is involved in regulation of basal pole cell divisions and endosperm proliferation. Interestingly, CLE8 positively regulates *WOX8* expression to promote basal cell lineage patterning (Fiume and Fletcher 2012). In maize ESR/CLE18 is expressed specifically in the ESR of the endosperm, but as yet, no function for this peptide has yet been defined (Bonello et al. 2002). To date, the receptor of CLE8 has not been identified, although obvious candidates include TOAD2/RPK1 and possibly ACR4.

It should be remembered that CLE peptides, like many other secreted signalling molecules, undergo both post proteolytic processing and post translational

modification (such as glycosylation), before becoming active ligands (Matsubayashi 2011; Ohyama et al. 2009). Although the enzymes involved are mainly unknown in plants, subtilisin serine proteases have been proposed to be involved in the cleavage of peptide precursors to give mature peptide ligands. One intriguing possibility is that ALE1 could be involved in processing a CLE peptide, such as CLE8, during seed development. This possibility remains to be tested.

1.3 Epidermal cell fate

1.3.1 Role of Epidermis

The epidermal layer, also known as the L1 layer, covers the surface of all aerial organs in plants and is the first line of defence against abiotic and biotic stress. One important characteristic of this population of cells is that, after early embryogenesis they undergo an almost exclusively anticlinal division pattern. In addition, epidermal cells can adopt a bewildering variety of different identities, depending on which plant organ, and where in that organ, they are found. This developmental plasticity is essential to L1 function, and highlights the likely interplay between organ identity regulators and epidermal cell fate regulators. For instance in *Arabidopsis*, a "basal" leaf epidermal cell can become a pavement cell, a stomatal guard cell, or a trichome, responsible for structural robustness and resistance, gas exchange and plant defence respectively. Finally, almost all epidermal cells produce a continuous cuticle layer which has an important role in plant development, defence and stress response and the composition and structure of which is highly variable from organ to organ.

Key regulators of L1 identity such as ATML1/PDF2, ACR4, and other epidermal specific genes are expressed both in the embryo when epidermal cell fate is established and in the epidermal layer of shoot and flower meristems and young organs, presumably to ensure that epidermal cell fate is maintained throughout plant development. The disruption of the epidermal layer has severe consequences for plant development. The disturbance of *DEK1* using RNAi lines lead to impairment of meristematic function (Johnson et al. 2005), and the loss of ATML1/PDF2 function in *atml1pdf2* double mutants or constitutive/ectopic expression of *ATML1* has severe defects in shoot apical meristem development leading to seedling lethal phenotypes

(Abe et al. 2003; Takada et al. 2013). These results suggest that the L1 layer is required for meristem maintenance and function.

Briefly, the shoot apical meristem is the structure responsible for the emergence of flowers and leafs. The central zone of the meristem is composed of undifferentiated cells regulated by the stem cell niche and is surrounded by a zone of proliferation from which organs emerge following specific spatial and temporal patterns. In *Arabidopsis* the L1 and L2 layers of the SAM generally undergo anticlinal divisions and together form the tunica that covers the internal tissue characterized by more random division planes and called the corpus.

Meristematic maintenance and organ primordium initiation require a complex network of genetic interactions. The determination of different meristematic regions is comparable to an orchestra where different types of instruments play together in harmony. In this comparison, transcriptional programs, signalling pathways, cell physical properties, and anisotropy and directionality of cell divisions are the different sections of the orchestra, whereas hormone signalling (auxin, gibberellin, brassinosteroïds and cytokinin) can be seen as the conductor. In general, the maintenance of the stem cells requires the WUS/CLV signalling pathway coupled with cytokinin signalling and organ initiation requires regulated auxin transport and response (reviewed in (Besnard et al. 2011; Ha et al. 2010)). In this process the L1 layer is a crucial vehicle of signalling, and the physical properties of epidermal cells critically influence organogenesis. For instance PIN1, the main auxin transporter in the meristem, is expressed mainly in the epidermal layer and directs auxin efflux in such a way as to produce regularly spaced auxin peaks which in turn induce organ primordium outgrowth (de Reuille et al. 2006; Reinhardt et al. 2003; Vernoux et al. 2011). This process is thought to act, in part, by the influence of auxin on the growth rates and growth anisotropy of epidermal cells (Kwiatkowska and Dumais 2003).

Radial patterning, and in particular the L1/L2/L3 organization of the meristem, is important for organ initiation and consequently leaf polarity. Interestingly, the effects of L1 specific cell ablations in tomato meristems indicate that the meristematic L1 layer is important for adaxial/abaxial pattering of leaf primordia (Reinhardt et al. 2005). The antagonistic relationship between HD-ZIP III and KAN transcription

factors is vital for the determination of adaxial and abaxial cell fate respectively (Emery et al. 2003). Small RNA-mediated gene regulation is critical for leaf polarity, since miRNA165/166 restricts expression of *HD-ZIP III* genes and a ta-siRNA-ARF gradient restricts *ARF3* and *ARF4* expression domain in an antagonistic manner, these two transcription factors promote abaxial cell fate and the respective double mutant, *afr3 arf4*, has a similar leaf phenotype to *kan1 kan2* (Chitwood et al. 2009; Emery et al. 2003; Pekker et al. 2005; Rhoades et al. 2002). These miRNAs are thought to move through different cell layers to regulate their targets and are predominantly expressed in the L1 layer. The transcription factors AS1 and AS2 (expressed in L1 layer) promote adaxial cell fate and are downregulated by KAN and YABBY genes are downstream these regulators (KAN, HD-ZIP III, ARF and AS) involved in leaf polarity (Byrne et al. 2000; Byrne et al. 2002; Eshed et al. 2001; Eshed et al. 2004; Iwakawa et al. 2007; Wu et al. 2008). Leaf polarization requires a signal from the meristem that promotes adaxial cell fate and should be transmitted through the L1 layer. (Reviewed in (Kidner and Timmermans 2010)).

Given the importance of the L1 layer for meristematic function, illustrated by the examples briefly outlined above, it is perhaps unsurprising that disrupting epidermal fate causes fundamental defects in meristem maintenance. Nonetheless, given that meristematic function is seminal to plant survival, this fact makes the study of mutants which affect L1 identity frustratingly complex in plant systems.

1.3.2 Cuticle formation

The cuticle is produced by almost all plant epidermal cells and essentially consists of cutin and waxes that form a continuous hydrophobic layer covering aerial surfaces of plants. The main function of cuticle is to protect the plant against biotic or abiotic stress and to control fluxes of gas, water and other molecules (Bernard and Joubes 2013; Lu et al. 2012; Reina-Pinto and Yephremov 2009; Wang et al. 2011). However, in addition to these physiological roles, there is a strong connection between cuticle production and plant development, especially regarding epidermal morphology and the prevention of organ fusion (reviewed in (Javelle et al. 2011b)). The impairment of the cuticle layer due to changes in wax and lipid metabolism disturbs epidermal differentiation signalling, and the patterning and morphology of

specialized epidermal cells such as stomata and trichomes (Bergmann et al. 2004; Bird and Gray 2003; Marks et al. 2009).

The formation of the plant cuticle consists in several loading steps for cuticular waxes and cutin (Figure 1.10). At early stages of leaf development the epidermis is covered by a thin wax layer, the procuticle. As the leaf expands the sequential deposition of polysaccharides, cutin and waxes occurs in consecutive layers that form the cuticle proper. In parallel, epicuticular waxes are deposited on the outer surface covering the cuticle proper, and often forming wax crystals of widely varying structure. The external portion of the epidermal cell wall is incorporated into the cuticle and is composed of a matrix of cutin layers and intracuticular waxes intercalated with cellulose microfibrils. Polymerized cutin is thought to provide structural rigidity to the cuticle and by extension to the epidermal layer.

The cutin polymer is composed of C16 and C18 hydroxy or epoxy fatty acids and glycerol, whereas cuticular waxes comprise a mixture of C24 to C34 alkenes, alcohols, ketones and wax esters derived from VLCFAs (very long chain fatty acids) (reviewed in (Beisson et al. 2012; Bernard and Joubes 2013; Bird and Gray 2003; Kunst and Samuels 2009; Pollard et al. 2008)). The precursors of cutin are fatty acids that are produced in plastids and then processed in the endoplasmic reticulum (ER). The enzymatic steps encompass acyl activation, reduction/oxidation or elongation, glycerol esterification and possibly oligomerization. The acylglycerols, oligomers or polymers are then transported from the ER to the plasma membrane and afterwards they cross the cell wall to integrate the cuticle layers. Despite the wealth of molecular information regarding specific component of the synthesis pathway, several important questions remain unanswered regarding the structure, assembly and biochemistry of cutin polymer. The polymerization of cutin is thought to involve primary and secondary ester bonds between monomers which should give rise to different polymer structures. However, the assembly mechanisms and cellular localization of the polymerization process and in particular the localization of polyesters synthases that are essential to the assembly of cutins, are unknown. The transport process of cutins from the ER to cuticle deposition is also unclear and is

likely to involve ABC transporters and LTPs (lipid transport proteins). (Reviewed in (Beisson et al. 2012; Pollard et al. 2008))

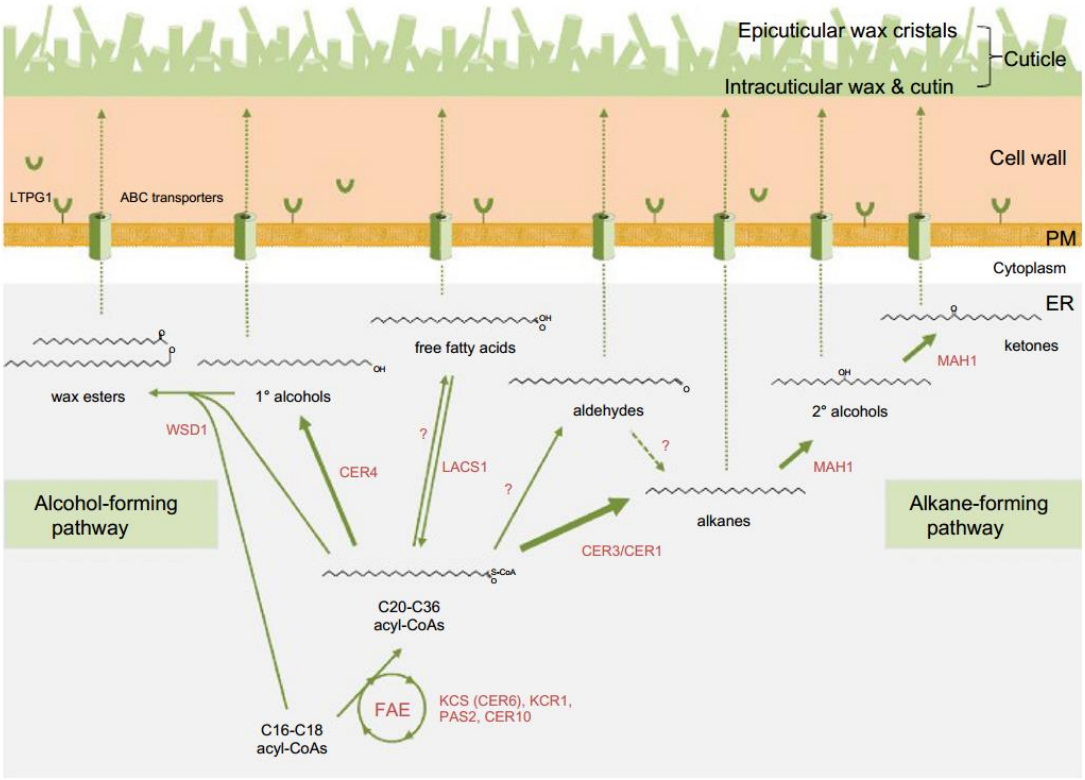


Figure 1.10 – Wax biosynthesis and transport. The VLCFAs are produced by the fatty acid elongase (FAE) complex localized in the ER and composed by four enzymes CER6, KCR1, PAS2 and CER10. Afterwards VLCFAs are modified into the aliphatic compounds of waxes via two distinct biosynthetic pathways. The alcohol-forming pathway produces primary alcohols and wax esters mediated by CER4 and WSD1 respectively. The alkane-forming pathway produces aldehydes, alkanes (mediated by CER3/CER1) and their derivatives the secondary alcohols and ketones via MAH1 activity. Finally the wax components are transferred from the ER to the plasma membrane (PM) and exported to the cuticle layer. The mechanisms of wax transport are unknown but this process may involve ABC transporters and LTP proteins such as LTPG1. (Adapted from (Bernard and Joubes 2013))

The comparison of lipid profiles between mutant plants and wild-type siblings has allowed the identification of many genes involved in the biosynthesis and transport of cutin and cuticular waxes (Figure 1.10). The first cuticle mutants identified were characterized by their shiny leaf and stem surfaces and therefore were named *glossy* and *cer* (*eceriferum*) in maize and *Arabidopsis* respectively. Most of these mutants were found to have impaired composition, distribution or load of cutin or waxes which helped to identify gene function. However, it is now recognized that many mutants involved in critical aspects of wax biosynthesis such as *fiddlehead (fdh)* or

cer3 do not display the classic eceriferum phenotype, but rather show severe organ fusion phenotypes or even embryo lethality (like *pas2* (*pasticcino2*) and *kcr1* (β-ketoacyl-CoA reductase)) due to impaired VLCFA formation during embryogenesis (Bach et al. 2008; Beaudoin et al. 2009; Beisson et al. 2012; Bernard and Joubes 2013; Chen et al. 2003; Yephremov et al. 1999). The basis for this embryo lethality is still not fully understood.

Wax biosynthesis involves three stages. The first stage is shared with other lipid biosynthetic pathways and consists of the synthesis of C16 and C18 fatty acids which occurs in plastids. The second stage involves the elongation of these fatty acids in the ER to make VLCFAs, the precursors of cuticular waxes. The final stage comprises several modifications of VLCFAs in the ER to produce alkanes, aldehydes, esters and ketones that will be exported to the cuticle (Figure 1.10).

In the plastid compartment, the long chain fatty acyl groups (C16-C18) are produced by fatty acid synthase (FAS), released by an acyl ACP thioestease (FATB), and activated to CoA thioester by long chain acyl-CoA synthetase (LACS) so that they can be exported to the endoplasmatic reticulum (ER) where they will be extended into VLCFAs via a reiterative mechanism involving the fatty acid elongase (FAE) complex. Each cycle of the FAE complex results in the addition of two carbons from malonyl-CoA to the growing acyl-CoA donor and this process encompasses four sequential steps. The first step, condensation, is performed by FDH/CER6/KCSs (β-ketoacyl-CoA synthetase) followed by a first reduction mediated by KCR1/KCRs and a dehydration mediated by PAS2/HCD (β-hydroxyacyl-CoA desydratase). A second reduction step mediated by CER10/ECRs (enoyl-CoA reductase) terminates the cycle resulting in a saturated acyl-CoA.

There are two main pathways for the conversion of VLCFAs into cuticular waxes; the alcohol-forming pathway and the alkane-forming pathway (Figure 1.10). In the alcohol-forming pathway the conversion of wax precursors into primary alcohols is mediated by CER4 whereas the conversion of wax precursors or primary alcohols into wax esters is mediated by WSD1 (Li et al. 2008; Rowland et al. 2006). The alkane-forming pathway involves the conversion of VLCFAs into alkanes, the

precursors of secondary alcohols and ketones. The first conversion may involve the CER1/CER3 complex or a secondary pathway via aldehyde modification, while the conversion of alkanes into secondary alcohols and ketones is mediated by the MAH1 (mid-chain alkane hydroxylase 1) enzyme (Bernard and Joubes 2013; Greer et al. 2007).

The mechanisms of transport of cuticular waxes from ER to the plasma membrane and later cuticle deposition remain controversial. Several mechanisms have been proposed including transport via junctions between the ER and plasma membrane, protein carriers, oleophilic droplets, or secretory pathways via the Golgi apparatus and other vesicles (Bernard and Joubes 2013). Recently, two ABC transporters CER5 and WBC11 (WHITE-BROWN COMPLEX HOMOLOG PROTEIN 11) together with the LTPG1 (glycosyl-phosphatidyl-inositol-anchored Lipid Transfer Protein) protein have been associated with wax secretion (Debono et al. 2009; Panikashvili et al. 2007; Pighin et al. 2004). The process of wax deposition at the plant surface is particularly mysterious, since highly hydrophobic wax molecules must transverse the extremely hydrophilic environment of the cell wall, to be secreted onto the surface of the plant. A possible role for LTPs in coating wax crystals to render them more hydrophilic has been discussed by several authors (Bernard and Joubes 2013; Schreiber 2005).

1.3.3 Epidermal cell differentiation

The outermost cell layer of plants, as mentioned above, has different roles in plant physiology and development that require the formation of specialized cell types. The establishment of different cell fates is controlled by spatial signals resulting in the formation of patterns of specialized and unspecialized cells. The leaf epidermis of *Arabidopsis* is composed of three main types of cells: stomata, trichomes and pavement cells (Glover 2000). In other organs other cell types are formed, for example the pavement cells of sepals are organized into populations of small and giant cells, whereas the abaxial and adaxial surfaces of petals are covered in distinct, but relatively uniform populations of pavement cells, which on the adaxial surface are conical to maximise light refraction (Gorton and Vogelmann 1996; Martin et al. 2002; Roeder et al. 2010). On the tip of the carpel a population of highly specialized

epidermal cells, the stigmatic papillae, perform a critical role in mediating pollen reception and germination, and are crucial for successful sexual reproduction (Higashiyama 2010; Watanabe et al. 2012).

In *Arabidopsis* leaves, the pavement cells are important for the physical robustness of the epidermis (reviewed in (Qian et al. 2009)). The main functions of pavement cells are the protection against biotic and abiotic stress, preventing water loss (thanks to the production of a cuticle), and the structural protection of underlying cell layers. Several mechanisms are involved in giving the pavement cell of *Arabidopsis* leaves their distinctive interdigitating morphology. These mechanisms are common to other epidermal cell types with diverse shapes, and involve the regulation of microtubule dynamics, GTPase signalling, vesicle transport, and endoreduplication. The overlap of mechanisms and regulators of epidermal differentiation between different cell types means that loss of function mutants of key regulators display combinations of trichome, pavement cell, stomatal or root hair phenotypes (reviewed in (Guimil and Dunand 2007; Qian et al. 2009; Smith and Oppenheimer 2005)). This section is focused on stomatal and trichome development, due to the published roles of HD-ZIP IV proteins in the differentiation of these two cell fates (Section 1.4.1).

Stomatal guard cells are essential to allow the gas exchange required for photosynthesis, whilst regulating water-loss from plant surfaces. This balance is achieved by tightly regulated changes in stomatal aperture. It is perhaps then unsurprising that the density of stomata in leaf surface is regulated by atmospheric conditions. The study of transpiration rate of several wild-type strains of *Arabidopsis*, by changing humidity, CO_2 concentration and exogenous application of ABA, showed that stomatal density increases with transpiration rate and ABA concentration in leaves (Lake and Woodward 2008). In particular, CO_2 levels have been found to alter stomatal density, although the way in which it is affected depends on the *Arabidopsis* strain used (Lake and Woodward 2008).

In addition to environmental control, stomatal formation is strictly regulated developmentally (Figure 1.11). In *Arabidopsis* this regulation means that stomatal pores are separated by at least one cell (the so-called one-cell spacing rule). This spacing is thought to be of vital importance to the function of the pore, which

undergoes rapid changes in turgor, and thus shape, in order to open and close; movements which might well be compromised if a cell directly adjacent to one of the guard cells belonged to a different pore. In *Arabidopsis* this spacing is achieved by the strict regulation of the orientations of the asymmetric cell divisions which precede stomatal differentiation (Figure 1.11A). This process therefore involves the controlled balance of cell proliferation and cell specification through specific signalling and regulatory pathways (reviewed in (Bergmann and Sack 2007; Casson and Gray 2008; Dong and Bergmann 2010; Nadeau 2009; Pillitteri and Torii 2012). Briefly the differentiation of stomatal guard cells involves the conversion of an epidermal cell into a meristemoid mother cell which, after an asymmetric division forms a meristemoid. The meristemoid then either differentiates into a guard mother cell or undergoes additional asymmetric divisions required for stomatal spacing. Finally the symmetric division of the guard mother cell followed by differentiation gives rise to the two guard cells (Figure 1.11A).

Stomatal formation involves an intricate system of cell-cell communication that requires several leucine-rich repeat receptor-like kinases LRR-RLKs, namely three receptor kinases of the ER (ERECTA) family (ER, ERL1 and ERL2 (ERECTA-LIKE)) and TMM (TOO MANY MOUTHS). Plants lacking these proteins show several defects associated with stomatal patterning. Unlike other LRR-RLKs, TMM lacks a cytoplasmic kinase domain (Nadeau and Sack 2002), suggesting that TMM signals via interactions with other RLKs. The loss of TMM function results in asymmetric division defects in the stomatal lineage due to disruption of cell-cell positional signalling and leads to an increase of stomatal density and stomatal clustering in leaves, cotyledons and siliques but to a total loss of stomata in stem epidermis (Geisler et al. 2000; Nadeau and Sack 2002; Shpak et al. 2005). ER is involved in multiple signalling pathways necessary for stress response, growth and plant development, and appears to be a key regulator of cell proliferation (reviewed in (van Zanten et al. 2009)). ERL1 and ERL2 are partially redundant with ER and their respective mutations enhance the er phenotype, with loss of the three genes leading to extreme dwarfism due to reduced cell proliferation (Shpak et al. 2004). The triple mutant *er erl1 erl2* displays strong stomatal spacing defects and excessive stomatal production (Figure 1.11A), these phenotypes are uniform in all epidermal

tissues that produce stomata (Shpak et al. 2005). The study of stomatal development in mutant combinations between *tmm* and *ER* family mutants indicates complex interactions between these four receptors suggesting that TMM and ER family members have different roles depending on the stage of stomatal development studied and the organ in which they are expressed (Shpak et al. 2005). These interactions could result from different dynamics of protein-protein interactions or the formation of diverse complexes in different organs.

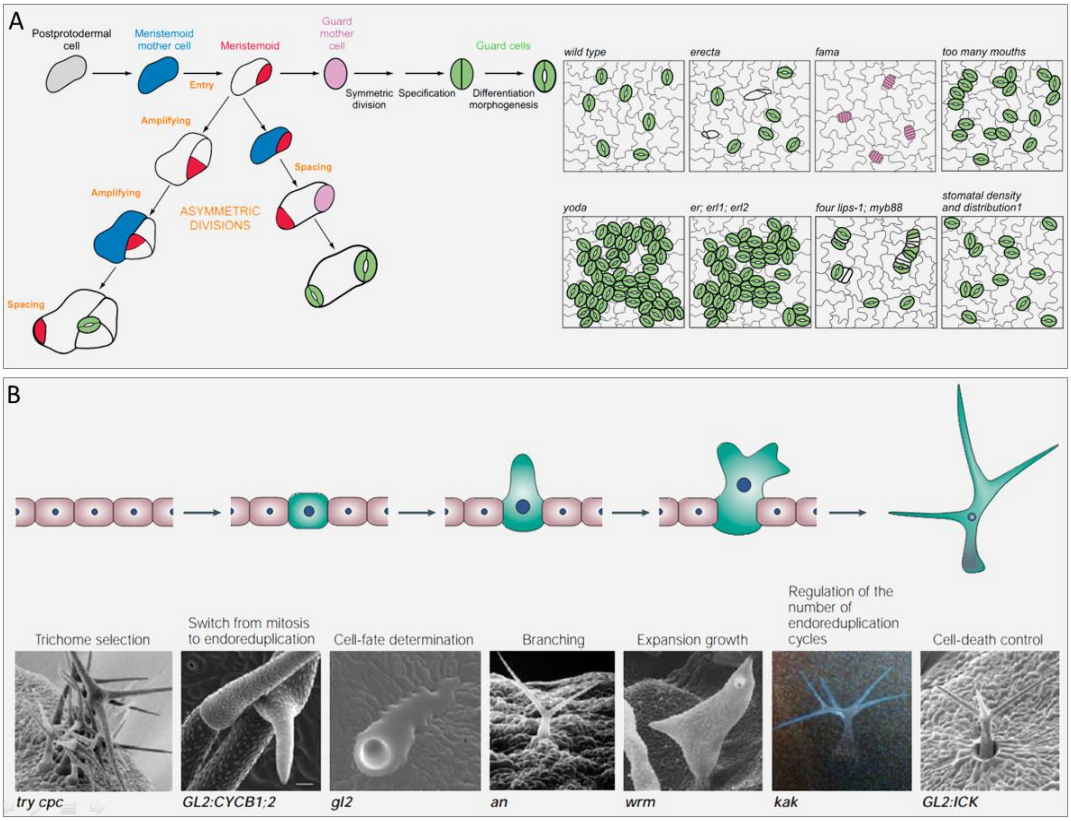


Figure 1.11 – Stomata and trichome development. A) The first step of stomatal patterning is the conversion of an epidermal cell into a meristemoid mother cell (blue) which undergoes an asymmetric division to create a meristemoid (magenta). In turn, the meristemoid is either converted into a guard mother cell (light pink) followed by symmetric division to form two guard cells (green) or undergoes one or more further asymmetric divisions required for stomatal spacing (orange). On the right the eight figures display different phenotypes associated with defects in stomatal development and patterning in leaves. B) Trichome development involves the selection of specialized cells that switch from mitosis to endoreduplication followed by cell fate determination. Trichome branching and morphology is dependent on endoreduplication events and specific regulation. Mutant phenotypes or overexpression lines display different trichome phenotypes related to trichome differentiation, patterning or morphology. (Adapted from (Bergmann and Sack 2007; Hulskamp 2004))

The positional cell-cell signalling involved in stomatal development and patterning involves the subtilisin protease SDD1 (STOMATAL DENSITY AND DISTRIBUTION 1) which has been proposed to process the ligands of TMM, ER, ERL1, ERL2 required for the activation of MAPK signalling. The loss of SDD1 results in the formation of stomatal clusters and an increase of stomatal density that reflect defects in spacing divisions (Berger and Altmann 2000). Several peptides have been shown to be involved in stomatal development. *EPF1* (*EPIDERMAL PATTERNING FACTOR 1*) and EPF2 are negative regulators of stomatal clustering and density, acting upstream TMM and ER family receptors, are therefore potential ligands of these receptor kinases, and might be processed by SDD1 (Hara et al. 2007; Hara et al. 2009; Hunt and Gray 2009). The ectopic expression of EPF4 and EPF5 also lead to a decrease in stomatal numbers suggesting that they act as negative regulators (Hara et al. 2009). Interestingly EPFL9 (EPIDERMAL PATTERNING FACTOR-LIKE 9 or STOMAGEN) is a positive regulator of stomata density and clustering, acting independently of EPF1 and SDD1 (Hara et al. 2009; Hunt et al. 2010). Unlike other EPF-Like peptides STOMAGEN is expressed in cells underlying the epidermis and its expression is environmentally regulated, suggesting that it may act as a key integrator of environmental signals into stomatal patterning pathways (Sugano et al. 2010). The current accepted model suggests that TMM and ER family receptors activate mitogen activated protein kinase (MAPK) signalling cascades important for the negative regulation of stomatal development. The most upstream element in this MAPK pathway is the YDA (YODA) MAPKKK (MAP kinase kinase kinase), which is thought to phosphorylate the MAP kinase kinases MKK 4 and 5 which in turn activate the MAP kinases MPK3 and 6 (Bergmann et al. 2004; Pillitteri and Torii 2012; Wang et al. 2007). However, because of the pleiotropic phenotypes associated with loss and gain of function alleles of these signalling elements, particularly YDA, and the complexity of the receptor kinase relationships upstream, the generation of convincing genetic evidence to support this model is challenging.

The process of stomatal differentiation requires the activity of a strict cascade of transcriptional regulators, which are presumably the targets of the signalling mechanisms outlined above. Briefly, while FLP (FOUR LIPS), MYB 88 and FAMA restrict cell cycling during the differentiation of guard cells, SPCH (SPEECHLESS)

controls asymmetric divisions and MUTE promotes the transition between meristemoid and guard mother cell (Lai et al. 2005; MacAlister et al. 2007; Ohashi-Ito and Bergmann 2006; Pillitteri et al. 2007). Additionally, two bHLH-LZ transcription factors SCREAM and SCREAM 2 also control cell fate transitions in stomata differentiation possibly interacting with the bHLH transcription factors MUTE, SPCH and FAMA (Kanaoka et al. 2008). Interestingly, both SCREAM and SCREAM2 are also involved in cold responses (Chinnusamy et al. 2003) potentially providing another link between stomatal patterning and environment responses.

Several classes of cell cycle regulators have also been implicated in stomatal differentiation, through their promotion of specific cell cycle transitions. However the relationship between the cell cycle and cell differentiation in the epidermis is not only associated with stomata but also with trichomes, pavement cells and giant cells in sepals (reviewed in (Jakoby and Schnittger 2004)). Recently, the regulation of epidermal patterning in sepals, and in particular giant cell fate, has been linked not only to cell cycle regulation but also to genes involved in the regulation of epidermal cell fate, namely *ATML1*, *ACR4*, *DEK1* and *HDG11* (Roeder et al. 2010; Roeder et al. 2012). The loss of function of any of these four genes causes a reduction in the frequency of giant cells in sepals (Roeder et al. 2012). These results indicate a link between cell differentiation, epidermal identity and cell cycle regulation in sepals. To support the theory that genes involved in basal epidermal identity can also play a subsequent role in the differentiation of diverse epidermal cell fates, other HD-ZIP IV genes are implicated in stomatal, trichome and root hair specification (Section 1.4.1) and the overexpression of DEK1 under the 35S promoter has been reported to result in the lack of trichomes (Di Cristina et al. 1996; Lid et al. 2005; Nakamura et al. 2006; Rerie et al. 1994).

The mechanism of trichome differentiation (Figure 1.11B) requires constant crosstalk between cell cycle control and pattern formation, and later between cell cycle control and the regulation of cell morphogenesis (reviewed in (Jakoby and Schnittger 2004)). Briefly, the first step of trichome development in leaves is the switch from mitosis to endoreduplication and activation of cell fate regulators such as the HD-ZIP IV transcription factor GL2 (GLABROUS 2). This is followed by cell

expansion and local outgrowth perpendicular to the epidermal layer. Afterwards the cell undergoes two branching events accompanied by endoreduplication cycles (Figure 1.11B). Trichome branching involves a complex regulatory network that includes the regulators STI (STICHEL), AN (ANGUSTIFOLIA) and coordinated endoreduplication (Folkers et al. 1997; Hulskamp 2004).

At the base of young leaves trichome precursors are distributed nearly uniformly. This spacing is dependent on positional signalling regulated by a process which is very similar to that involved in root hair patterning (reviewed in (Grebe 2012; Hulskamp 2004)). Mutant trichome and root hair phenotypes have helped to identify the signalling pathways and regulatory networks involved in these processes (Guimil and Dunand 2007; Hulskamp et al. 1994; Marks et al. 2009; Schellmann and Hulskamp 2005). In *Arabidopsis,* trichome and root hair patterning are both based on the formation of inhibitory fields and similar regulatory networks involving positive and negative regulators of cell fate (Grebe 2012; Hulskamp 2004). Two MYB transcription factors GL1 (GLABRA1) and MYB23, two bHLH transcription factors GL3 (GLABRA3) and EGL3 (ENHANCER OF GL3) and the TTG1 (TRANSPARENT TESTA GLABRA1) protein, form complexes that positively regulate trichome cell fate (Grebe 2012; Payne et al. 2000; Walker et al. 1999; Zhang et al. 2003). With exception of MYB23, at early stages of leaf development, these positive regulators are initially expressed in all cells at the leaf basis and later their expression is enhanced in cells that have acquired trichome cell fate. Negative regulators of trichome differentiation include three transcription factors TRY (TRYPTICON), CPC (CAPRICE) and ETC1 (ENHANCER OF TRY AND CPC1) which act together to inhibit trichome cell fate in neighbouring cells, interfering with GL1 transcriptional activity (Kirik et al. 2004; Schellmann et al. 2002). The leaves of *trycpc* double mutants thus display clusters of trichomes (Figure 1.11B) and the surface of *try cpc etc1* leaves is almost full of trichomes (Kirik et al. 2004; Schellmann et al. 2002).

There are two models describing trichome patterning, namely the activator-inhibitor model or the activator depletion model, both based on cell-cell transport of transcriptional factors via plasmodesmata (reviewed in (Torii 2012)). One hypothesis

considers that the activators induce the transcription of inhibitors (negative regulators) that are able to move to neighbouring cells, while GL1 and GL2 activators are stationary (Schellmann et al. 2002; Torii 2012). The depletion model is based on cell-cell movement of the positive regulator TTG from neighbouring cells, non-trichome cells. Then GL3 activator traps TTG in trichome precursor cells, promoting activator complex stability and consequently trichome cell-fate (Bouyer et al. 2008; Torii 2012). The stochastic differences in cell expression levels of genes in the activator complex, enhanced by cross-regulatory transcriptional regulation and intercellular transport of positive and negative regulators, creates a population of cells in which the activation complex reaches a threshold level which enables induction of the downstream regulator GL2 that positive regulates trichome cell fate and trichome differentiation (Grebe 2012; Pesch and Hulskamp 2009). The precise mechanism via which the balance and movement of these multiple transcriptional regulators leads to wild-type trichome spacing is still a subject of some controversy (Torii 2012).

1.4 HD-ZIP IV family

The superfamily of homeobox transcription factors in plants as recently been grouped into fourteen families characterized by different domain architectures and the conserved DNA binding homeodomain (Mukherjee et al. 2009). Four families belong to the plant specific class of HD-ZIP transcription factors (Figure 1.12) sharing the homeodomain (HD) and a leucine zipper domain implicated in protein-protein interactions. Each family is further characterized by specific combinations of protein domains (described in (Ariel et al. 2007; Ciarbelli et al. 2008; Henriksson et al. 2005; Nakamura et al. 2006; Prigge et al. 2005)). *In vitro* studies suggest that the HD-ZIP dimerization domain is essential for the transcription factor activity of these proteins, and therefore the DNA binding to pseudopalindromic motifs likely involves homo or heterodimerization which implies the potential for multiple regulatory combinations (Johannesson et al. 2001; Palena et al. 1999; Sessa et al. 1993; Tron et al. 2004).

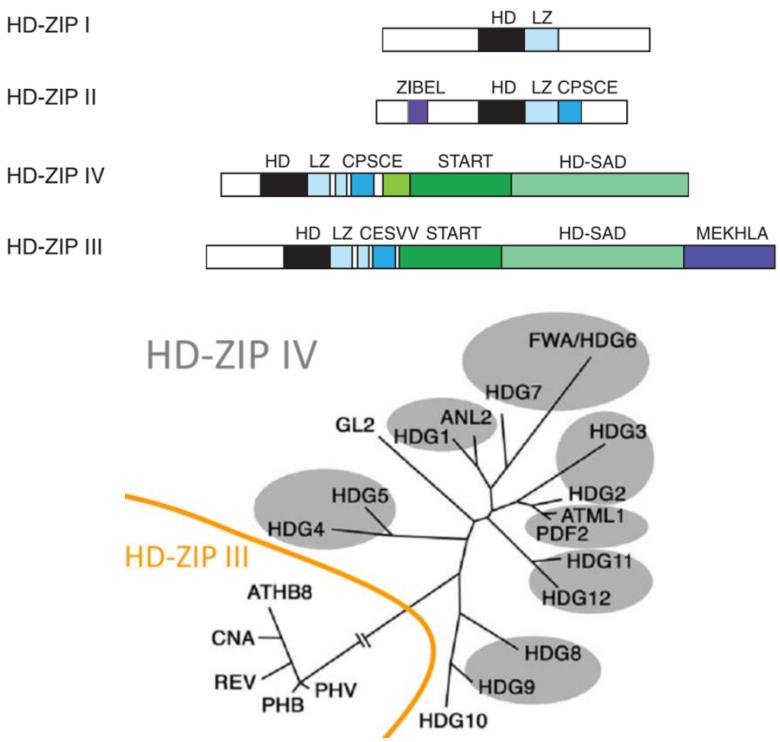


Figure 1.12 – Structure of HD-ZIP transcription factors. The four classes of HD-ZIP proteins are characterised by different combinations of motifs. All HD-ZIP genes contain a DNA binding homeodomain (black) and leucine zipper dimerization domain (blue). The HD-ZIP III and IV proteins share the lipid binding START domain and HD-SAD domain (green). The phylogenetic tree of HD-ZIP III and IV protein sequences suggest that there could be functional redundancy between pairs of HD-ZIP IVs. (Adapted from (Mukherjee et al. 2009; Nakamura et al. 2006))

As mentioned previously, several HD-ZIP transcription factors from classes II and III have important roles in embryo patterning (Section 1.2.5). In general HD-ZIP IV transcription factors are expressed specifically in the epidermal layer in the shoot, and a number of these genes are associated with epidermal specification (Abe et al. 2003; Nakamura et al. 2006). In addition, HD-ZIP II transcription factors are involved in auxin signalling, light response and shade avoidance whereas HD-ZIP III proteins participate in SAM regulation and lateral organ initiation, leaf polarity and auxin transport (Morelli and Ruberti 2002; Otsuga et al. 2001; Prigge et al. 2005; Sawa et al. 2002; Sessa et al. 2005). HD-ZIP I genes are implicated in the response to abiotic stress (water and light), photomorphogenesis, etiolation and abscisic acid signalling which is critical for many stress responses (Henriksson et al. 2005; Himmelbach et al. 2002; Olsson et al. 2004; Wang et al. 2003b). The phylogenetic tree resulting from either the full protein sequences alignment of all 47 members of *Arabidopsis* HD-ZIP class of transcription factors, or the alignment of the sequences

of the HD domain of these proteins in several plant species (Ariel et al. 2007; Mukherjee et al. 2009), reflects the structural differences between the four classes of HD-ZIP proteins (Figure 1.12). The HD-ZIP I and II families appear to be relatively closely related and significantly different from the HD-ZIP III and IV families. The later are reasonably similar, especially at the level of the homeodomain sequence. The structural similarities between the HD-ZIP I and II and the HD-ZIP III and IV families appear to be associated with similar functions since HD-ZIP I and II proteins are involved in stress response and hormone signalling whereas HD-ZIP III and IV proteins are involved in embryo patterning and plant development.

HD-ZIP III and IV proteins contain a START (steroidogenic acute regulatory protein-related lipid transfer) domain followed by the conserved SAD (START-adjacent domain). These domains have been proposed to be involved in lipid binding, thereby leading to the suggestion that the HD-ZIP III and IV proteins could be involved in the transcriptional regulation of lipid metabolism. The plant START domain is structurally similar to animal START domains although in plants, no lipid ligands have yet been identified (Schrick et al. 2004). In animals START domains function as lipid exchange or lipid sensing domains important for lipid metabolism, lipid transfer between cell compartments, or cell signalling. Mutations or misexpression of START proteins can lead to pathological disorders (Alpy and Tomasetto 2005; Soccio and Breslow 2003).

Independent studies suggest that there is extensive functional redundancy in the HD-ZIP IV family of transcription factors. The first analysis of the whole family indicates that most likely there have been duplication events in the gene family leading to the formation of several paralogous gene pairs (Figure 1.12 in grey), for instance *ATML1* and *PDF2* (Nakamura et al. 2006). As mentioned previously ATML1 and PDF2 have overlapping functions in the establishment of epidermal cell fate during embryogenesis (Abe et al. 2003). In order to test further gene redundancy, Nakamura and colleagues investigated other double mutant combinations and concluded that in general, double mutants in this gene family are either aphenotypic or do not exacerbate previously described single mutant phenotypes (Nakamura et al. 2006). In this section I will describe in detail what is

currently known about the role of HD-ZIP IV proteins in plant development, with a particular focus on epidermal differentiation post germination, and flower development.

1.4.1 HD-ZIP IV role in epidermal differentiation

A significant number of *HD-ZIP IV* genes are expressed specifically in the epidermal layer of different organs (Nakamura et al. 2006) suggesting that this family has an important role in epidermal specification. As outlined in previous sections, the differentiation of epidermal cell fates depends on complex regulatory networks involving positional signalling, transcription regulation programs, cell cycle control, and in the case of stomata, cell lineage. Positional signalling is likely to be dependent on an intact L1 identity which assures anticlinal divisions, epidermal cell-cell adhesion, and the production of an intact cuticle. It is therefore probable that there is overlap between the pathway required for L1 identity and those involved in positional signalling of epidermal cell fates. This is likely to be particularly true within the HD-ZIP IV family.

ATML1 and PDF2 are known to be key regulators of epidermal identity, and recently it has been proved that ATML1 is sufficient to induce the expression of epidermal specific genes and produce epidermal cell types in non epidermal tissues (Abe et al. 2003; Takada et al. 2013). Interestingly, the most recent of these studies have shown that at least ATML1 and the related HDG2 protein are involved both in epidermal identity and in the differentiation of specific cell fates (Peterson et al. 2013; Takada et al. 2013).

Briefly, the constitutive ectopic expression of *ATML1* using estradiol-inducible lines (Figure 1.13) proved that ATML1 is a master regulator of epidermal specification (Takada et al. 2013). When transgenic lines were grown in the presence of estradiol the overexpression of *ATML1* caused seedling lethality due to defects in root growth and SAM development. The analysis of seedlings from two independent lines showed that ectopic *ATML1* expression induced the ectopic expression of the *HDG2*, *CER5* and *FDH* epidermis specific genes, most likely via L1 box regulation. Importantly, ATML1 induced epidermal differentiation in inner layers, where

trichome and stomata like structures were detected (Figure 1.13 A, B). These specialized structures expressed appropriate cell fate markers (Takada et al. 2013). Together these results prove that ATML1 is sufficient to induce epidermal cell fate. Moreover the authors hypothesise that this gene is involved in the repression of mesophyll cell fate (Takada et al. 2013).

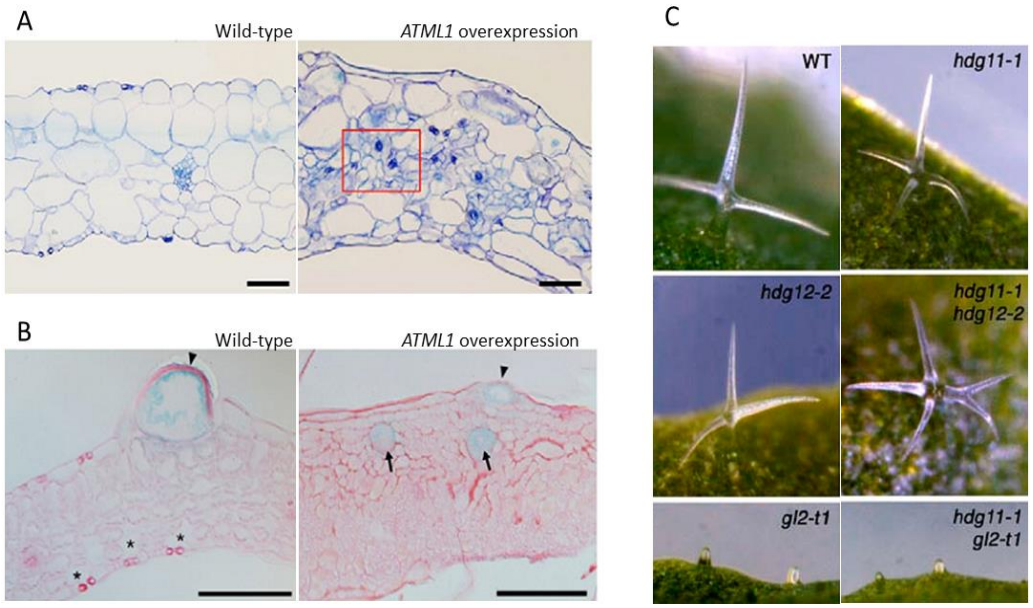


Figure 1.13 – HD-ZIP IVs are involved in trichome and stomatal development. *ATML1* overexpression induces the ectopic formation of stomata (A) and trichome-like cells (B), confirmed by the expression of the *GL2* trichome marker (blue). C) The loss of *HDG11* induces multiple trichome branching and this phenotype is enhanced by *hdg12*. *hdg11* is epistatic to *gl2*. (Adapted from (Nakamura et al. 2006; Takada et al. 2013))

In general *HDG2* expression is epidermal specific and overlaps with the expression of other *HD-ZIP IV* genes including *ATML1* and *PDF2* suggesting that *HDG2* works together with these other transcription factors. In seedlings, the *HDG2* promoter driving a *GUS* marker was expressed in SAM, leaf primordia and developing leaves. Particularly strong expression was observed in trichomes and stomatal meristemoids. Later, GUS staining was found in inflorescence meristems, carpels, ovules and seeds including embryos (Nakamura et al. 2006). Mutations in *HDG2* show a glassy trichome phenotype caused by defects in trichome cell walls, and mild defects in stomatal differentiation that are enhanced by the *atml1* mutation but not by *pdf2* (Marks et al. 2009; Peterson et al. 2013). The overexpression of *HDG2*, like that of *ATML1*, induces ectopic stomatal differentiation in the inner layers of cotyledons and

true leafs (Peterson et al. 2013; Takada et al. 2013). In addition, HDG2 was shown to activate the transcription of the *TMM* and *MUTE* genes involved in stomatal development (Peterson et al. 2013). The seedlings of *HDG2* overexpression lines occasionally displayed multiple epidermal layers again indicating that HDG2 induces ectopic epidermal fate (Peterson et al. 2013). Given the trichome phenotype of *hdg2* mutants it might be expected that HDG2 was able to ectopically induce trichome cell fate, however no trichome differentiation was observed in the inner layers of these plants, nor was *GL2* ectopically expressed (Peterson et al. 2013).

Loss of function of the *HD-ZIP IV* gene *HDG11* causes the production of multiple branched trichomes and this phenotype is enhanced in the *hdg11-1hdg12-2* double mutant (Figure 1.13 C) suggesting that HDG11 and HDG12 are involved in regulating trichome morphology (Nakamura et al. 2006). GL2 is also involved in cell-type specific differentiation in the epidermis, being required for trichome production, the differentiation of the seed coat, and the production of root non-hair cells (Di Cristina et al. 1996; Rerie et al. 1994). GL2 is expressed in non-hair cells of roots where it represses root hair cell fate. Consequently, the loss of function of GL2 disrupts root hair patterning causing an increased frequency of root hairs (Di Cristina et al. 1996; Masucci et al. 1996). The loss of GL2 also causes trichome abortion and the production of undifferentiated trichomes (Figure 1.11B) suggesting that GL2 triggers trichome differentiation (Rerie et al. 1994). The *anl2* (*anthocyaninless2*) mutant also displays ectopic root hairs, as well as extra cells between the epidermal and cortical layers, indicating that ANL2, like GL2, is involved in the maintenance of epidermal cell fate in roots (Kubo and Hayashi 2011). Finally, several genes required for epidermal cell fate have recently been associated with giant cell identity in sepals and epidermal cell fate in roots. Of these *ATML1* and *HDG11* appear to repress giant cell fate, probably due to regulation of the cell cycle (Roeder et al. 2012).

In summary, the HD-ZIP IV family is involved both in epidermal identity and in the differentiation of several cell types including stomata, trichomes and root non hair cells. Recent studies suggest that at least HDG2 and ATML1 are involved in both specifying basal L1 identity and in mediating epidermal cell differentiation,

supporting a close relationship between the regulation of these processes. The further study of HD-ZIP IV functions, targets and functional redundancy will help to clarify this connection.

1.4.2 HD-ZIP IV role in flower development

Recent studies suggest that several *HD-ZIP IV* genes are involved in flower development which is consistent with the fact that most *HD-ZIP IV* genes are expressed in flowers or siliques (Nakamura et al. 2006). In particular, *PDF2* together with *HDG1, HDG2, HDG5* and *HDG12* could be involved in floral organ identity whereas *HDG3* possibly participates in anther dehiscence (Kamata et al. 2013; Li et al. 2007).

The first analysis of *HD-ZIP IV* family included the study of expression patterns of *HD-ZIP IV* promoters and a selection of genetic interactions between paralogous genes that highlighted gene redundancy in this family, since most combinations were aphenotypic (Nakamura et al. 2006). Recently the same laboratory published intriguing results indicating that loss of function of several *HD-ZIP IV* genes disturbs flower organ identity (Kamata et al. 2013). The phenotypes of *pdf2-1* knockout lines combined with *hdg1-1*, *hdg2-2*, *hdg2-3, hdg5-1* or *hdg12-2* displayed sepaloid petals, carpelloid stamens or infertile stamens with different degrees of severity (Figure 14), whereas the strongest phenotypes were found in *pdf2-1hdg2-3*, the *pdf2-1hdg12-2* displayed weaker phenotypes. The expression of *APETALA3(AP3)* MADS box transcription factor encoding gene, involved in sepal and stamen identity, was reduced in these double mutants (Kamata et al. 2013). In particular, this work showed variation in the frequency and length of stamens in flowers and distinct stamen phenotypes in single mutants and double mutants. With the exception of the stamen frequency in flowers, the other reported phenotypes were absent in single mutants and, intriguingly, in the same combinations of *hdg* mutations with the *pdf2-2* allele or double mutants between *pdf2-1* and *hdg1-1* or *hdg5-2* (Kamata et al. 2013). Moreover the combinations between *atml1-1* or *atml1-2* and the above-mentioned *hdg* alleles had no effect in flower organ identity. The discrepancy in phenotypes between alleles of the same gene is disturbing. It is possible that some of the T-DNA insertion mutants used produce stable but incomplete transcripts. For instance, using

primers that target the sequence upstream of the *pdf2-1* T-DNA insertion, the *PDF2* transcript detected by RT-PCR for all mutant combinations was at the same level as in wild-type. Moreover, *in situ* hybridization experiments proved that *PDF2* expression in the *pdf2-1* allele in inflorescences and flower meristems is epidermis specific as in wild-type (Kamata et al. 2013). Both *pdf2-1* and *pdf2-2* contain T-DNA insertions which disrupt the START domain, and therefore they should display the same phenotype. One possibility is that *pdf2-1* acts as a dominant negative or positive allele. The study of further *PDF2* alleles should clarify this question.

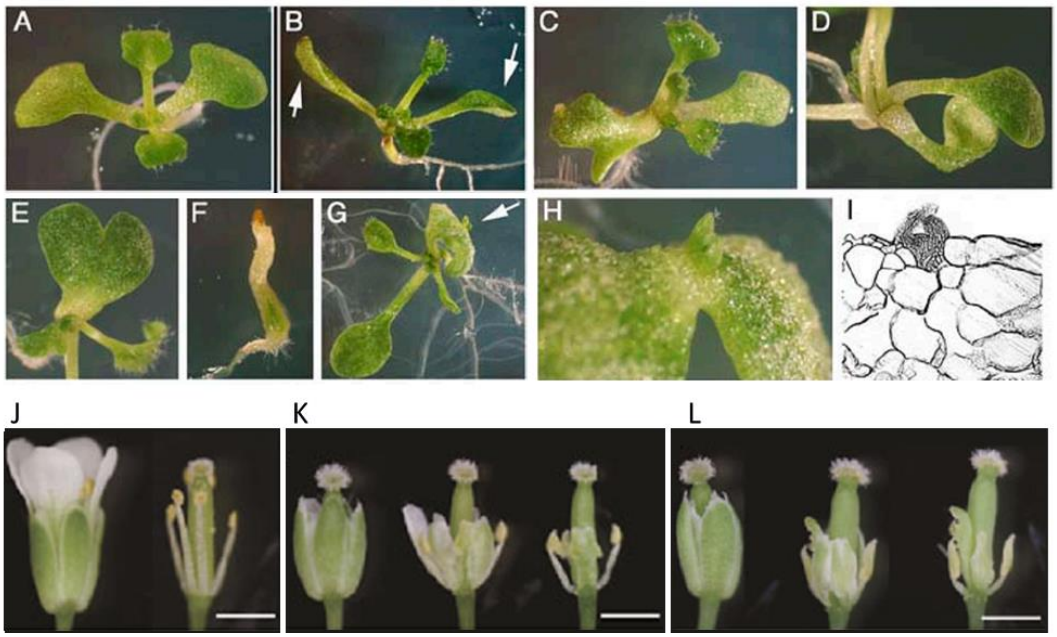


Figure 1.14 – HD-ZIP IV mutant phenotypes. Seedlings of *hdg2-3hdg3-1* (A) display wild-type phenotypes, whereas *atml1-1hdg3-1* (B) seedlings show developmental problems including cotyledon hyponasty (white arrows). The *pdf2-1hdg3-1* double mutant (C-I) shows various phenotypes ranging from misshapen cotyledons, to pin phenotypes (F) and adventitious shoot development (G-I). Moreover compared to the wild-type flowers (J), pdf2-1hdg2-3 (K) and pdf2-1hdg1-1 display abnormal flower phenotypes. Scale bar is 1mm. (Adapted from (Kamata et al. 2013; Nakamura et al. 2006))

The screen for mutants displaying hyponastic phenotypes identified the mutant *ucl1* (*upcurved leaf1*) an *HDG3* overexpression line (Li et al. 2007). Besides the upcurved leaves, the plants are sterile due to defects in anther dehiscence, and consistent with this phenotype, anther dehiscence regulators are downregulated in the *ucl1* mutant background (Li et al. 2007). While the expression of *pHDG3-HDG3* intriguingly showed the same phenotype as *ucl1*, the ectopic and constitutive expression of

HDG3 resulted in the exacerbation of the ucl1 phenotype (Li et al. 2007). Therefore, it is possible that HDG3 plays a negative role in anther dehiscence and is involved in leaf morphogenesis. Although the *hdg3-1* mutation is aphenotypic, the seedlings of *hdg3-1atml1-1* display hyponastic cotyledons and *hdg3-1pdf2-1* double mutants were reported to display several developmental problems (Figure 1.14) with morphological cotyledon defects, cotyledon fusions, and occasionally adventitious shoot formation (Nakamura et al. 2006). These results suggest that HDG3 possibly acts together with ATML1 and PDF2 in seedling development. However, it should be noted that the *pdf2-1* allele was again used in these experiments.

1.4.3 Overview of HD-ZIP IV family and future perspectives

The summary of the published work on HD-ZIP IV genes is included in Table 1.15. Additional roles of HD-ZIP IV genes include anthocyanin production, stress response and lipid metabolism. These are summarized briefly below.

Table 1.15 – Characterization of the HD-ZIP IV family. The abbreviations Sd, Rt, Lf, Fl and Sl stand for seedlings, root, leaf, flowers, and siliques respectively. (Di Cristina et al. 1996; Kamata et al. 2013; Kinoshita et al. 2004; Kubo and Hayashi 2011; Kubo et al. 1999; Li et al. 2007; Marks et al. 2009; Nakamura et al. 2006; Peterson et al. 2013; Rerie et al. 1994; Roeder et al. 2012; Yu et al. 2008)

Gene	RT-PCR	Phenotypes
ANL2	leaves, stems, buds, and roots	− *anl2* phenotype consists in reduced anthocyanin in subepidermal cells and epidermal cells in the shoot − *anl2* shows aberrant root structure, in particular extra cells in the root between the epidermal and cortical layers
HDG1	Sd, Rt, Lf, Fl, Sl	− *hdg1-2 anl2* showed no additional phenotypes to those reported for *anl2* − *pdf1-1hdg1-1* display homeotic conversion of flower organs
HDG2	Sd, Rt, Lf, Fl, Sl	− Mutations of HDG2 result in abnormal trichome cell walls *hdg2-2* − *pdf1-1hdg2-2* and *pdf1-1hdg2-2* display homeotic conversion of flower organs − Overexpression of HDG2 induces ectopic production of stomata and multiple cell layers
HDG3	Sd and Sl	− *hdg3-1 atml1-1* and *hdg3-1 pdf2-1* have defects in cotyledon development − Overexpression of *HDG3* induces leaf hyponasty and defects in anther dehiscence
HDG4	after flower development	No discernible growth and development phenotypes shown in mutant lines
HDG5	Sd, Lf, Fl, Sl	− *pdf1-1hdg5-1* display homeotic conversion of flower organs
FWA	Endosperm	*fwa*-1 shows complete absence of cytosine methylation in *FWA* promoter and coding regions. This hypomethylation will activate the gene which in turn leads to late flowering
HDG7	Rt and Sl	No discernible growth and development phenotypes shown in mutant lines
HDG8	after flower development	No discernible growth and development phenotypes shown in mutant lines
HDG9	after flower development	No discernible growth and development phenotypes shown in mutant lines

HDG10	after flower development	No discernible growth and development phenotypes shown in mutant lines
HDG11	Sd, Rt, Lf, Fl, Sl	− *hdg11-1* appeared to have more branched trichomes; this phenotype is enhanced in *hdg11-1 hdg12-2* plants − *hdg11-3* sepals display reduced giant cell frequency − Overexpression of HDG11 confers drought resistance
HDG12	Sd, Rt, Lf, Fl, Sl	− *hdg11-1 hdg12-2* have more branched trichomes
GL2	Sd, Rt, Lf, Fl, Sl	− mutations in *GL2* specifically alter the differentiation of the hairless epidermal cells, causing them to produce root hairs − *gl2* shows abnormal trichome expansion

The *ANL2* loss of function reduces the anthocyanin in subepidermal and epidermal cells in shoots, while ATML1 overexpression causes the accumulation of anthocyanin (Kubo et al. 1999; Takada et al. 2013) suggesting that both genes act together in regulating anthocyanin production. Since anthocyanins are typically produced in response to stress, these observations provide an intriguing link to stress responses. A further link is provided by the fact that the overexpression of *HDG11* confers drought resistance due to a reduction in stomatal density, higher levels of abscisic acid and an enhanced root system (Yu et al. 2008).

The identification of *cfl1* (*curly flag leaf1*) mutant in rice resulted in the identification of *HDG1* as having a potential role in cuticle development. The overexpression of *AtCLF1* or *OzCLF1* in *Arabidopsis* resulted in severe defects in cuticle development similar to the *cfl1* mutant in rice. Since CFL1 and HDG1 were found to interact physically it was suggested that AtCFL1 and HDG1 work together in regulating cuticle development. In addition HDG1 binds to the L1box sequence and in so doing possibly regulates the expression of *BDG* (*BODYGUARD*) and *FDH* genes involved in cuticle formation (Wu et al. 2011). The study of cuticle composition in the *anl2* mutant also showed that cutin levels are lower than in wild-type plants, and mutation of the tomato homolog of *ANL2*, *CD2* (*CUTIN DEFICIENT2*) causes phenotypes associated with defects in epidermal cell fate, including severe cuticle phenotypes (Nadakuduti et al. 2012). The hypothesis that *HD-ZIP IV* genes are involved in cuticle biosynthesis and lipid metabolism is supported by the identification of OCL1 targets in maize which are involved in these two processes (Javelle et al. 2010), and the presence of the L1 box binding site in the promoters of several genes involved in cuticle biosynthesis. Possible links between cuticle formation and the maintenance of epidermal cell fate were discussed earlier in this introduction. It is therefore tempting to speculate that the formation of the cuticle

could provide part of a feedback mechanism, relying on the activity of *HD-ZIP IV* genes, which could act throughout the development of the plant.

1.5 Aims and objectives

The goal of this project is to clarify the mechanisms involved in epidermal specification during embryogenesis. The work starts from the premise that the central core of epidermal specification is the transcriptional regulation mediated by the transcription factors ATML1 and PDF2. I aim to understand more about the signalling pathways acting upstream of these transcription factors, to regulate their epidermis-specific expression. In particular I will investigate their relationships with the receptor kinase ACR4. I also aim to uncover the identities of potential new regulators of epidermal cell fate in plants and, by so-doing, to understand more about the mechanisms and pathways involved in epidermal cell fate establishment and maintenance.

Currently, it has been proposed that epidermal cell fate specification requires a positive feedback loop. ATML1 and PDF2 have been proposed to positively regulate their own expression and to activate the expression of *ACR4*. In turn ACR4 mediated signalling could activate the transcription of ATML1 and PDF2 although this has yet to be proved. I aim to use a variety of different strategies including mathematical modelling, to dissect this potential feedback loop.

In the search for further regulators of epidermal cell fate, the *HD-ZIP IV* gene family is an obvious starting point. The identification of transcriptional targets of ATML1 and PDF2 will involve the analysis of experimental data obtained by ChIP (Chromatin ImmunoPrecipitation) and transcriptomic analysis coupled with genome-wide search for the presence of potential DNA binding sites like the L1 box. Another important objective of this project is the identification of HD-ZIP IV transcription factors that could act together with ATML1 and PDF2 in epidermal cell fate using both a genetic and a biochemical approach.

Ultimately I will try and integrate the data produced in this work in order to describe the regulatory network that defines epidermal cell fate in embryo.

Chapter 2 – Materials and methods

2.1 Plant material and growth conditions

Most of the mutant lines used in this project are in the Col0 background. The mutant alleles of *HD-ZIP IV* genes (described in Table 3.2 in Chapter 3) were used to investigate genetic interactions amongst *HD-ZIP IV* family and between *HD-ZIP IV* mutants and *acr4-2*. Additionally, mutant lines of four *WOX* genes and *ACR4* were also used to determine genetic interactions between *ACR4* and *WOX* family (Table 2.1). For ChIP experiments (Section 2.5) I used two lines developed by the Ingram group before my arrival, namely *pATML1-GFP:ATML1* (Gifford et al. 2005) and a construct that includes *FDH* promoter and *fdh* 5'UTR driving the fusion gene *GFP:ATML1*. For Co-Immunoprecipitation (CoIP) experiments I used transgenic plants expressing different combinations of *GFP-PDF2*, *GFP-ATML1*, *FLAG-PDF2*, and *FLAG-ATML1* under *PDF1* promoter (Section 2.2.3 and 2.6) and in addition *pATML1-GFP:ATML1* and *pACR4-ACR4:GFP* (Gifford et al. 2003; Gifford et al. 2005).

Table 2.1 – *wox* and *acr4* mutant alleles used in this project.

Gene	Allele	Background	Reference
ACR4	*acr4-2*	Col0	(Gifford et al. 2003)
WOX1	*wox1-1*	Col0	(Vandenbussche et al. 2009)
WOX2	*wox2-1*	WS	(Breuninger et al. 2008; Haecker et al. 2004)
	wox2-2	Col0	(Breuninger et al. 2008; Haecker et al. 2004)
PRS/WOX3	*prs*	Col0	(Vandenbussche et al. 2009)
WOX8	*wox8-2*	WS	(Breuninger et al. 2008)

Two alternative methods were used for seed sterilization. In the first method seeds were kept in a solution of 70 % ethanol and 0.05 % Triton X100 (Sigma, USA) for 15 min, mixed every 5 min by inverting the tubes. Then they were rinsed three times in a fresh solution of 96 % ethanol and pipetted onto sterile filter paper (previously imbibed in 96 % ethanol and dried under a sterile hood) until they dried. Alternatively, the seeds were kept in a Cl_2 saturated atmosphere for at least 2 hours under a chemical hood. Seeds were transferred into open 1.5 ml tubes and placed, together with a small beaker, containing 50 ml of bleach (NaCLO 2.6 % (Oxena,

France)) and around 5 ml of concentrated hydrochloric acid (37 %), inside a sealed container.

For the majority of experiments, sterilized seeds were sprinkled on plates containing MS (Murashige and Skoog) medium (1x MS basal salt mixture,), 0.8 % Plant agar, and 0.5 % Sucrose with appropriate antibiotics (all four reagents from Duchefa, The Netherlands). Seeds were stratified for 3 days at 4 °C before transferring to a growth chamber under long day conditions (16 h light/ 8 h dark and 21 °C) for seven days. The seedlings were transferred into soil (Favorit "Argile10" soil mixture, Germany) and grow in long day conditions (16 h light at 20.5 °C, 50 % humidity and 8 h dark at 19.5 °C, and 55% humidity). For the first few days the trays were covered in plastic bags to increase the humidity level.

For Hygromycin selection, plates contained 0.3 % of sucrose MS medium and were grown in long day conditions for 4 hours, followed by four days under dark conditions (plates covered in aluminium foil). After a further 3 days under long day condition, seedlings were transferred into soil.

For confocal microscopy on plant meristems, plants were grown in short day conditions (8 h light at 18.6 °C, 70 % humidity and 16 h dark at 16 °C, and 79% humidity) for 4 to 6 weeks and then moved to long day chambers. For these experiments, sterile seeds were kept at 4 °C in a 0.1% agarose (Euromedex, France) solution for three days to stratify. Then they were sown directly onto soil and transferred to the short day room. After 10 days, seeds were thinned leaving 1 or 2 seedlings per pot.

Genetic crosses were carried out using a two-step procedure, starting with the emasculation of flowers of the acceptor (female) plants and followed 24h later by the pollination of carpels using flowers of the donor plants. Once the siliques were mature they were harvested and stored in appropriate envelopes.

2.2 DNA techniques

2.2.1 DNA extraction

The method of extraction of genomic DNA was chosen according to different applications and number of samples. For plant genotyping I extracted DNA from small leaf samples that were collected in microfuge tubes (Eppendorf, Germany) or collection tubes placed in racks of 96 wells (Qiagen, Germany) and were immediately frozen in liquid nitrogen and kept in -20 °C. The first step of DNA extraction protocol consisted of leaf grinding. For small numbers of samples, microfuge tubes were kept in liquid nitrogen, and using pre-cooled plastic pestles, the frozen leaves were crushed and ground to a fine powder. For a large number of samples, I used small metal beads (Qiagen, Germany) and performed the grinding step in a TissueLyser II (Qiagen, Germany), performing two cycles of 30s at 30 Hz.

For rapid DNA extractions (one-off genotyping) 400 µl of extraction buffer (100 mM Tris HCl pH 8, 0.5 M NaCl, 50 mM EDTA and 0.7 % SDS) was added to the samples which were then vortexed, incubated for 15 min at 37 °C, and centrifuged (Sigma, USA, 2-16 K Centrifuge) for 15 min at 4000 rpm. 300 µl of supernatant was transferred to new tubes, and 300 µl of isopropanol was added. Samples were mixed by inverting the tubes several times and left at room temperature for 2 min. Samples were then centrifuged for 15 min at 4000 rpm, the supernatant was removed. The pellet was washed with 70 % ethanol and the pellet dried overnight at room temperature. Finally the pellet was dissolved in 100 or 500 µl of water or 100 mM Tris HCl (protocol adapted from (Edwards et al. 1991)). Alternatively, for large sample numbers, I used the BioSprint 96 kit (Qiagen, Germany) following the respective protocol. DNA samples were stored at -20 °C.

2.2.2 Polymerase chain reaction (PCR)

PCR reactions were performed in a 2720 Thermalcycler (Applied Biosystems, USA) machines. For genotyping I generally followed the following PCR program: 5 min at 95 °C followed by 30 cycles of 30 s at 95°C / 30 s at 55-60 °C / 1 min per 1000 bp at 72 °C and the final extension step of 7 min at 72 °C. Go-Taq polymerase was used

following the manufacturer's instructions (Promega, USA). The annealing temperature was chosen as a function of the lowest Tm of the pair of primers.

Table 2.2 – Description of the primers used for PCR. Primers used for plant genotyping, ChIP experiments and cloning are shown in black, blue and grey respectively.

HD-ZIP IV	*atml1-3*	Atml1L atgtatcatccaaacatgttcgaatctc / Atml1R tccttgtacctattgttctctgctc
	pdf2-3	Pdf2L gaaagctaaaccggactgaacc / Pdf2R ctgagtttcaagttccatcaccc
	anl2-2	Anl2L aagagagacatcaccaagcatgg / Anl2R gaagaacatgagagtagatctgg
	hdg1-3	Hdg1L agtacgacacaagtttctctcgg / Hdg1R atcaattgagacatccacgactgc
	hdg2-2	Hdg2L gtggaagagtttggttttagacg / Hdg2R tttcacgtgttgggactaatgg
	hdg3-1	Hdg3L aacaacggagacaacaacaacg / Hdg3R agaatttgatctgaacagggtcc
	hdg4-2	Hdg4L ctcaatcttgttttggtatccg / Hdg4R tccttcgagaaccctaatttcc
	hdg5-3	Hdg5L tcttatccgtgatcaacatcacc / Hdg5R aagaatcttcaccgttcatcgc
	fwa-t1	Hdg6L cagaaaaggaaagaagaaaggaagc / Hdg6R tagtgccagatggagtaaatggg
	hdg7-2	Hdg7L gtgtgatcttcagaagacaaaagc / Hdg7R ttatacaagacataggcaacggc
	hdg8-1/2	Hdg8L tacgagccatctagctcatatgg / Hdg8R atccgagcacatgaattgtgttgg
	hdg9-3	Hdg9L atgacgagcctatcatcaatagacg / Hdg9R aaacagtacggaggtcactcaatgc
	hdg10-2	Hdg10L tggaagtgactgataaagtgtgg / Hdg10R tattgttggctccagactgttgc
	hdg11-2	Hdg11L ttattaaccgtttccaccgactcc / Hdg11R agcaacaacatgcttattctgc
	hdg12-2	Hdg12L ttagctgcaatactcgaaaccc / Hdg12L ttagctgcaatactcgaaaccc
ACR4	*acr4-2*	acr4 F ttctccttgtaaagcttgtgc / acr4 R attccacaagcataaaacttccc
WOX genes	*prs*	AtPrsF gggaactggagtaggagaagctc / AtPrsR catccaatctcgaccgtacgatgag
	wox2-1,2	m-s aagtaaacgcaggaacagcaagcagttca / wox2-as cgaaacgagtagaagtagaac caccagaa / m-as cgaaacgagtagaagtagaaccaccagaa
	wox8-2	wox8-2s cacctagagaggagattcaaagaatccg /wox8-as ccaccaactcattgttttgaaccg/ wox8-Tdna cattttataataacgctgcggacatctac
T-DNA	SAIL gcttcctattatatcttcccaaattaccaataca, SALK attttgccgatttcggaac, GABI KAT atattgaccatcatactcattgc, and WISCONSIN aacgtccgcaatgtgttattaagttgtc	
ChIP	*pACR4*	L1ACR4F gagaaaccaaaagagaaagagag / L1ACR4R cacctacatttgccattattcag
		ACR4 F agatcaatgttgtacaccatgtcg / ACR4R tccgtctttaaagttgacttgagacc
	pPDF1	L1PDF1F ttgaagtaaatacaaacaggtgg L1PDF1R tagtaatacggactgagagagag
		PDF1F cttgctgcatcttcttctttgg / PDF1R gatttggcgattttcgaactgg
	pDEK1	L1DEK1F cactcaatattttatccaatcc / L1DEK1R tatgagctttgagcctataacc
		DEK1F gagttttcaaagaaacttaagcc / DEK1 R atcttgtcttgtgaattcacc
	pPDF2	L1PDF2F tattgatgagcgcatagagagg / L1PDF2R acacattagtatatgtggaagc
		PDF2F tacctgttactagtcatagcc / PDF2R cgatgcaatttgtatcctcc
	pPAS1	L1PAS1F ttttctctgtgttatactcgg / L1PAS1R aacagatagtgacagaagaggg
		PAS1F gaggaaagaacatcctaaattcc / PAS1 R ctatatgacctgagttttgttgc
	pFDH	L1FDH tatgtcactcttaattcctccg / L1FDHR aaccaaaccaaagttactagc
		FDHF cttacaattaatcgattggtgg / FDHR ggtagtaataatttaggtggg
	pGDH	L1GDHF gattatccgttcacattgaaggg / L1GDHR acgtattgtagaatcttgagctcgc
	(-control)	GDHF aacagatagacagttagttgg / GDHR atgtttgggtaagattcagtggagg
	Actin	actinF cgtttcgctttccttagtgttagct / actinR agcgaacggatctagagactcaccttg
ACR4 promoter	L1mut1F attcagaagcagcagttttggtcgttattaatgccatctcagtacttcatg/ L1mut1R catgaagtactgaga Tggcattaataacgaccaaaactgctgcttctgaat / L1mut2F agaagcagcagttttggtcgcattctctgccatctcag tacttcatgac/ L1mut2R gtcatgaagtactgagatggcagagaatgcgaccaaaactgctgcttctg pACR4-RS1 Gaggttagcttacagatcaatg / pACR4-RS2 tctgaaccatatagttagttggg pACR4-RS3 tgaaatcttgagtctaaagaggac / pACR4-RS4 tatgtgcatgaaacctcacaagc pACR4-RS5 aaaggaagctttgaaggaatgag / pACR4-RS6 aaacttttgggttagattgcac pACR4rev-RS1 catgagctactgttgaatgtataa / pACR4rev-RS2 cccaattactctctcaacac		
Cloning primers	gATML1B2r ggggacagctttcttgtacaaagtggcaatgtatcatccaaacatgttcgaatc gATML1B3 ggggacaactttgtataataaagttgaaagtagaagagaatctcaagtttgg gPDF2B2r ggggacagctttcttgtacaaagtggcaatgtaccatccaaacatgtttgag gPDF2B3 ggggacaactttgtataataaagttgttcaggttaagaacacacatcattgcc FLAGf ggggacaagtttgtacaaaaaagcaggctcaacaatggactacaaggacgacgatgacaaaaacccagctttct tgtacaaagtggtcccc / FLAGr ggggaccactttgtacaagaaagctgggtttttgtcatcgtcgtcgttgtagtccattg ttgagcctgctttttgtacaaacttgtcccc pPDF1B4 ggggacaactttgtatagaaaagttgatagcggaatagctggcaac pPDF1B1r ggggactgctttttgtacaaacttggagagaagtttgttgcaaatgg		

The primers were designed with a GC content between 40 and 60 %, primer length between 20 and 25 nt and a 3'end enrichment in G or C bases. I then used the program NetPrimer (USA), available online, to look for secondary structures between the pair of primers, namely hairpins, repeats, self-dimers and cross-dimers. A standard PCR reaction mix contained 4 µl Green GoTaq Flexi Buffer (Promega, USA), 0.2 µl of 20 mM dNTP mix (20 mM dATP, dTTP, dCTP and dTTP (Promega, USA)), 0.1 µl GoTaq polymerase (Promega, USA), and 1 µl of each primer (10 µM stock) for 20 µl PCR reaction. The primers used in this project are described in Table 2.2.

For cloning, I used the high fidelity polymerase (Phusion Hot Start from Finnzymes, Finland) to ensure fidelity. The master mix contained 4 µl Phusion HF Buffer (Finnzymes, Finland), 1 µl of each primer (10 µM stock), 0.2 µl dNTP 20 mM (Promega, USA) and 0.2 µl Phusion Hot Start DNA polymerase (Finnzymes, Finland) for 20 µl reaction. The program used was as follows: 30 s at 98°C, and 30 cycles of 10 s at 98 °C / 30 s at 60-72 °C (depending on primer sequence) / 15-30 s for each 1000 bp of the PCR product, and a final extension step of 5-10 min at 72°C.

DNA separation

The products of restrictions or PCRs were separated by electrophoresis on agarose gels. In general, the gels contained 1 % agarose dissolved in TAE buffer (0.04 M Tris acetate and 0.001 M EDTA at pH8) and a small amount of EtBr (~0.1 µg/ml). The gel was poured in appropriate moulds (Fisher Scientific, USA) and was run in TAE buffer. DNA samples were mixed with an appropriate volume of 6x loading buffer (0.5 % Organge G, 50% Glycerol, 49.5% TE buffer pH8, (Sigma, USA)) before loading on gels. Additionally, I used a 1 kb size marker mix (Smartladder) or 100 bp ladder (both from Promega, USA). The electrophoresis was performed at room temperature at 80 to 160 volts depending on the size of the gel. DNA fragments were visualized on transilluminator under UV light.

PCR products were purified using a NucleoSpin Gel and PCR Clean-up kit (Machery-Nagel, Germany) following the manufacturer's instructions.

2.2.3 Plasmid construction

DNA restriction

Restriction enzymes (New England Biolabs, USA, or Invitrogen, USA) were used to assemble and test plasmid constructs. Restrictions were carried according to the manufacturer's protocol (choice of appropriate buffer, reaction temperature and BSA requirement). Digestions were incubated for 1 to 3 hours at the specified temperature. Double digestions were made simultaneously when the both enzymes required the same reaction buffer. Alternatively sequential digestions were carried out, which implied the purification of the DNA fragment resultant from the first digestion followed by the second digestion. DNA purifications were carried out using the Nucleospin Gel and PCR clean-up kit from Machery-Nagel (Germany) following the manufacturer's instructions.

The DNA ligation of fragments obtained by restriction was carried out using T4 DNA ligase (New England Biolabs) following the manufacturer's protocol. Ligation reaction was performed overnight at 16°C.

Transgene construction using GATEWAY technology

The cloning strategy to obtain constructs expressing *PDF2* or *ATML1* gene fusions under the *PDF1* promoter first involved the amplification of DNA fragments and recombination into pDONR (Invitrogen, USA) vectors. The genomic region of *ATML1* or *PDF2* was amplified using gATML1B2r/gATML1B3 or gPDF2B2r/ gPSD2B3 primer pairs. PCR products were purified as described previously and recombined into pDONR P2R-P3 (Invitrogen, USA) to obtain pENTR-ATML1-L2R-L3 and pENTR-ATML1-L2R-L3 respectively. BP clonase reactions (Invitrogen, USA) were carried out following manufacturer's instructions. The promoter of *PDF1* was amplified with pPDF1B4/pPDF1B1R and recombined into pDONR P4-P1R (Invitrogen, USA) to obtain pENTR-pPDF1-L4-L1R (this clone was made by Dr Pradeep Das prior to my arrival in the laboratory). I used the ENTR-GFP-L1-L2 containing GFPS65C (Fobis-Loisy et al. 2007; Heim et al. 1995). The amplification of the FLAG tag involved the denaturation at 96 °C for 5 min of the

primers FLAGf and FLAGr followed by a re-annealing step of 10 min at room temperature. The product was then recombined directly into pDONR 221 (Invitrogen, USA) to obtain pENTR-FLAG-L1-L2. Triple LR combinations were carried out into destination vectors to obtain N-terminally tagged versions of ATML1 or PDF2 driven by the *PDF1* promoter in binary vectors. I used two destination vectors: pART27 (Gleave 1992) and pH7m34GW (Karimi et al. 2002) which carry Kanamycin or Hygromycin resistance respectively in plants.

ACR4 promoter fragments were amplified, purified and cloned into pENTR/D-TOPO (Invitrogen, USA) following manufacturer's instructions, so they could be recombined into pKGWFS7 (Karimi et al. 2002). The *ACR4* promoter was amplified using different combinations of primers in order to obtain two sets of six fragments of different lengths. The two sets were obtained using two different reverse primers corresponding to the 3'end of the promoter, pACR4rev-RS1 or pACR4rev-RS2. The six fragments were obtained using the 5' primers pACR4-RS1 to pACR4-RS6 which are complementary to different sites in the *ACR4* promoter. For the mutated versions of the full length promoter, two overlapping fragments were amplified using two pairs of primers pACR4-RS1/L1mut1R (or L1mut2R) and L1mut1F (or L1mut2F)/pACR4rev-RS1to obtain 5' and 3' fragments of *ACR4* promoter with the required mutation in the L1 box region. These fragments were purified, mixed, and subjected to a second round of PCR of 19 annealing cycles, using and pACR4-RS1/pACR4rev-RS1. The DNA products were then cloned into pENTR/D-TOPO (Invitrogen, USA) and recombined into pKGWFS7 (Karimi et al. 2002). Ultimately I obtained 14 constructs corresponding to different versions of ACR4 promoter driving GUS and GFP markers.

Transformation of plasmids into E. coli

The products of ligations were transformed by heat-shock. First 25 µl of thermo-competent TOP10 cells (Invitrogen, USA) were thawed on ice, the pre-chilled ligation was added to the cells and they were incubated for 30 min on ice. A heat-shock was performed at 42 °C for 1 min and the tubes were placed on ice

immediately for 5 min. 1 ml of LB medium (1% tryptone, 0.5 % yeast extract, 0.5 % NaCl at pH 7) was added to the cells before incubation at 37 °C for 1 hour.

After bacterial transformation, 100 µl of cells were plated in LB selection plates containing appropriate antibiotics. The remaining cells were centrifuged at 7000 rpm for 3 min and the pellet was resuspended in 100 µl of LB and plated in LB selection plates. Finally the plates were incubated overnight at 37 °C, upside-down. The selection plates contained LB medium and 100 µg/ml of Ampicillin (Duchefa, The Netherlands), 50 µg/ml of Kanamycin (Duchefa, The Netherlands) or 250 µg/ml of Spectinomycin (Duchefa, The Netherlands).

Plasmid purification and sequencing

DNA minipreps were performed using NucleoSpin Plasmid kit (Macherey-Nagel, Germany) following the manufacturer's instructions. The plasmids were then quantified using a NanoDrop ND-1000 UV-Vis spectrophotometer (Thermo Scientific, USA) system. For sequencing, samples of 30 to 100 ng/µl of DNA in a volume of 30 µl were processed by GATC Biotech (Germany) using an ABI 3739xl sequencer (Applied Biosystems, Germany). The sequencing results were analysed using DNASTAR software (Lasergene, USA).

2.3 Plant transformation

The protocol used for *Agrobacterium*-mediated plant transformation via floral dipping was adapted from (Clough and Bent 1998). In this project all constructs were initially transformed into Col0 wild-type plants, and transferred to other genotypes by crossing. Plants (50 seedlings per transformation) were grown in long day conditions in a container measuring 12/20 cm. Plants were transformed a few days after the appearance of the first open flowers (around 3 weeks after seedling transfer from plates).

Binary plasmids were transformed into electrocompetent cells of the *Agrobacterium* strain C58/pMP90. To prepare electrocompetent cells, LB plates were inoculated from glycerol stock of *Agrobacterium* and incubated at 28 °C for 2 to 3 days. One colony was resuspended in 2 ml LB medium and incubated at 28 °C with agitation

for 6 hours. 100 µl of this pre-culture was subsequently transferred into 100 ml of 0.1 % sucrose LB medium, and incubated overnight at 28 °C with agitation until the OD_{600} reached 1-1.5. The culture was cooled on ice for 15 min and then centrifuged (Sorvall Evolution RC from Thermo Scientific, USA) for 20 min at 4°C and 4000 g. The pellet was resuspended in HEPES 1mM, pH 7. This washing step was repeated three times. Subsequently the pellet was washed in 10 ml of 10 % glycerol and resuspended in a final volume of 500-750 µl of 10 % glycerol. Aliquots of 40 µl were flash-frozen in liquid nitrogen and kept at -80.

Before electroporation, electrocompetent bacterial cells and electroporation cuvettes were kept on ice for 10-15 min. 1-5 µl of plasmid were added to the cells which were immediately transferred to the base of the electroporation cuvette. Electroporation (1 pulse of 2.2 kV) was performed using a MicroPulser Electroporator (Bio-Rad, USA) and 1 ml of LB was immediately added to the cuvette. The cells were then transferred gently to an Eppendorf tube and incubated for a period of 2 hours at 28 °C with shaking. Finally 100 µl of cells were plated on LB plates containing 50 µg/ml Rifampicin, 20 µg/ml Gentamycin, 250 µg/ml Spectinomycin and incubated at 28 °C for 2 to 3 days. Colonies were verified for presence of binary vectors by PCR and then one colony for each construct was transferred into a sterile 100 ml flask containing 10 ml of LB medium (with antibiotics) and incubated overnight at 28 °C under 200 rpm/min agitation. The evening of the next day, 10 ml of this culture were transferred into 2000 ml flasks containing 800 ml of LB medium (with antibiotics) and incubated under the same conditions. The following day the OD_{600} was monitored until it reached 1.6 to 2. The cells were then centrifuged at 5000 rpm for 15 min and the pellet was resuspended in the infiltration medium (5 % Sucrose, 10 mM MgCl2, 0.02 to 0.05 % Silwet (Agridyne, France) to obtain an OD_{600} of 0.8. The solution was poured into 1.5 L aluminium containers in which plant inflorescences were immersed for 2 min. The plants were then closed into a plastic bag overnight and the following day bags were opened and subsequently removed. Plants were allowed to set seeds, and collected seeds were sterilized using several washes of 70 and 96 % ethanol solutions (30 min total in 70% Ethanol and 5 min total in 96% ethanol), and dried on sterile filter paper under a sterile hood. They were then spread thinly over MS plates containing, 0.5 % sucrose, 50 µg/ml of Kanamycin or 30

µg/ml Hygromycin and 10 µg/ml of Amphotéricine B (to kill remaining bacteria). The seeds were stratified for 3 days at 4 °C and grown in long-day conditions. Transformed individuals were selected from plates and transferred to soil (T0 plants). Seeds from these plants (minimum 15 individuals per transformation) were sown on MS plates containing appropriate antibiotics and lines segregating 75% antibiotic resistant seedlings were retained and brought to homozygosity over subsequent generations.

2.4 Quantitative gene expression analysis

Seedling material for Q-PCR was grown in liquid culture. For each sample 10 sterile seeds were place into each well of a Costar 6 cell culture plate (Corning, USA) with 2.5 ml of MS medium. Four biological replicates per genotype were distributed randomly in the plates. The plates were then sealed with micropore tape and seeds were stratified for 3 days at 4 °C. The plates were incubated for 10 days in long day growth chambers, with horizontal circular agitation. Seedlings were then collected, blotted carefully on tissue paper and flash-frozen in liquid nitrogen in 2 ml Eppendorf tubes with two large metal beads (Qiagen, Germany). Samples were stored at -80C.

Table 2.3 – Primers used for Q-PCR.

	Primer sequences
ATML1	atml1-ATGF ggatatacaggcagaagaaaatcgag/ atml1-ATGR gatgatacatgatgatgatggatgc
	atml1-QF tcttctccacaagtaagagaa / atml1-QR tctcgattttcttctgcctgt
PDF2	pdf2-ATGF cacgaaaatatattgatcagtgccttg / pdf2-ATGR caaacatgtttggatggtacattgtta
	pdf2-QF tctccttgctccgagtcaata / pdf2-QR tgcatctgtacagctctcttg
ACR4	acr4-QF ataacattcccgagcgtgac / acr4-QR atctccttcagatgccgatg
EIF4	EIF4 FW gaactcatcttgtccctcaagt / EIF4 RV ttcgctcttctctttgctctcc
ChIP	qACTF agccaggtcaccatagcatt / qACTR gctctacgatttttctcactcttttat
	qACR4F cagaccgcaaggtaacaaca / qACR4R acacttctttagtcggcagaaaa
	qATML1F tgcatttcgcaataactcctt / qATML1R tggtggaggaggaacttagg
	qPDF2F caggtggagttgcaataagaaat / qPDF2R taaatgcattgcaccctctct
	qPDF1F gcaaagggatttttgaattaagttt / qPDF1R tgcaacacatcaactaaccaca

The frozen plant material was ground using a TissueLyser II (Qiagen, Germany) for two cycles of 30 s at 30 Hz. RNA extraction was performed using the Spectrum Plant

Total RNA kit (Sigma, USA) according to the appropriate protocol for seedling tissues. RNA preps were then treated using Turbo DNA-Free DNaseI (Ambion, USA) following the manufacturer's instructions. RNA quantification was performed using a NanoDrop ND-1000 UV-Vis spectrophotometer (Thermo Scientific, USA). Reverse transcription of 1 µg of the RNA extraction was performed using a SuperScript VILO cDNA synthesis kit (Invitrogen, USA).

Q-PCR reactions were carried out in a StepOne Plus Real Time PCR System (Applied Biosystems, USA). cDNA was diluted 1:30 with sterile water and the mastermix was prepared using Platinum SYBR Green Q-PCR Supermix (Roche Diagnostics, France) for a volume of 20 µl according to the appropriate protocol. Each sample contained 5 µl of diluted cDNA. The PCR reactions were performed in an optical 96-well plate (Eurogentec, Belgium). The thermal profile used for PCR reactions was as follows: 2 min at 50 °C, 2 min at 95 °C, 40 cycles of 15s at 95 °C, 20 s at 60 °C and 20 s at 72 °C. Dissociation curves were recorded using a ramp speed of 1 °C/min from 60 to 95 °C. The data was then analysed using the StepOne Software v2.2 (Applied Biosystems, USA). The *EIF4A* gene was used as reference gene. The PCR efficiency (E) was estimated from the standard curve amplification using the equation $E=10^{-1/slope}$, and the level of expression (R) of each gene of interest (GOI) was calculated using the equations $R=E^{-\Delta Ct}$,$\Delta Ct=Ct_{Goi}-C_tG_{EIF4A}$.The primers used in Q-PCR are shown in Table 2.3.

2.5 Transcriptomic analysis

The plants used for RNA sequencing were grown in long day conditions (Section 2.1). The inflorescences of *pdf2*, wild-type and *pPDF1-GFP:PDF2* transgenic plants were harvested and immediately frozen in liquid nitrogen. RNA extraction and DNAse treatments were as described previously, but RNA quality and quantity was ascertained using a Bioanalyser (Agilent, USA). One biological sample of each background was processed by GATC Biotech (Germany). On average 18 million single end reads were obtained per sample.

The data processing was based on a Tuxedo pipeline that includes the Bowtie, TopHat and Cufflinks to analyse the differential gene expression in RNA-seq

experiments (Trapnell et al. 2012). The different tools are available online in NGS TOOLBOX BETA on the Galaxy server. The first step was to upload the raw data (fastq Illumina files) on the Galaxy server and the data was groomed for quality using the *FastQC* tool. *TopHat* was then used to align the RNA-seq reads to the TAIR 10 transcriptome. The Cufflinks package includes several tools required for comparison of transcript levels between biological samples. *Cufflinks* was used to assemble the aligned reads into transcripts and *Cuffcompare* to match the assembled transcripts to the annotation database of the reference transcriptome (again, TAIR10). Finally *Cutdif* was used to determine differential expression between at least two samples and analyse whether the difference of each gene or transcript expression is statistically significant between the samples. The output is an Excel file that displays, for each gene localization/annotation, the RPKM (reads per kilobase per million) for each biological sample and the respective results of p-value and q-value statistical tests for the differential expression. Roberta Galletti analysed the data produced in this project, and in addition GATC carried out their own in house analysis. Although not identical, the results from the two analysis methods were globally comparable.

2.6 Chromatin Immunoprecipitation (ChIP)

The ChIP protocol was adapted from (Gendrel et al. 2005), and the version of the protocol I used in this project (Table 2.4) was optimized by Ralf Muller, a former member of Goodrich laboratory (IMPS, Edinburgh). The solutions required for ChIP experiments are described in Table 2.5.

The plant material used for ChIP experiments were seedlings, inflorescences or siliques of transgenic plants expressing GFP fusion proteins to either ATML1 or PDF2. Wild-type plants were included as controls. I used plants expressing four different constructs: *pATML1-GFP:ATML1*, *pFDH-GFP:ATML1*, *pPDF1-GFP:ATML1*, *pPDF1-GFP: PDF2*. Seedlings were grown for 7 days in long day conditions or continuous light, while the inflorescences and siliques were harvested from plants grown in long day conditions. The plant material was harvested and collected into a solution of 1x PBS at 4 °C. I then followed steps 1-6 of ChIP protocol (Table 2.4) to fix the material. Ground fixed plant material could be stored

at -80 °C. Alternatively, the plant material was harvested, frozen in liquid nitrogen, ground and kept at -80 °C. Then the ground material was suspended in Extraction buffer 1 and fixed according to steps 2-4 and the protocol was then followed from step 10 onwards. The sonication at step 19 was performed in a Bioruptor (Diagenode, Belgium).

Table 2.4 – Description of the ChIP protocol. From Goodrich laboratory, adapted from (Gendrel et al. 2005)), ChIP solutions are described in Table 2.5

Step	***Chromatin Immunoprecipitation protocol (from University of Edinburgh)***
0	Make autoclaved stock solutions of Extraction Buffers 1-3, Nuclei Lysis Buffer and ChIP Dilution Buffer, Low Salt Wash Buffer, High Salt Wash Buffer, LiCl Wash Buffer and TE buffer.
1	Harvest the tissue (ca 1 g) onto piece of miracloth (ca 12x12 cm) on ice. Fold the miracloth up like a bundle and use a green coated florist wire to secure it closed. Put the bundle into a 2 L beaker containing 360 ml ice-cold 1x PBS.
2	Add 10 ml (**1**) of 37% formaldehyde respectively to give 1% formaldehyde that covers the harvested plant material.
3	**Crosslink** under vacuum for 15 - 30 min. At this stage, plant material should become water-logged.
4	**Stop the crosslinking** by adding Glycine to a final concentration of 0.125M. [25 ml of 2M Glycine in 370 ml], vacuum for 5 min.
5	Remove the plant material from the vacuum and **rinse off** the formaldehyde with 1x PBS. Repeat 3X. Following the rinses, remove as much water as possible from the plant material by placing on a paper towel.
6	**Grind** the plant material in liquid Nitrogen to a fine powder. Store the plant powder in a 50 ml *greiner* tube at –80°C.
7	Add β-Mercapto-EtOH, PEFABLOC and plant PI to Extraction Buffer 1-3 and ChIP Dilution Buffer. Place solutions on ice (from step 7 to 29 the samples and solutions are kept on ice or at 4ºC, with exception of steps 16 and 24)
8	**Prepare beads** (1 tube *no antibody*; 1 - 3 tubes for *antibodies*): wash 20 µl beads (per sample/antibody) 3x in ChIP dilution buffer; add antibody (1 - 2 µg per antibody, none in the *no antibody* control) in 50 - 100 µl total volume). Incubate rotating at 4°C for 1 - 4 hours.
9	Add the tissue powder to about 20 ml of **Extraction Buffer 1** in a 50 ml *greiner* tube. Leave on ice for 5 min. Vortex occasionally
10	**Filter** the solution through miracloth into a 50 ml *greiner* tube. Repeat.
11	**Spin** the filtered solution for 20 min at 6000 g in a centrifuge at 4°C.
12	Gently decant supernatant (into a fresh tube) and resuspend the pellet in 1 ml of **Extraction Buffer 2**. Transfer the solution to a 1.5 ml Eppendorf tube.
13	Centrifuge at 12,000 g for 10 min at 4°C.
14	Remove supernatant and resuspend pellet in 300 µl of **Extraction Buffer 3**.
15	In a clean Eppendorf, add 300 µl of **Extraction Buffer 3**. Take the 300 µl solution (resuspended pellet) from step 14 and carefully layer it on top of the clean 300 µl of Extraction Buffer 3.
16	Spin for 1 hour at 16,000 g at 4°C. **Make up Nuclear Lysis Buffer and start the cooling system of sonicator.**
17	Remove the supernatant and resuspend the chromatin pellet in 300 µl of Nuclei Lysis Buffer.
18	Resuspend the pellet by pipetting up and down and vortexing (keep solution cold between vortexing). **Keep 2 µl aliquot** to run on a gel (*Input0*).
19	Once resuspended, **sonicate the chromatin solution** for 20x 30 sec ON/30 sec OFF (20 min)

	on settings marked on sonicator.
20	During the sonication: **Wash** prepared antibody coated beads 3x with ChIP dilution buffer. Let beads attach to the magnet, then discard buffer.
21	Spin chromatin solution at 12,000 g for 5 min at 4°C to pellet debris. Remove supernatant to new tube. **Remove 200 µl for ChIP and 20 µl from each "Input" DNA.** Store "Input" DNA on ice. (Keep rest to check sonication efficiency on gel. Use loading dye with RNase added Compare with aliquot from step 18, *Input0*).
22	**Dilute** the 200 µl for ChIP: Make up to 2 ml with ChIP Dilution Buffer in a 2 ml tube. The point here is to dilute the 1% SDS to 0.1% SDS with ChIP dilution buffer.
23	**ChIP:** Add 950 µl (for 2 samples) or 650 µl (for 3 samples) to antibody (or NO antibody) coated beads and incubate with rotation at 4°C for 1 hour.
24	**Prepare Elution Buffer**. For 20 ml: 2 ml of 10% SDS, 0.168 g of Sodium Bicarbonate, bring to final volume of 20 ml with distilled water.
25	Attach beads to magnet and **wash two times**, using the sequence of buffers listed below. Use 950 µl of each buffer per wash and wash at 4°C with gentle agitation. Let beads attach to magnets in between washes.
26	a) **Low Salt Wash Buffer**: Two Washes, 5min.
27	b) **High Salt Wash Buffer**: Two Washes, 5 min.
28	c) **LiCl Wash Buffer**: Two Washes, 5 min.
29	d) **TE Buffer**: One wash.
30	**Elute immune complexes** by adding 500 µl Elution Buffer to the pelleted beads. Vortex briefly to mix and incubate at 65°C for 30 min. Vortex occasionally.
31	Let beads attach to magnet and carefully **transfer the supernatant fraction** (eluate) to another tube.
32	Add 20 µl 5 M NaCl to the eluate and **reverse crosslink** at 65°C for overnight. Make up "Input" samples from step 27 to 500 µl with elution buffer, add 20 µl 5 M NaCl and place at 65°C with the other samples.
33	Add 10 µl of 0.5 M EDTA (pH 8), 20 µl Tris-HCl 1M (pH 6.5), and 2 µl of 10 mg/ml **proteinase K** to the eluate and incubate for one hour at 45°C.
34	**Recover DNA**: Add an equal volume of **phenol/chloroform** vortex and spin for 5 min. Transfer upper layer to new tube and add 1/10th salt, 2 x ethanol and 4 µl glycogen. Leave at -20°C for 1 hour to **precipitate**.
35	Spin at 4°C for 1 hour then wash pellets with 70% ethanol and leave to air dry.
36	**Resuspend the pellets in 25 µl of water.** Dilute 1 µl in 100 µl water and use 5 µl of this dilution per PCR (15 - 25 µl reaction).

Table 2.5 – Description of solutions and reagents used in ChIP experiments.

Basic Solutions and Reagents	
2 M Glycine (Sigma, USA)	1:200 plant PI (proteinase inhibitor cocktail, Sigma, USA)
37 % Formaldehyde (AnalaR Normapur)	10 mg/m Proteinase K (Invitrogen, USA)
Anti-GFP antibody (Invitrogen, USA or Abcam, UK)	10 % SDS (Euromedex, France)
Protein A or G Dynabeads (Invitrogen, USA)	10 % Sodium deoxycholate (Sigma, USA)
0.5 M EDTA (pH 8, Euromedex, France)	2 M Sucrose (Merck, Germany)
20 mg/ml Glycogen (Sigma, USA)	1 M Tris-HCl (pH 6.5, Euromedex, France)
8 M LiCl (Sigma, USA)	1 M Tris-HCl (pH 8, Euromedex, France)
1 M $MgCl_2$ (Biochemical, BDH)	10% Triton X-100 (Sigma, USA)
5 M NaCl (Euromedex, France)	14.3 M β-MEtOH (Sigma, USA)
10 % NP-40 (Sigma, USA)	$NaHCO_3$ (Sigma, USA)
1 M PEFABLOC (Roche, Germany)	

Plant ChIP solutions		
Extraction Buffer 1	*Extraction Buffer 2*	*Extraction Buffer 3*
0.4 M sucrose 10 mM Tris-HCl pH 8 10 mM MgCl2 1 mM EDTA (pH 8) 5 mM β-MEtOH 0.2 mM PEFABLOC 1:200 plant PI	0.25 M sucrose 10 mM Tris-HCl pH 8 10 mM MgCl2 1 % Triton X-100 1 mM EDTA (pH 8) 5 mM β-MEtOH 0.2 mM PEFABLOC 1:200 plant PI	1.7 M sucrose 10 mM Tris-HCl (pH 8) 2 mM MgCl2 0.15% Triton X-100 1 mM EDTA (pH 8) 5 mM β-MEtOH 0.2 mM PEFABLOC 1:200 plant PI
Nuclei Lysis Buffer	*ChIP Dilution Buffer*	*Elution Buffer*
50 mM Tris-HCl (pH 8) 10 mM EDTA (pH8) 1 % SDS 0.2 mM PEFABLOC 1:200 plant PI	1.1 % Triton X-100 1.2 mM EDTA (pH 8) 16.7 mM Tris-HCl (pH 8) 167 mM NaCl 0.1 mM PEFABLOC 1:200 plant PI	1 % SDS 0.1 M $NaHCO_3$ *TE Buffer:* 10 mM Tris-HCl (pH 8) 1 mM EDTA (pH 8)
Low Salt Wash Buffer	*High Salt Wash Buffer*	*LiCl Wash Buffer*
150 mM NaCl 0.1 % SDS 1 % Triton X-100 2 mM EDTA (pH 8) 20 mM Tris-HCl (pH 8)	500 mM NaCl 0.1 % SDS 1 % Triton X-100 2 mM EDTA (pH 8) 20 mM Tris-HCl (pH 8)	0.25 M LiCl 1 % NP-40 1 % sodium deoxycholate 1 mM EDTA (pH 8) 10 mM Tris-HCl (pH 8)

By step 36, for each biological sample, three samples had been generated corresponding to the ChIP sample, and two controls. The first control was the input obtained after the sonication and the second control was a mock Immunoprecipitation to which no antibody had been added. Additionally, I used the wild-type as a negative control of ChIP. The *ACTIN7* gene was used as a reference for semi-quantitative or Q-PCR quantification of enrichment at specific promoters. The first analysis of ChIP results included semi-quantitative PCR using two pairs of primers: actinF/actinR and L1ACR4F/L1ACR4R. Alternatively, I used Q-PCR (primers in Section 2.4) to quantify the ChIP enrichment.

2.7 Co-Immunoprecipitation

CoIP coupled with nLC-MS/MS (Nano Liquid chromatography-tandem mass spectrometry) was carried out at the University of Wageningen in the Laboratory of Biochemistry (Plant Development) in collaboration with Dolf Weijers and Cristina Llavata Peris as part of the SIREN network.

The plant material, consisted of wild-type plants and two transgenic lines expressing *pATML1-GFP:ATML1* and *pACR4-ACR4:GFP* (Gifford et al. 2003; Gifford et al. 2005). Plants were grown under long day conditions at the University of Wageningen. Siliques and inflorescences were harvested and immediately collected in 50 ml *greiner* tubes that were kept in liquid Nitrogen. The protocol used for the CoIP experiments and the set of preparation steps required for the nLC-MSMS analysis of the samples (De Rybel et al. 2013; Zwiewka et al. 2011), are described in the following Table 2.6. The results obtained from the nLC-MS/MS apparatus, were analysed using the software Max Quant (Cox and Mann 2008).

Table 2.6 – CoIP protocol and preparation steps for nLC/MSMS from the University of Wageningen. The Extraction Buffer contains 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 % NP40, protease inhibitor mix cocktail, and the Elution Buffer contains 50 mM NH4HCO3.Protocol from University of Wageningen described in (De Rybel et al. 2013; Zwiewka et al. 2011).

Immunoprecipitation of GFP tagged nuclear localized proteins with magnetic monoclonal anti-GFP beads (Miltenyi Biotech, Germany)	
1	Thoroughly grind plant material in liquid nitrogen with mortar and pestle. Transfer the ground material to a new tube and place it in liquid nitrogen.
2	Weigh the desired amount of powder on a scale in a pre-cooled mortar. Usually about 1-3 g ground plant material is used (for low expressed proteins use 3 grams of tissue and use only tissue where your protein is expressed to limit background). Immediately after, add as little Extraction Buffer as possible to solubilize the plant material and grind again very thoroughly with mortar and pestle.
3	Transfer the protein extract to a 14 ml tube (if you have more than 3 ml protein extract, divide over several 14 ml tubes) and sonicate three times 15 seconds on ice with 15 seconds pause in between. It is very important to keep the samples on ice the whole time, especially during sonication, to prevent breakdown of proteins.
4	Keep the samples 30 minutes on ice to extract the proteins. Take an aliquot (named INP: input) for Western blot.
5	Dilute the NP40 to 0.2% in the protein extract by adding Extraction Buffer without NP40 and transfer the protein extract to centrifuge tubes.
6	Centrifuge 10 minutes at 4 ºC at 10.000 rpm. Transfer the supernatant to another centrifuge tube. Centrifuge again for 10 minutes at 4 ºC at 10.000 rpm to remove all cell debris (if there is any debris remaining, the column will get stuck or run very slow). Transfer the supernatant to a new tube. Take an aliquot for Western blot (named S: supernatant).
7	Add 50-100 µl anti-GFP Microbeads (take 50 µl for 1 gram of plant material and 100 µl for 3 grams of plant material, make sure to resuspend the beads before pipetting them out of the glass pot). Mix well by swirling the tube. Rotate for two hours in the cold room.
8	Place µ Column in the magnetic field of the µMACS Separator (Miltenyi Biotech).
9	Add 200 µl Extraction Buffer with 0.1% NP40 to the column.
10	Apply cell lysate onto the column and let the lysate run through. Columns are "flow-stop" and do not run dry. Take an aliquot for Western blot of the run-through (FT: Flow Through). Pipet for each column that is used 1 ml Elution Buffer (= 50 mM NH_4HCO_3) into an Eppendorf tube and heat to 95ºC on a heat block.
11	Rinse column with 4 x 200 µl Extraction Buffer with 0.1% NP40

12	Rinse column with 2 x 500 µl 50 mM NH_4HCO_3
13	Remove the column from the magnet, immediately place a 0.5 ml Low Bind Eppendorf tube below the column and add 50 µl pre-heated hot Elution Buffer to the column. Make sure that the "end" of the column is in contact with the eluate all the time. In that way much more beads are eluted with the first 50 µl of elution buffer. Add another 50 µl pre-heated hot Elution Buffer to the column. Take an aliquot of the eluate for Western blot. Store the rest of the eluate at -20.
Preparation of Immunoprecipitation samples for Mass Spectrometry (reduction, carboxymethylation and trypsin digestion)	
14	Use the first 50 µl of the eluate of the immunoprecipitation (which contains most magnetic beads) to prepare for MS.
15	Add 1 µl 50 mM DTT (dithiotreitol) in 50 mM $NH_4HCO_3$3 pH8 (7,7 mg/l) to each IP sample and incubate 1-2 hours at 37°C.
16	Add 1 µl 100 mM iodocetamide in 50 mM NH_4HCO_3 pH8 (19 mg/l) to each IP sample and incubate at least 1 hour at room temperature in the dark.
17	Add 1 µl 200 mM cysteine in 50 mM NH_4HCO_3 pH8 (24 mg/l) to each IP sample to stop the alkylation.
18	Add 1 µl trypsin sequencing grade (0,5 µg/µl in 1 mM HCl) to each IP sample and incubate overnight at 20°C on a shaker. Do not incubate longer than approximately 16 hours because this will increase the amount of chymotrypsinic cleavages.
19	The next morning add approximately 1,5-3 µl 10% TFA to make the pH approximately pH 3. Check the pH with a pH paper (try to avoid a pH of 1 or 4).
20	Centrifuge 5 min. at maximum speed and pipet the supernatant into a new Low Bind tube. Repeat this 4-5 times to make sure that there are no beads in the samples anymore.
21	The samples are ready to inject into the MS (18 µl will be injected). Store the samples at -20.

The validation of the protein interactions between ATML1 and PDF2 transcription factors was accomplished using a CoIP approach coupled with Western blotting in collaboration with Gwyneth Ingram and Roberta Galletti (RDP, ENS).

Inflorescence material was collected from wild-type and plants expressing different combinations of FLAG and GFP fusion proteins (GFP-PDF2, GFP-ATML1, FLAG-PDF2, and FLAG-ATML1 all under *PDF1* promoter). The lines expressing one fusion protein were used as controls for the CoIP. Material was collected into microfuge tubes and immediately flash-frozen in liquid nitrogen. The protocol for the CoIP was similar to that used in Wageningen. Tissue was carefully ground in pre-cooled mortars, and the resulting powder was resuspended in extraction buffer (described in Table 2.6) and kept on ice for 30 min. Samples were then diluted with extraction buffer (without NP-40) to give a final concentration of 0.2 % of NP-40. The samples were centrifuged two times for 10 min at 14 000 g and the supernatant was transferred into new tubes. Total protein concentration was determined using a Bio-Rad Protein assay (Biorad, USA), and an aliquot of the sample was taken as an Input sample for Western Blotting analysis. 500 µg of total protein was incubated for

2 h at 4°C with anti-GFP magnetic beads (Miltenyi Biotech, Germany) with agitation. Using the µMACS Separator (Miltenyi Biotech, Germany) system, the samples were loaded into the appropriate columns and washed four times with 200 µl of extraction buffer (with 0.1 % NP-40). Bound proteins were eluted with Leaemli Buffer that had been pre-heated to 95 °C.

For the protein separation, the Co-IP samples (eluate) and 1 % of the Input were loaded on gels containing 7.5 % polyacrylamide and 0.1 % SDS and run at 150 V using Biorad (USA) MiniProtean III equipment. Proteins were transferred onto nylon membrane using an iBlot Dry System (Invitrogen, USA), the program used was as follows: 20 V for 1 min, 23 V for 5 min and 25 V for 8 min. Membranes were blocked using PBS solution containing 0.2 % Tween 20 (Sigma, USA) and 5 % no-fat milk (Regilait, France) for 1 hour. Membranes were then incubated overnight at 4°C with 1/2000 or 1/1000 dilutions of anti-FLAG M2 (Sigma, USA) or anti-GFP (Roche, Germany) monoclonal antibodies, respectively. Membranes were washed three times for 5 min in PBS containing 0.2 % Tween 20, followed by a 1 hour incubation with a 1/10 000 dilution of horseradish peroxidase-conjugated anti-mouse Igs (Promega, USA). Membranes were again washed three times before being incubated with Super Signal West Femto (Thermo Scientific, USA) reagents according to the manufacturer's instructions and exposed to Super RX film (Fujifilm, Japan).

2.8 Histochemical techniques and imaging

2.8.1 Toluidine blue staining assay

The integrity of the epidermal layer of seedlings was analysed by assessing the levels of permeability to toluidine blue dye. The technique used in these experiments has been described in (Xing et al. 2013).

After seed sterilization (Section 2.1), the seeds of different backgrounds were spread on the same plate (15 cm diameter) containing MS medium at pH 5.8, 0.3 % sucrose and 0.4 % Phytagel (Sigma, USA). The plates were kept at 4 °C for 3 days and transferred to the long day growth room for 7 days. The plates were immersed in staining solution (0.05 % toluidine blue (Merck, Germany) and 0.4 % Tween-20

(Sigma, USA)) for 2 min, and rinsed gently under the tap until the running water was clear (1 to 2 min). The roots and seed coats were carefully removed from 20 seedlings from each background, and the seedlings were collected into 1 ml of 80 % ethanol. The seedlings were incubated at room temperature for at least 2 hours until all the blue colour and chlorophyll were released into the solution. The solution was analysed using a spectrophotometer measuring the absorbance at 626 and 430 nm wavelengths (absorbance maxima for toluidine blue and chlorophyll respectively). The ratio between these values is correlated with the seedling permeability to toluidine blue. These experiments were carried out using biological triplicates.

2.8.2 GUS staining

In this project I used GUS (β-glucuronidase) staining experiments to analyse the expression of different versions of the *ACR4* promoter in seedlings, inflorescences and seeds. The GUS staining experiments included the preparation of three solutions: the phosphate buffer (0.5 M of NaH_2PO_4 and 0.5 M Na_2HPO_4 at pH 7), the X-Gluc solution (2 mM X-Gluc, 0.1 % Triton X-100, 10 mM EDTA, 100 mM of phosphate buffer, and pH 8) and Ferri/Ferro solution (1 mM $K_4Fe(CN)_6$, 1 mM $K_3Fe(CN)_6$ dissolved in phosphate buffer). The GUS staining solution is a mix of X-Gluc solution and Ferri/Ferro solution in equal proportions. After mixing the two solutions, the GUS solution was distributed in 24-well or 96-well plates (Corning, USA). The plant material was then collected, immersed in the GUS solution and incubated for 12 to 24 h at 37 °C. The stained samples were washed with 70 % ethanol several times and kept in 70 % ethanol for observation under the microscope. For the GUS staining in embryos I followed the protocol of (Vielle-Calzada et al. 2000), which involves incubation in the GUS staining solution at 37 °C for 3 days and seed clearing with 20 % lactic acid and 20 % ethanol. For GUS staining experiments in seedlings, sterilized seeds were stratified for 3 days and grown in liquid culture for 7 to 10 days in MS 5% sucrose medium in 24-well plates (Corning, USA) under long day conditions prior to staining.

2.8.3 Imaging

To optimize embryo visualization seeds were cleared using a fresh solution of 8 g chloral hydrate (Merck, Germany) dissolved in 2 ml water and 1 ml glycerol (100 %). To obtain undamaged ovules and young seeds, siliques were stuck onto double-side tape on a microscopic slide, and opened carefully along the valve margins with needles. Valves were opened out, and the replum with seeds attached, was removed with forceps and dropped into a drop of clearing solution on a new slide. Cover slips were applied gently over the solution containing the seeds and slides were kept overnight at 4 °C. The seeds were then visualized under DIC optics using a Zeiss AX10 microscope and AxioVision software (Zeiss, Germany). GUS stained inflorescences and seeds were observed using the same microscope. Dried seeds, seedlings and flowers were observed and photographed using a Leica (Germany) MZ12 stereomicroscope and an electronic camera with Pixia software.

For visualization of GFP signal in inflorescences and embryos I used a Zeiss LSM700 confocal microscope and associated software (Zeiss, Germany). Inflorescences were first vertically fixed into small plates containing 1 % agarose gel, and most flower buds were removed with forceps or small needles until the shoot apical meristem was visible. The dissected meristem was then submersed in water and transferred to a small plate containing apex culture medium (0.5x MS, 1% sucrose, 0.8 % agarose, and vitamins (1000X vitamin stock contains 5 g myo-inositol, 0.05 g nicotinic acid, 0.05 g pyridoxine hydrochloride 0.5 g thiamine hydrochloride and 0.1 g glycine dissolved in 50 ml of water)). Samples were observed using a dipping objective. For the visualization of embryos, seeds were dissected from opened siliques as described previously and immersed in water on a microscopic slide. Cover slips were applied and tapped with the back end of a pair of the forceps in such a way that the embryos popped out of the seeds without being crushed.

The SEM (scanning electron microscope) pictures were taken using Hirox SH-3000 SEM equipped with a COOLSTAGE Specimen Cooling Unit MK3 (Deben).

Chapter 3 – HD-ZIP IV family

3.1 Introduction

3.2 Results

3.2.1 Expression analysis of HD-ZIP IV family using in silico data.

3.2.2 Redundancy within the HD-ZIP IV family

3.2.2.1 Analysis of genetic relationships within the HD-ZIP IV gene family

3.2.2.2 Dosage sensitivity of ATML1/PDF2 during embryogenesis

3.2.2.3 Production of tagged HD-ZIP IV lines

3.2.3 ATML1 interacts with other HD-ZIP IV proteins

3.2.3.1 Identification of ATML1 protein interactions

3.2.3.2 Confirmation of ATML1 and PDF2 transcription factor dimerization

3.3 Discussion

3.1 Introduction

As stated in the introductory chapter, the transcription factors ATML1 and PDF2 have an important role in epidermal cell fate specification. These two proteins belong to the class IV HD-ZIP family of transcription factors which is characterized by the presence of a DNA binding domain (homeodomain), dimerization domain (leucine zipper) and START domain (Nakamura et al. 2006; Schena and Davis 1992). Some other HD-ZIP IV transcription factor encoding genes have expression patterns which overlap with those of these two epidermal regulators and therefore it has been proposed that they could act together with ATML1 and PDF2 to regulate epidermal cell fate. There are three possible and not mutually exclusive mechanisms for this interaction. They could act in independent and potentially functionally redundant pathways with ATML1/PDF2, they could be direct targets of ATML1/PDF2 or they could act together, and form heterodimers with either ATML1 or PDF2.

The *HD-ZIP IV* family is composed of 16 genes. Although the function of some genes is well described in the literature, that of others remains unknown. Briefly, *ATML1* and *PDF2* are expressed specifically in epidermis at different stages of embryogenesis, in shoot meristems and in young organ primordia (Abe et al. 2003; Takada and Jurgens 2007). Single *pdf2-1* and *atml1-1* mutants are aphenotypic but the double mutant displays severe epidermal defects that lead to lethality. Therefore it has been proposed that these genes are involved in epidermis specification (Abe et al. 2003). Moreover, *hdg3-1atml1-1* and *hdg3-1pdf2-1* double mutants are reported to present cotyledon defects suggesting that these three genes could act together in cotyledon development (Nakamura et al. 2006). Three HD-ZIP IV proteins have been implicated in trichome morphology. The *hdg11-1* single and *hdg11-1hdg12-2* double mutants show mild and severe multiple trichome branching phenotypes respectively (Nakamura et al. 2006). Additionally, GL2 is involved in trichome cell identity and hairless cell fate specification in roots. *gl2-1* mutants show an increase in root hair cell numbers and in trichome abortion (Di Cristina et al. 1996; Masucci et al. 1996; Rerie et al. 1994). From the phenotype of *anl2-1* mutants, Kubo and colleagues have suggested that ANL2 is involved in anthocyanin accumulation in

seedlings and mature tissues and in cellular organization of the root (Kubo et al. 1999).

In this chapter I analyse the HD-ZIP IV family in depth, and identify the strongest candidates for roles in epidermal cell fate specification during embryogenesis. I will analyse gene expression and gene redundancy and I will try to identify proteins that interact with ATML1.

3.2 Results

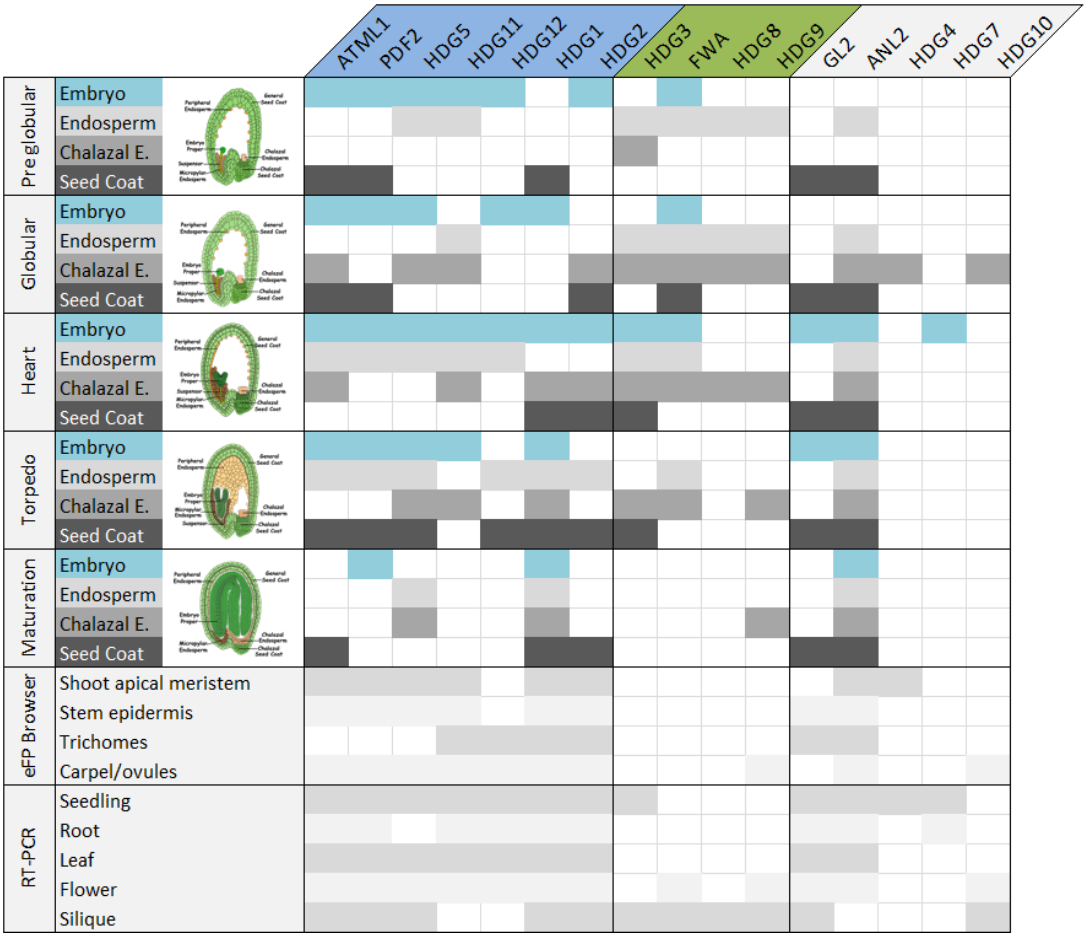
3.2.1 Expression analysis of HD-ZIP IV family using *in silico* data.

Experimental support for the role of *PDF2* and *ATML1* in the regulation of epidermal cell fate is strong based on the double mutant phenotype and on their L1 specific gene expression. Thus it appears clear that once the expression of ATML1/PDF2 is established, these two transcription factors trigger pathways involved in epidermal cell fate specification and maintenance. Nevertheless the network behind epidermal cell fate remains vague. In particular, the pathways activated by these proteins and the mode of regulation of the *ATML1* and *PDF2* genes are unknown. I concluded that an analysis of information regarding gene expression at different stages of development, and in different tissues, could be useful in permitting the recognition of potential regulatory patterns in the HD-ZIP IV family (Figure 3.1 and Table 3.1).

My approach was to combine the description of the *HD-ZIP IV* gene family provided by Nakamura and colleagues (Nakamura et al. 2006) with the micro-array derived expression data assembled in different bioinformatics resources (Table 3.1 and HDZIPIV.xlsx file). The tool eFP Browser (Electronic Fluorescent Pictograph Browser) assembles microarray and other high-throughput expression data from different sources in an intuitive mode (Winter et al. 2007). I used the information relating to *Arabidopsis thaliana* from three different features: “Developmental Map”, “Tissue Specific” and “Seed” (Winter et al. 2007). I also used “Genevestigator” to characterize the expression of *HD-ZIP IV* family members. In particular I used the “Development” and “Anatomy” profiles (Hruz et al. 2008). Finally, a seed specific dataset (which forms the basis of the “Seed” data in the eFP Browser) was assembled by the project “Gene Networks in Seed Development”, and covers the gene

expression profiles of specific seed compartments (isolated using Laser Capture Microdissection) at five stages of seed development (Le et al. 2010). More information is available on the site associated with this project than is presented on the eFP Browser. From comparative analysis of the expression profiles shown in these resources, I clustered genes that have expression profiles overlapping with those of *ATML1/PDF2* into the *Embryo* group, and genes which show expression specifically, or majoritarily, in endosperm during seed development into the *Endosperm* group (Table 3.1 and Figure 3.1).

Table 3.1 –Summary of expression analysis of HD-ZIP IV family. Seed profiles obtained from the *Gene Networks in Seed Development* project describing five stages of embryogenesis (Le et al. 2010), in this table *"Endosperm"* refers to both peripheral endosperm and the ESR. A selection of data from the eFP Browser is also incorporated into this table, regarding expression in shoot meristem, stem, trichomes, stigma and ovules (Marks et al. 2009; Suh et al. 2005; Swanson et al. 2005; Winter et al. 2007; Yadav et al. 2009). RT-PCR results are from the Nakamura study on HD-ZIP IV class genes (Nakamura et al. 2006).



In addition to *ATML1* and *PDF2*, the *Embryo* group contains five *HD-ZIP IV* genes (*HDG1*, *2*, *5*, *11*, and *12*). According to the protein sequence alignment of the HD-ZIP IV families, (Figure 3.1) three proteins encoded by genes of this group are closely related to ATML1 and PDF2 (HDG2, 11 and 12) whereas HDG1 and HDG5 are more distantly related. The seed expression data shows that the expression patterns of these genes is dynamic. In general, their expression starts at the earliest stages of embryogenesis (except for *HDG1* which initiates at the globular stage), consistent with potential roles in epidermis specification. According to Seed Networks the expression of *HDG2* starts at the heart stage, but according to promoter-GUS fusion results (Nakamura et al. 2006) this gene is expressed at the dermatogen stage. Interestingly at the heart stage 12 out of the 16 *HD-ZIP IV* genes are actively expressed in the embryo (Table 3.1), which suggests that the *HD-ZIP IV* family may have an important role at this stage of embryogenesis.

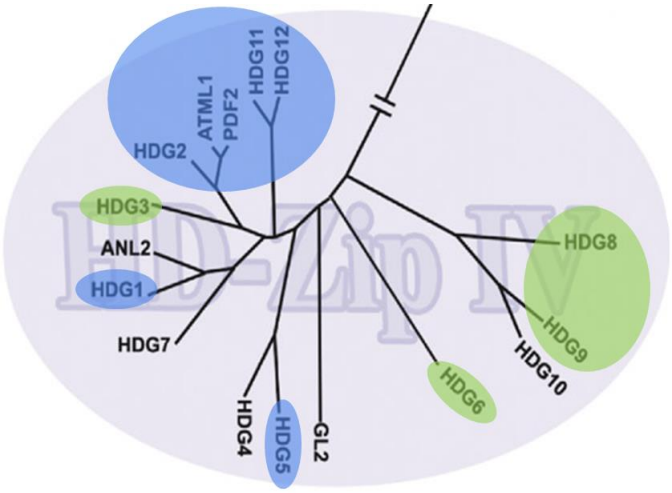


Figure 3.1 – HD-ZIP IV family representation. (Green – *Endosperm* group, blue – *Embryo* group; Phylogenetic tree adapted from (Ariel et al. 2007))

Another particularity of the *Embryo* group is that all genes are expressed in the shoot apical meristem (SAM) and carpels, in the same zones as *ATML1/PDF2*. Results from the transcriptional analysis of three regions of the SAM, made using fluorescence-activated cell sorting with three lines expressing GFP under the promoters of *CLV3* (central zone, including outer cell layers), *WUS*, and *FIL* (peripheral zone, including outer cell layers) are available on the eFP browser (Winter et al. 2007; Yadav et al. 2009). As expected, *ATML1* and *PDF2* are present

in *CLV3* and *FIL* datasets, but are absent in WUS domain, which is consistent with L1 expression. The same tendency is observed for *HDG1*, *HDG2*, *HDG5*, and *HDG11*. With exception of *HDG1*, the genes clustered in the *Embryo* group are expressed in SAM epidermis according to promoter-GUS experiments (Nakamura et al. 2006). All genes were expressed in the carpel but only *HDG11* and *HDG12* were expressed in stigma tissue (Nakamura et al. 2006; Swanson et al. 2005; Winter et al. 2007). According to Swanson dataset *HDG1* is expressed in carpels, which is not consistent with promoter-GUS experiments in flowers that show expression of *HDG1* promoter exclusively in the epidermal layer of stamens (Nakamura et al. 2006; Swanson et al. 2005).

Four HD-ZIP IV transcription factors (HDG3, 6, 8 and 9) were assembled into the *Endosperm* group (Table 3.1 and Figure 3.1). From comparison of protein sequences in the HD-ZIP IV family these proteins are not as clearly linked in terms of sequence similarity as the *Embryo* group (Figure 3.1). Nevertheless, the two proteins HDG8 and HDG9 are quite similar which is reflected by their expression pattern and may point to either heterodimerization and/or functional redundancy (Table 3.1). These four genes are mainly expressed in the endosperm and chalazal pole of the developing seed (Table 3.1). *Endosperm* genes are not expressed at the maturation stage of seed development with the exception of *HDG9*, which remains expressed in chalazal tissues. Only at the heart stage is *HDG3* expression detected in the embryo and seed coat. The expression profile of *HDG6/FWA* (*FLOWERING WAGENINGEN*) also starts at early stages of development when it is expressed in embryo, ESR and peripheral endosperm. At the globular stage *FWA* is expressed in the whole seed. According to the literature, *FWA* expression is specific to the endosperm (Kinoshita et al. 2004) which is not consistent with the microarray data (Le et al. 2010).

Five HD-ZIP IV genes are not included in either the *Embryo* or *Endosperm* group. The genes, *ANL2*, *GL2*, *HDG4*, *7* and *10* did not fit the criteria imposed for each group but this does not exclude an involvement of these genes in epidermis specification. Although the *ANL2* expression domain largely overlaps the expression of genes in *Embryo* group, its expression in the embryo is only detected at later

stages of embryogenesis, long after epidermal specification is thought to occur. *HDG7* expression was not detected at any stage of seed development (Le et al. 2010), although *HDG7* promoter activity was detected in the shoot apical region of the embryo at late stages of embryogenesis (Nakamura et al. 2006). *HDG4* and *HDG10* are expressed uniquely in the chalazal endosperm at the globular stage and *GL2* is expressed mainly in the seed coat, consistent with its role in this tissue (Qing and Aoyama 2012; Rerie et al. 1994). Briefly, these five genes have rather diverse expression patterns both during seed development and in other tissues.

In conclusion, my analysis of *in silico* and published results regarding the expression of the *HD-ZIP IV* gene family suggests the presence of two gene clusters within which there may be a higher probability of functional redundancy. The *Embryo* group of genes contains the best candidates for a similar role to *ATML1* and *PDF2*. I also suggest that a second group of four genes (the *Endosperm* group) exists. These genes are expressed in ESR at early stages of development. Finally the expression pattern of *HD-ZIP IV* genes in seed development seems to be spatially and temporally very dynamic.

3.2.2 Redundancy within the HD-ZIP IV family

The HD-ZIP family of transcription factors is characterized by the DNA binding domain (HD) and the dimerization domain. Gene duplication events and competition for the same binding sites have been described for several other classes of HD-ZIP transcription factors and may contribute to the functional redundancy found between these proteins (Ariel et al. 2007; Ciarbelli et al. 2008; Henriksson et al. 2005). However, redundancy is often incomplete due to diversification of expression patterns. In this section I analyse different combinations of *HD-ZIP IV* mutants to understand if genes of this family act together with *ATML1* and *PDF2* in specifying epidermal cell fate. Furthermore I describe my findings regarding the relationship of *ATML1* and *PDF2* during embryogenesis via double mutant analysis.

3.2.2.1 Analysis of genetic relationships within the *HD-ZIP IV* gene family

The analysis of *HD-ZIP IV* genetic interactions started with the acquisition of homozygous mutant lines for each gene (Table 3.2 and Figure 3.2). The transcription

factor GL2 was excluded from this analysis due to its late expression in the embryo and lack of expression in the shoot apical meristem. Due to genotyping problems we abandoned *fwa-t1* mutants, leaving fourteen *HD-ZIP IV* genes to study. In most cases we chose to study only one mutant allele per gene, concentrating on alleles which clearly disrupted the open reading frame. This decision was made to avoid an unmanageable multiplication of lines. The analysis the *atml1-3/pdf2-3* mutant is included in a separate section together with the single mutant analysis of *pdf2-3* and *atml1-3*.

Table 3.2 – Description of the mutant alleles for the *HD-ZIP IV* gene family

Genes	Allele	Mutation	Reference
ATML1	*atml1-1*	T-DNA insertion	(Abe et al. 2003; Kamata et al. 2013; Nakamura et al. 2006)
	atml1-2	Mutant screen (Ler)	(Roeder et al. 2012)
	atml1-3	SALK_033404	(Kamata et al. 2013; Roeder et al. 2012) and this study
	atml1-4	SALK_128172	(Roeder et al. 2012)
PDF2	*pdf2-1*	T-DNA insertion	(Abe et al. 2003; Kamata et al. 2013; Nakamura et al. 2006)
	pdf2-2	SAIL_503_E09	(Kamata et al. 2013)
	pdf2-3	SALK_109425	This study
ANL2	*anl2-1*	l element (Ler)	(Kubo and Hayashi 2011; Kubo et al. 1999)
	anl2-2	SALK_000196	(Kubo and Hayashi 2011; Nakamura et al. 2006), and this study
HDG1	*hdg1-1*	SALK_006757	(Kamata et al. 2013; Nakamura et al. 2006)
	hdg1-2	SALK_062171	(Kamata et al. 2013; Nakamura et al. 2006)
	hdg1-3	SALK_147739	This study
	hdg1-4	SAIL_326_F03	This study
HDG2	*hdg2-1*	Mutant screen	(Weigel et al. 2000)
	hdg2-2	SALK_127828	(Kamata et al. 2013; Nakamura et al. 2006; Peterson et al. 2013) and this study
	hdg2-3	SALK_138646	(Kamata et al. 2013; Nakamura et al. 2006; Peterson et al. 2013)
	hdg2-4	SALK_120064	(Peterson et al. 2013)
HDG3	*hdg3-1*	SALK_033462	(Kamata et al. 2013; Nakamura et al. 2006) and this study
HDG4	*hdg4-1*	WiscDsLox328F6	(Nakamura et al. 2006)
	hdg4-2	GK-168F12	This study
HDG5	*hdg5-1*	SALK_032993	(Kamata et al. 2013; Nakamura et al. 2006)
	hdg5-2	SAIL_238_A05	(Kamata et al. 2013)
	hdg5-3	SALK_020840	This study
FWA	*fwa-t1*	SALK_064256	(Nakamura et al. 2006) and this study
HDG7	*hdg7-1*	T-DNA insertion	(Nakamura et al. 2006)
	hdg7-2	SAIL_871_E10	This study
HDG8	*hdg8-1*	SAIL_586_E12	(Nakamura et al. 2006) and this study
	hdg8-2	SALK_088578	This study
HDG9	*hdg9-1*	SAIL_534_D06	(Nakamura et al. 2006)
	hdg9-2	SALK_079210	(Nakamura et al. 2006)
	hdg9-3	SALK_011611	This study
HDG10	*hdg10-1*	SALK_116071	(Nakamura et al. 2006)
	hdg10-2	SALK_151998	This study
HDG11	*hdg11-1*	SAIL_865_G09	(Kamata et al. 2013; Nakamura et al. 2006)
	hdg11-2	SALK_044434	(Nakamura et al. 2006; Yu et al. 2008) and this study
HDG12	*hdg12-1*	SAIL_870_C06	(Nakamura et al. 2006)
	hdg12-2	SALK_127261	(Kamata et al. 2013; Nakamura et al. 2006) and this study
	hdg12	SALK_061323	(Yu et al. 2008)

Most of the T-DNA insertions I worked with (Figure 3.2) are localized in exons with the exception of *hdg10-2* and *hdg8-2* (in introns). In particular, the T-DNA insertions in *hdg1-3*, *hdg2-2*, *hdg7-2*, *hdg8-1*, and *pdf2-3* are located in the START domain encoding region and the *hdg3-1*, *atml1-3*, and *anl2-2* insertions are located in the homeodomain encoding region.

Several of the *HD-ZIP IV* mutant lines chosen for my analysis were already described in the literature and match the previous description. The novel T-DNA insertion lines which I used were all indistinguishable from wild-type except for the fact that some *atml1-3* and *pdf2-3* seedlings showed weak cotyledon defects and some embryo phenotypes (Section 3.2.2.2).

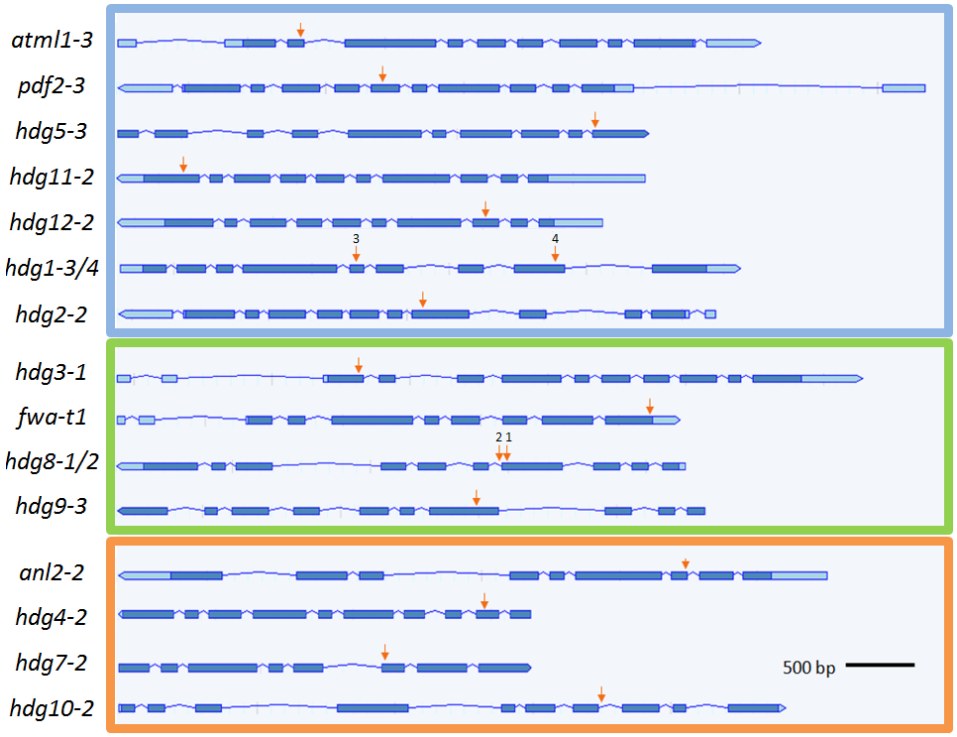


Figure 3.2 – Gene structure of the *HD-ZIP IV* family and localization of T-DNA insertions. UTRs are indicated by a light blue box, exons and introns are represented by blue boxes and blue lines respectively. The direction of the gene transcription is indicated by arrows (adapted from TAIR gene representation). The localization of each insertion mutation used in this study is illustrated with an orange arrow.

I originally aimed to obtaining crosses between mutants in all the 14 genes retained (Figure 3.2). However, this turned out to be unfeasible due to time and space constraints, and I subsequently decided to concentrate only on double mutant

combinations involving *atml1-3* or *pdf2-3* and additionally between genes from the *Embryo* group. Following the results from Immunoprecipitation using *pATML-GFP:ATML1* transgenic lines (Section 3.4), I included combinations between *anl2-2* and genes from the *Embryo* group. In total I aimed for 37 double mutant combinations (Figure 3.3). From this double mutant analysis some potential interactions were detected which need to be explored further and confirmed.

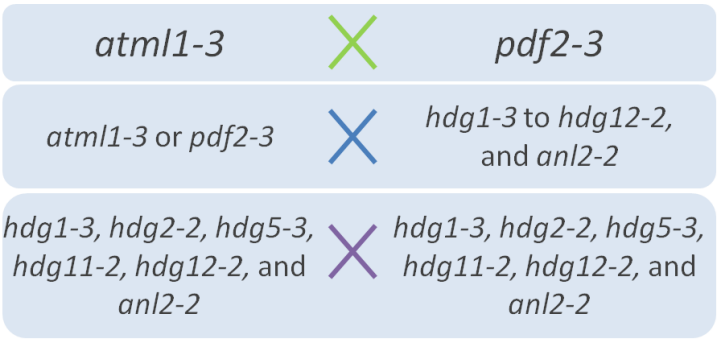


Figure 3.3 – HD-ZIP IV double mutant plan.

The genotyping of most HD-ZIP IV double mutants resulted in the identification of at least one double homozygous plant. However, several crosses are not complete. In particular, only the F2 generation of *anl2-2 x hdg11-2* and *anl2-2 x hdg12-2* crosses is available. The progeny of F3/F4 plants homozygous for one mutation and heterozygous for another (*aaBb*) was obtained for two crosses: *hdg12-2 x pdf2-3* and *hdg2-2 x hdg12-2*. The genes *PDF2* (At4g04890) and *ANL2* (At4g00730) are localized in the same chromosome relatively close to each other which makes genetic recombination challenging. The F2 generation obtained from the *anl2-2 x pdf2-2* cross consisted only of double heterozygous, *pdf2* or *anl2* plants. I therefore concentrated on generating an *anl2-2atml1-3* double mutant.

The double mutant *hdg11-2hdg12-2* was also difficult to obtain. However, an *hdg11-2hdg12-2* double mutant plant was recently obtained in the progeny of a plant homozygous for *hdg12-2* and heterozygous for *hdg11-2*. As previously reported (Nakamura et al. 2006), this plant showed multiple trichome branching, but also had notable defects in ovule development and embryo defects, leading to dramatically reduced fertility. This genotype is currently being studied further in the laboratory. The analysis of a significant number of plants homozygous for one mutant and

heterozygous for the other will help to understand if *hdg12-2pdf2-3* double mutants also have real development problems.

Some double mutants were tested for defects in seedling epidermal integrity using the toluidine blue staining carried out on seedlings according to (Xing et al. 2013). This allowed me to ascertain that double mutant combinations of *hdg5-3* with either *atml1-3* or *pdf2-3* mutants result in a weak seedling phenotype (mainly characterized by an increased proportion of monocotyledonous seedlings) and cuticle permeability defects (Figure 3.4). Notably the permeability to toluidine blue of double mutant *pdf2-3hdg5-3* seedlings is significantly higher (t-test inferior to 2 %) than *atml1-3hdg5-3* and respective single mutants. The seedlings of *hdg5-3*, *atml1-3* and *pdf2-3* single mutants are also significantly more permeable to toluidine blue when compared to the wild-type seedlings although the difference between single mutant backgrounds is not significant (Figure 3.4). The genetic interaction between *HDG5* and *ATML1/PDF2* supports the theory that other *HD-ZIP IV* genes could act functionally redundantly with *ATML1* and *PDF2*, however again this interaction needs to be studied more closely.

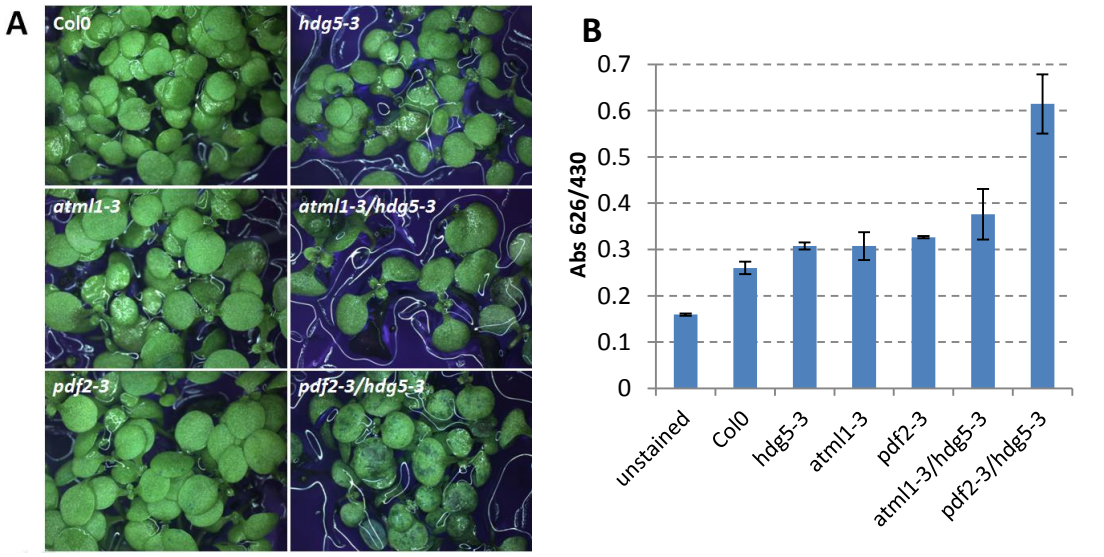


Figure 3.4 – Defects in seedling epidermal integrity of *HD-ZIP IV* mutant combinations. After toluidine blue staining (A), the quantification of seedling permeability of *atml1-3hdg5-3*, *pdf2-3hdg5-3* double mutants and respective single mutant to toluidine blue is represented by the average of ratios between toluidine blue (Abs 626) and chlorophyll (Abs 430) concentration (B).

Interestingly, although my results for previously described double mutants were globally in agreement with those in the literature, I was totally unable to see the severe phenotypes previously described for *pdf2-1/hdg3-1* and *atml1-1/hdg3-1* double mutant combinations in *pdf2-3hdg3-1* and *atml1-3hdg3-1* double mutants. As discussed in the introduction, there is some doubt as to the nature of the *pdf2-1* allele which could explain the discrepancy regarding the *pdf2-1/hdg3-1* result, but this does not explain the difference in our results with *atml1-3hdg3-1* mutants compared to *atml1-1/hdg3-1* mutants.

As double mutant analysis of *HD-ZIP IV*s is on-going, here I have only the phenotypes identified so far, which does not exclude the possibility that other double mutants have interesting phenotypes. The genetic interactions which I have demonstrated need to be explored and validated further. The generation of higher order mutant combinations may be required to unravel new genetic relationships between HD-ZIP IV class proteins.

3.2.2.2 Dosage sensitivity of ATML1/PDF2 during embryogenesis

According to the literature, the transcription factors ATML1 and PDF2 act redundantly to specify and maintain epidermal cell fate identity (Abe et al. 2003; Tanaka et al. 2007). Single mutants in these genes were described as being aphenotypic and the atml1-1pdf2-1 phenotype was characterized by the lack of an epidermal cell layer in seedlings, leading to lethality (Abe et al. 2003). My results suggest that the atml1-3pdf2-3 phenotype is more severe than that previously reported for *atml1-1pdf2-1*, and that there is a haploinsufficiency for *ATML1* in *pdf2-3* mutant background and for *PDF2* in *atml1-3* mutant background which implies a strong dosage dependency for these genes.

My analysis of *pdf2-3* and *atml1-3* mutants suggests that both display a weak cotyledon phenotype (Figure 3.5). They are also more permeable to toluidine blue compared to wild-type (t-test inferior to 2 %) which may indicate some defects in epidermis integrity. The offspring of a plant homozygous for *pdf2-3* and heterozygous for *atml1-3* showed higher permeability to toluidine blue than the respective single mutants (Figure 3.5C).

The double mutant *atml1-3pdf2-3* is lethal. I never identified a double homozygous plant in the F2 generation of *atml1-3* x *pdf2-3* or *pdf2-3 x atml1-3* crosses, nor in the progeny of F3 plants homozygous for one of these two mutants and heterozygous for the other. Therefore I had to work with the lines homozygous for *pdf2-3* and heterozygous for *atml1-3* (*ppAa*) or lines homozygous for *atml1-3* and heterozygous for *pdf2-3* (*aaPp*). From here on I will refer to these genotypes as *ppAa* and *aaPp* respectively. Both *ppAa* and *aaPp* plants produce around 25% hollow, shrivelled seeds which failed to germinate. Additionally, a number of abnormal shaped seeds were found in both backgrounds but more frequent in *ppAa* progeny. When seeds from these backgrounds were germinated, a significant number of abnormal seedlings were identified (Figure 3.5). Most of these seedlings survived and were genotyped *ppAa* or *aaPp*, the same genotype as the parent. These data are indicative of haploinsufficiency for *ATML1* or *PDF2* in backgrounds already mutated for the other gene.

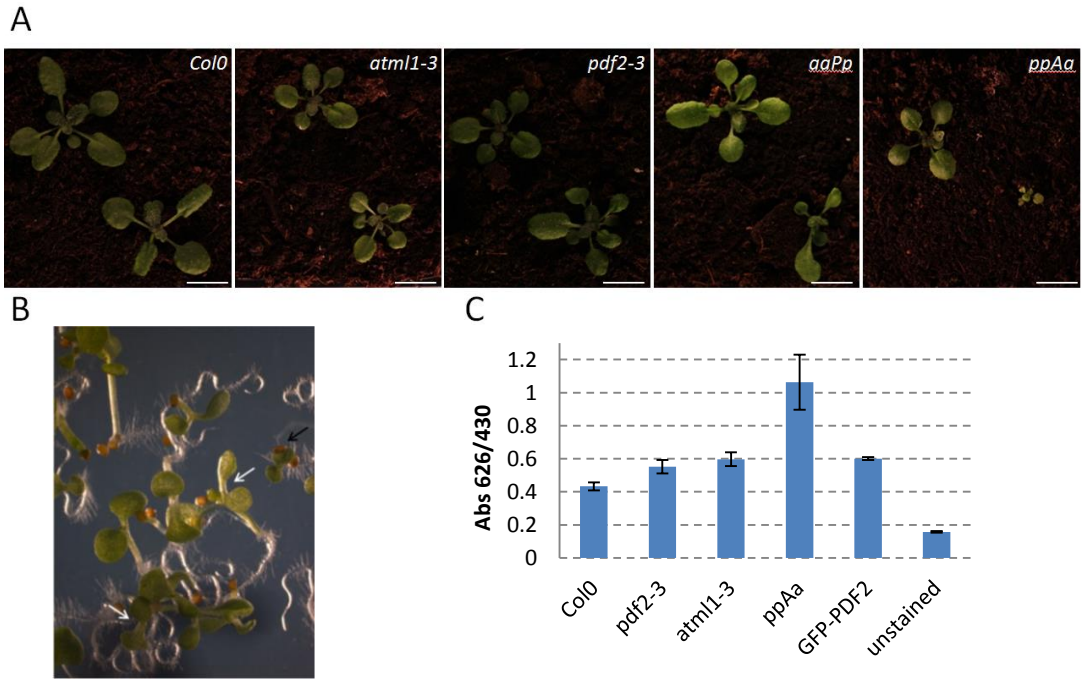


Figure 3.5 – Characterization of *pdf2-3* and *atml1-3* mutant alleles. A) Comparison of progeny of F3 plant *atml1-3pdf2-3* homozyous/heterozygous (*aaPp* and *ppAa*) with respective single mutants and wild-type. B) Seedling phenotypes of progeny from a *ppAa* parent (white arrows show three cotyledon seedlings). C) The average of ratios between toluidine blue (Abs 626) and chlorophyll (Abs 430) concentration is displayed for *ppAa* double mutant, the respective single mutants and pPDF1-GFP:PDF2 transgenic line. The scale bar corresponds to 1 cm.

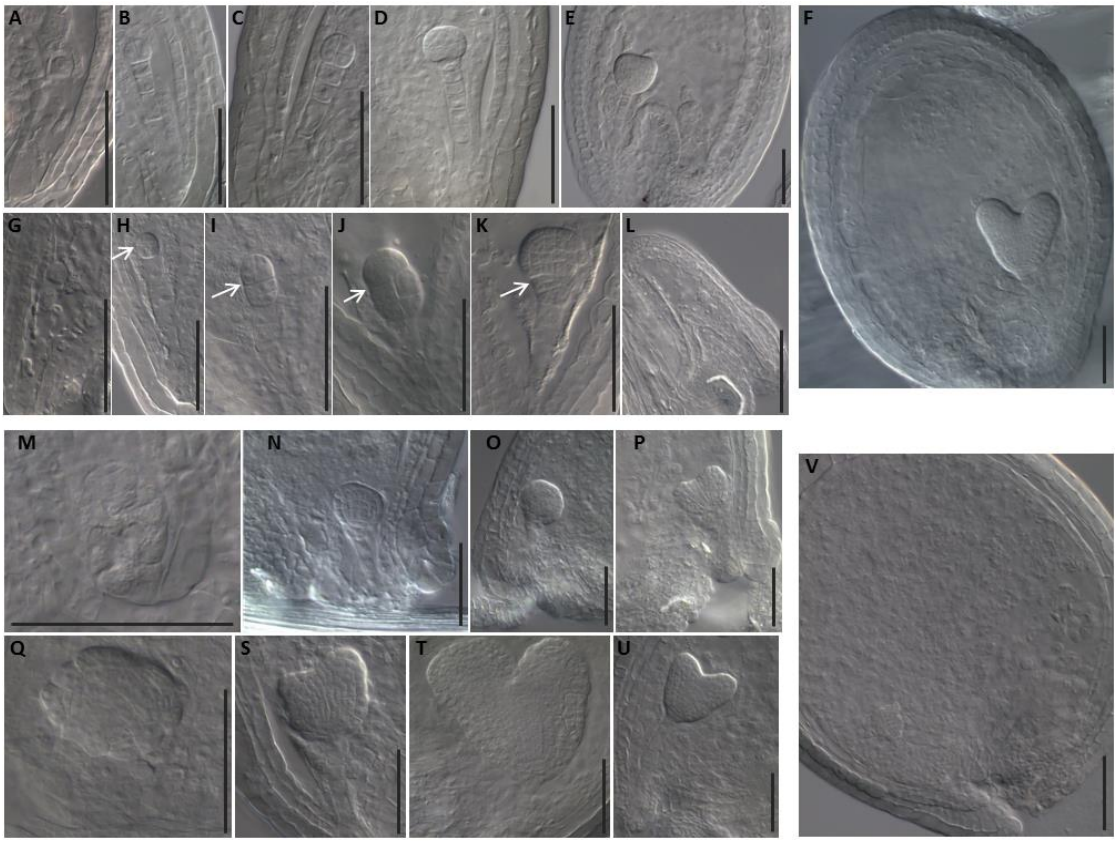


Figure 3.6 – Embryogenesis in *HD-ZIP IV* mutant backgrounds. The different stages of embryogenesis in wild-type correspond to A-F. In general, the phenotype of *atml1-3* and *pdf2-3* single mutants is similar to the wild-type (A-F), however a small number of *atml1-3* or *pdf2-3* embryos (G-L) have some defects in cell divisions (white arrows) or seed shape. M-V show the different phenotypes of embryos found in the siliques of plants homozygous for *pdf2-3* and heterozygous for *atml1-3*. Several embryos appear to be pressed into the micropylar region (M-P). Embryos likely corresponding to the double homozygous genotype arrest at the globular stage (Q, V) whereas the embryos that are likely homozygous for *pdf2-3*and heterozygous for *atml1-3* show division problems in the epidermis (S, T) (corresponding to about 50% of seeds analysed). The embryos likely to be homozygous for *pdf2-3* and wild-type for *ATML1* (U) resemble the wild-type (about 25% of seeds analysed). The scale bar corresponds to 50 µm.

The study of embryogenesis in the mutant backgrounds involved clearing of seeds at different developmental stages followed by observation using DIC optics (Figure 3.6). At early stages of embryogenesis, a small proportion (4-6%) of *atml1-3* and *pdf2-3* embryos showed defects in their basal pole, suspensor, or epidermis layer. Such defects were not observed in wild-type seeds. Some embryos displayed abnormal division planes during the first division of the embryo proper (perpendicular rather than horizontal to the apical/basal axis). Additionally, some seeds exhibit a flattened shape, embryos that appeared “squeezed”, and endosperms

that developed more slowly than in other seeds (Figure 3.6). Interestingly at the heart stage, despite these apparent defects in early development most seeds and embryos appeared to be viable and similar to the wild-type in phenotype suggesting that ATML1 and PDF2 function is more important, or at least less redundant, at early stages of the development when the epidermis is being specified.

To find when and how *atml1-3pdf2-3* embryos arrest their development I examined then siliques of *ppAa* plants at different stages of embryogenesis (Figure 3.6). At early stages of development some embryos exhibited orientation problems. The strongest phenotypes looked as if embryo growth had occurred in the direction of the micropylar pole (Figure 3.6 M) due to suspensor bending or problems in suspensor development (different cell size or cell number). At the heart stage of embryogenesis, around ¼ of the embryos had arrested at the globular stage with severe epidermal defects. The disorganization of the outermost layer in these cases was correlated with a lack of differentiation of root and apical poles and consequently a lack of cotyledon primordium development. We predict that these embryos are *atml1-3pdf2-3* double homozygotes and correspond to the empty/shrivelled seeds of *ppAp* progeny. Interestingly, about half of the embryos presented an intermediate phenotype characterized by irregularities in the epidermis due to abnormal division planes that were first observed at the globular stage (Figure 3.6S, and T). Typically, the development of these embryos was slower compared to wild-type looking embryos and at later stages they produced shorter cotyledons. This group of embryos are probably of the *ppAa* genotype and when germinated should correspond to most of the abnormal seedlings we observed. The remaining 25 % of embryos have phenotypes which are comparable to those of wild-type embryos and should correspond to the *pdf2-3* genotype (Figure 3.6U).

In summary, I have described haploinsufficiency of *ATML1* or *PDF2* in the *pdf2-3* and *atml1-3* mutant backgrounds respectively, and a much earlier embryo lethal phenotype for *atml1-3pdf2-3* double homozygous mutants than that previously reported for *atml1-1pdf2-1* double mutants.

3.2.2.3 Production of tagged HD-ZIP IV lines

To understand *ATML1* and *PDF2* dosage dependency and more importantly *ATML1*/*PDF2* function and regulation I wanted to produce lines overexpressing *ATML1* and *PDF2*. My first strategy was to produce transgenic lines expressing ATML1 and PDF2 tagged with either a FLAG tag or GFP at the N-terminus and driven by the 35S promoter. Most primary transformants of these four constructs did not survive after seedling transfer to soil, apparently presenting problems concerning meristem function. The seedling phenotype included long strap-like cotyledons and malformed first leaves. Some plants survived but transgene expression was not detectable in these lines, suggesting that the transgenes were silenced. These results indicate that the ectopic expression of *PDF2* or *ATML1* generates critical development problems. In order to solve this problem the promoter of *PDF1* was selected to drive the ATML1 or PDF2 fusion proteins (Figures 3.7-3.10).

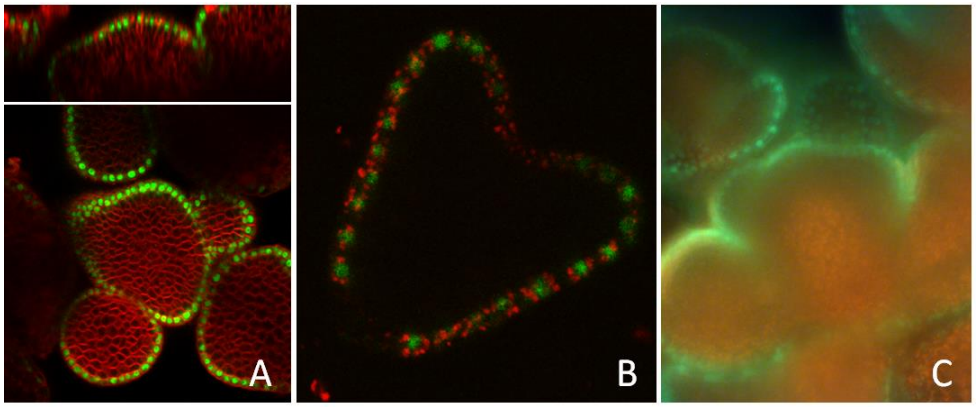


Figure 3.7 – Expression pattern of the *PDF1* promoter. The GFP-PDF2 driven by *pPDF1* is detected in the epidermis of the shoot apical meristem (A) and embryo (B) (confocal microscopy). (C) The expression of GFP-ATML1 driven by the *PDF1* promoter in a complemented *atml1-3pdf2-3* double mutant (fluorescence microscopy).

The expression of *PDF1* was previously reported to be L1 specific (Abe et al. 1999; 2001; Tanaka et al. 2007) and thus overlap with the expression domains of *PDF2* and *ATML1*. Additionally, because the removal of an L1 box from the *PDF1* promoter eliminated its expression, and *PDF1* was not expressed in the *atml1-1pdf2-1* double mutant, *PDF1* was proposed to be positively regulated by ATML1/PDF2 (Abe et al. 2003; Abe et al. 2001). Therefore we hypothesised that driving *PDF2* and *ATML1* expression from the *PDF1* promoter might cause strong but epidermis specific

overexpression due to the creation of a positive feedback loop. Constructs for the expression of GFP-ATML1, GFP-PDF2, FLAG-ATML1, or FLAG-PDF2 under the *PDF1* promoter were generated and for each construct at least 30 independent transgenic plants were produced. Several (at least 5) independent homozygous lines for each construct were subsequently generated and analysed (Figure 3.7-3.10). When lines expressing GFP-PDF2 or GFP-ATML1 were studied by fluorescence or confocal microscopy, GFP nuclear signal was restricted to the L1 layer of embryos and shoot apical meristems in seedlings and inflorescences, consistent with previous reports regarding the epidermal specificity of the *PDF1* promoter (Figure 3.7).

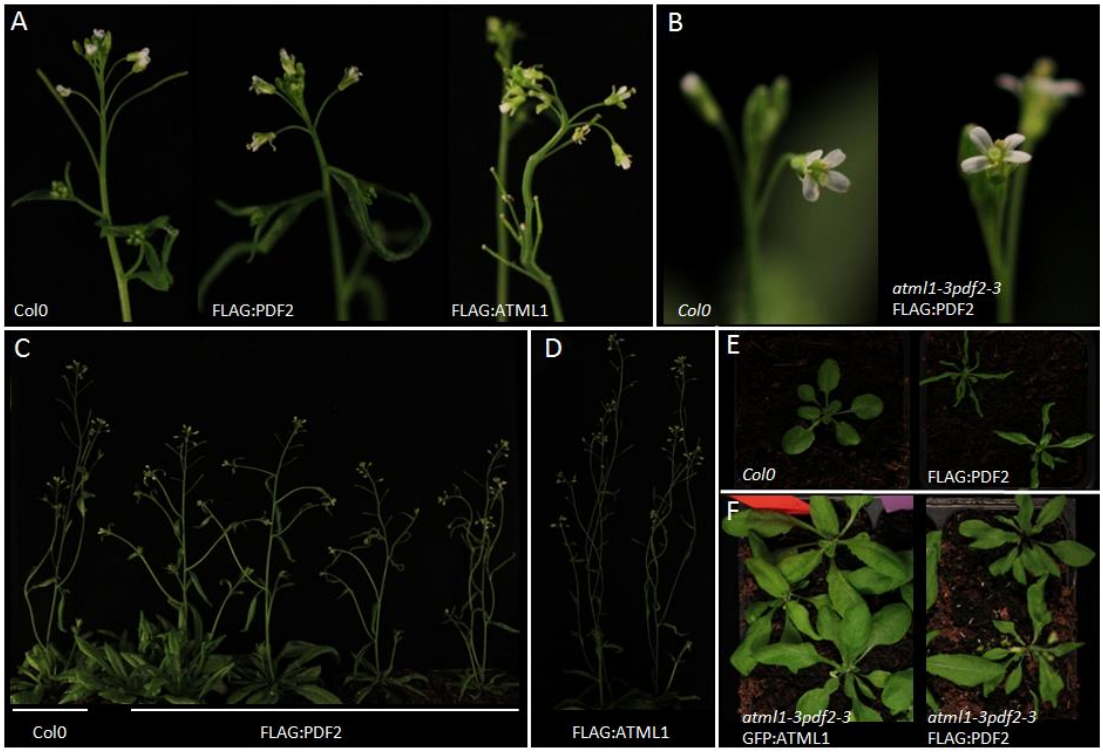


Figure 3.8 – Phenotypes of plants expressing fusions of ATML1 or PDF2. The flowers of plants expressing FLAG fusions to PDF2 or ATML1 display long strap-like floral organs compared to the wild-type flowers, additionally these lines can show stem fusions (A). The lines overexpressing PDF2 (C) and ATML1 (D) show leaf epinasty and spiralling of leafs. Four independent lines overexpressing FLAG-PDF2 are shown in (C) compared to the wild-type. *pPDF1-FLAG:PDF2* and *pPDF1-GFP:ATML1* complement the *atml1-3-pdf2-3* lethal phenotype, displaying the flower (B) and rosettes phenotype (F) comparable to the wild-type and FLAG:PDF2 expressing lines(E).

Wild-type plants expressing each fusion combination were crossed with plants homozygous for *atml1-3* and heterozygous for *pdf2-3* to test the functionality of the construct. In the F2 generation of these crosses I identified double mutant *atml1-3*

pdf2-3 plants containing either the *pPDF1:FLAG-PDF2* or *pPDF1: GFP-ATML1* transgene. Therefore these two fusion proteins are able to rescue the lethal phenotype of the double mutant. Rescued double mutants showed relatively normal rosette phenotypes (Figure 3.8), although they had abnormal floral organs (see below) and showed a reduced seed set. Although there was a high degree of seed lethality in these rescued plants, several hundred viable seed were harvested and, upon germination, their genotypes were re-tested and confirmed. The FLAG-ATML1 and GFP-PDF2 lines used in the crosses did not give rise to rescued F2 plants. We propose either that not enough F2 plants were screened, or that the transgene expression levels in the lines chosen for the crosses were insufficient to rescue epidermal specification during early embryogenesis.

From this work my first conclusion is that, consistent with the phenotypes of single mutants, *ATML1* and *PDF2* are functionally redundant in the specification of epidermis, since either of these genes is enough to restore the lethal phenotype of the respective double mutant. Secondly the *PDF1* promoter is active early enough in embryogenesis to recapitulate the expression of *ATML1/PDF2*. In the absence of ATML1 or PDF2, other HD-ZIP IV transcription factors may be responsible for the activation of the *PDF1* promoter. Alternatively, *PDF1* may not be a direct target of ATML1 and PDF2. In fact, in seedlings, recent data from other group members suggest that the *PDF1* promoter may in fact be slightly repressed by ATML1 and PDF2, possibly explaining why overexpression under this promoter remains very moderate (Gwyneth Ingram, personal communication).

The majority of *GFP:ATML1* or *GFP:PDF2* transgenic lines showed a wild-type phenotype, with exception of some lines that displayed narrow leaves and occasional stem fusions. The total *ATML1* RNA expression levels in inflorescences of wild-type plants expressing GFP or FLAG-tagged ATML1 (tested in the T3 generation) were generally not significantly different from the expression levels of the native gene in wild-type plants. However, *TAG-ATML1* primary transformants with strong phenotypes were sterile (due to a pollen defect), which may explain why no lines showing overexpression of tagged ATML1 were obtained in the T2 generation. Lines

expressing GFP/FLAG fusions with PDF2 generally overexpress total *PDF2* 2-3 fold (Figure 3.9A).

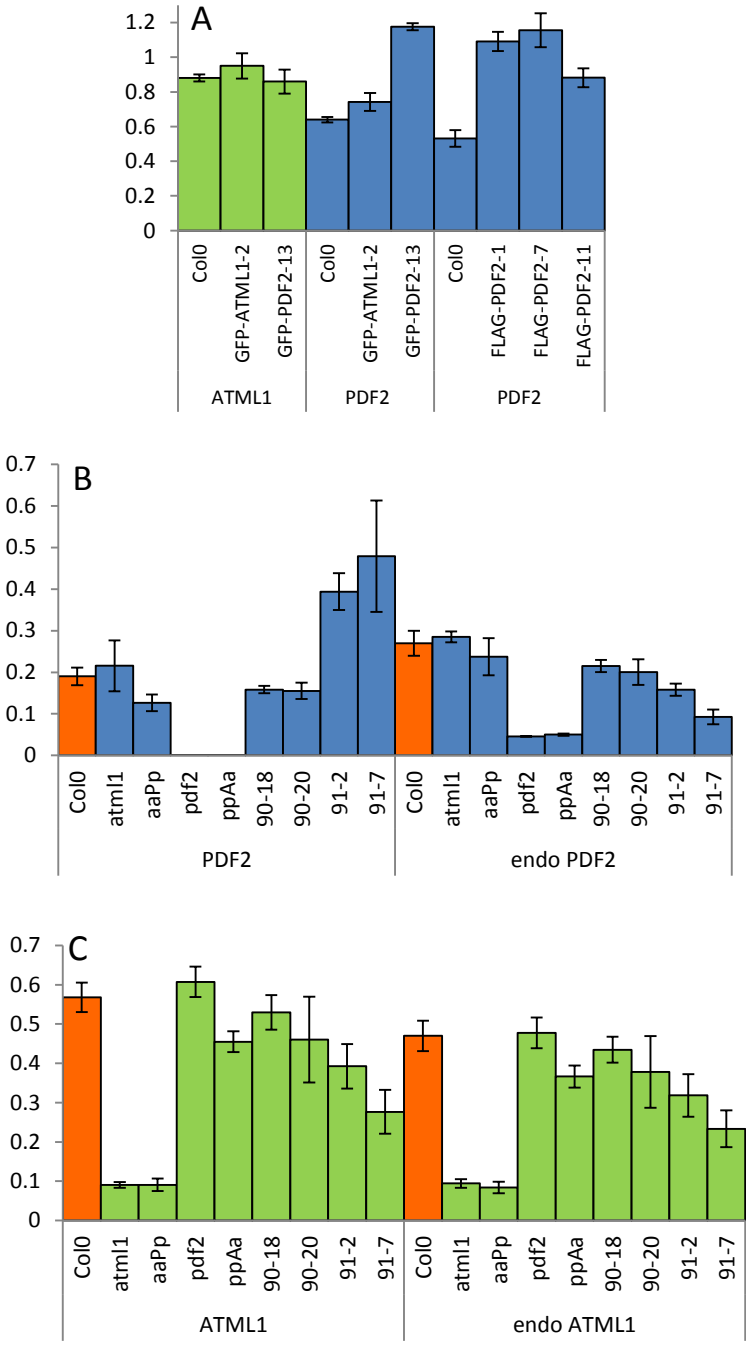


Figure 3.9 – Transcription levels of *ATML1* and *PDF2* in mutant and transgenic lines. A) Total *PDF2* and *ATML1* expression in inflorescences of wild-type plants, and wild-type plants expressing FLAG/GFP fusions with ATML1 or PDF2 under the *PDF1* promoter. Comparison of endogenous *PDF2* (B) and *ATML1* (C) seedling expression with total *PDF2/ATML1* expression for Col0, *atml1-3* and *pdf2-3* single mutants, *atml1-3pdf2-3*mutant/heterozygote combinations (genotypes *aaPp* and *ppAa*) and lines expressing FLAG constructs under the *PDF1* promoter (90 and 91 represent FLAG:ATML1 and FLAG:PDF2 fusion proteins respectively). Q-PCR is normalized to *EIF2* expression.

To understand possible regulatory relationships between *ATML1* and *PDF2* we analysed the transcript levels of *PDF2* and *ATML1* in seedlings of different backgrounds (Figure 3.9B, and C), namely the transgenic lines expressing FLAG fusions of ATML1/PDF2, *atml1-3* and *pdf2-3* single mutants, and *atml1-3pdf2-3* mutant/heterozygote combinations (*aaPp* and *ppAa*). I designed primers that allowed quantification of the transcript level of only the native *ATML1/PDF2* gene (endo *ATML1/PDF2*), and primers that allowed quantification of total expression of each gene (endogenous + transgenic). For *atml1-3* single mutants, a dramatic decrease in *ATML1* endogenous transcript was observed, as was the case for *PDF2* transcript in *pdf2-3* mutants. However the presence of T-DNAs in the coding region in each case makes it impossible to ascertain if this effect is due to decreased transcription or decreased mRNA stability. Interestingly, the levels of *ATML1* and *PDF2* are slightly higher in the *pdf2-3* and *atml1-3* mutant backgrounds respectively. However this difference is not significant enough to suggest a strong compensation of *ATML1* or *PDF2* in the absence of *PDF2* or *ATML1* respectively. The overexpression of *PDF2* in FLAG-PDF2 transgenic lines resulted in the repression of both the endogenous *ATML1* and *PDF2* genes. Globally, these results suggest that *ATML1* and *PDF2* are autoregulated by a negative feedback mechanism in seedling tissues and in inflorescence tips.

I observed that overexpression of both *PDF2* and *ATML1* gave specific developmental phenotypes. These effects appeared to be slightly stronger for FLAG fusion proteins and, indeed, as mentioned above, gave rise to sterility in the case of FLAG:ATML1 expressing lines. The production of narrow leaves was observed in both GFP and FLAG lines but this phenotype was stronger in plants expressing FLAG fusions (Figures 3.8 and 3.10). In extreme cases we observed both epinastic curling and spiral growth of leaves (Figure 3.10). The leaves of these transgenic lines have trichomes with only one or two branches, and the lobes of epidermal pavement cells are less prominent than in wild-type plants (Figure 3.10). The trichome phenotype is the opposite of that previously reported for double *hdg11-1hdg12-2* mutants, which have multiple branched trichomes (Nakamura et al. 2006).

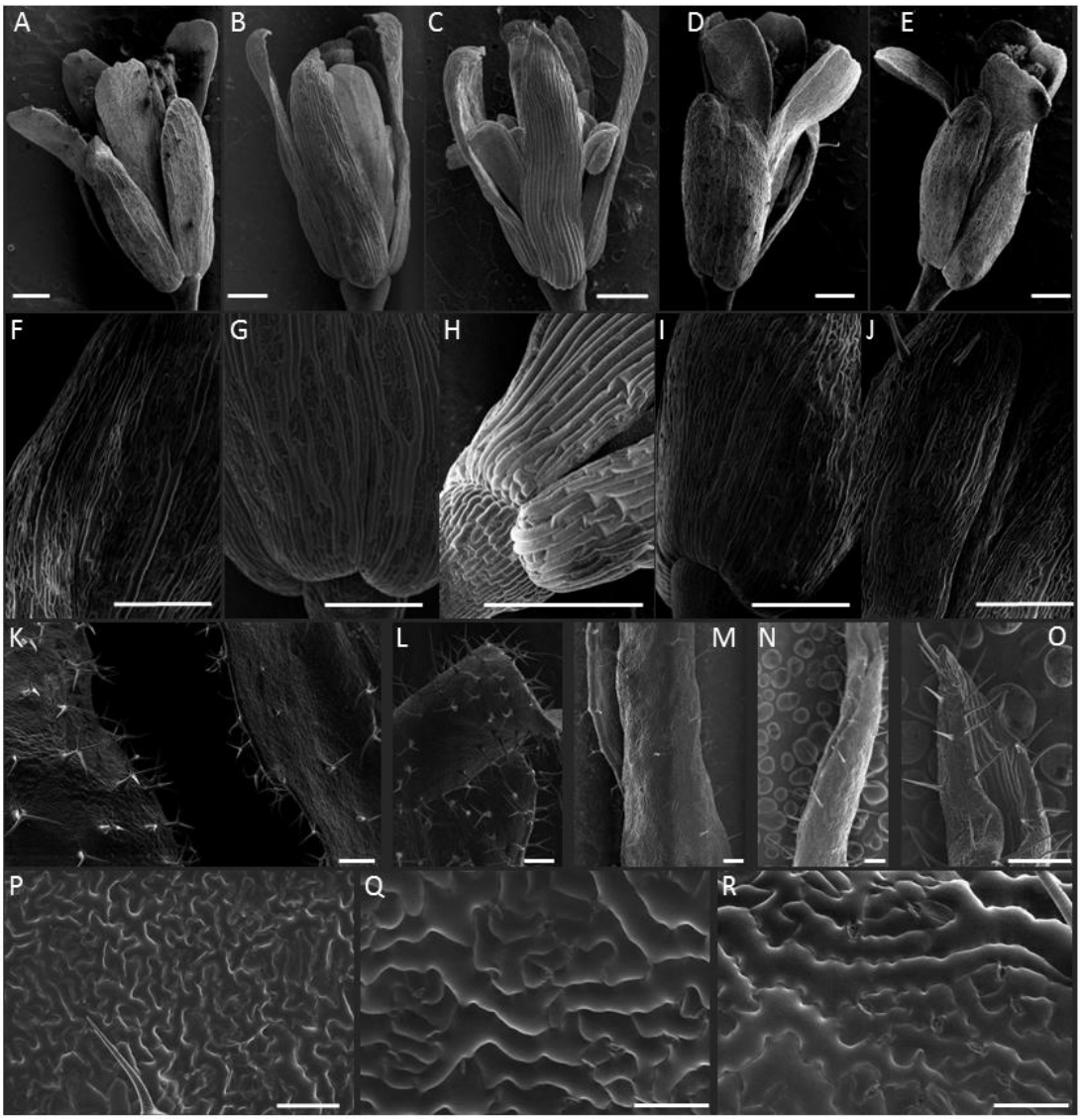


Figure 3.10 – The effects of dosage dependence of ATML1/PDF2 in leaf and flower development. The wild-type control is represented in A, F, K, and P; *atml1-3* and *pdf2-3* are represented in D, I and E, J respectively; the phenotypes of *ATML1* overexpression are shown in B, G, L, M and Q; and the phenotypes of *PDF2* overexpression are shown in C, H, N, O and R. The dose of either ATML1 or PDF2 disturbs the frequency of giant cells in sepals, the higher the amount of ATML1/PDF2 the higher is the frequency of giant cells (A-J). The overexpression of PDF2 or ATML1 corresponds to a decrease in trichome branching (L-O), epinasty and curling of leaves (L-O) and decrease in lobe number of pavement cells in leaf epidermis (Q,R). The scale bar of A-0 and P-Q correspond to 200 µm and 50 µm respectively.

Regarding flower development, the lines expressing *FLAG:ATML1* and *FLAG:PDF2* showed long strap-like floral organs (Figures 3.8 and 3.10). As previously mentioned, several lines expressing *FLAG:ATML1* showed substantial fertility problems. The frequency of giant cells in sepals of *FLAG:ATML1* and FLAG:PDF2 overexpressing transgenic lines was much higher than in the wild-type. The sepals of

some lines overexpressing FLAG:PDF2 were almost completely covered in giant cells (Figure 3.10). This is the opposite phenotype to that observed in *pdf2-3* and *atml1-3* mutants, where giant cell frequency is significantly lower than in wild-type, as reported by (Roeder et al. 2012) in other mutant alleles for these genes. The opposite phenotypes observed in loss of function and overexpression lines again support my conclusion that the fusion proteins used in this study are functional.

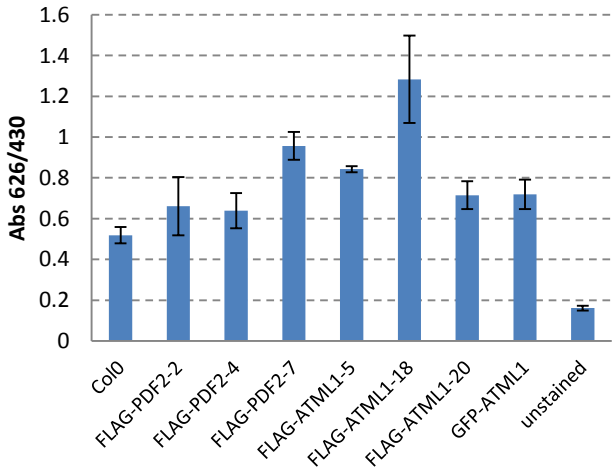


Figure 3.11 - Defects in seedling epidermal integrity of ATML1/PDF2 fusion protein overexpressing plants. The quantification of seedling permeability is represented by the average of ratios between toluidine blue (Abs 626) and chlorophyll (Abs 430) concentration.

Because seedling permeability to toluidine blue of *atml1-3*, *pdf2-3* and *ppAa* lines is increased, I also tested this characteristic in lines over expressing ATML1/PDF2 fusion proteins. Interestingly, like the *atml1-3* or *pdf2-3* (Figure 3.5), most seedlings expressing FLAG/GFP fusion proteins to ATML1 or PDF2 were significantly more permeable to toluidine blue than the wild-type. The correlation between overexpression and seedling permeability was not strong, and indeed the transgenic line expressing FLAG-ATML1 which was most permeable to toluidine blue, does not show a global overexpression of *ATML1* (Figures 3.9, 3.11). Nevertheless the transgenic lines with the strongest adult phenotypes showed higher seedling permeability. These results were surprising and suggest that any disturbance in the expression levels of *ATML1/PDF2* affect seedling epidermal development. It is important to bear in mind that toluidine blue permeability can be caused by a variety of different defects, and is a very general indicator of epidermal abnormality.

3.2.3 ATML1 interacts with other HD-ZIP IV proteins

Independent experiments *in vitro* have shown not only that HD-ZIP transcription factors are able to form dimers, but that the dimerization domain is required for the binding of HD-ZIP transcription factors to pseudopalindromic DNA binding sites via the homeodomain (Palena et al. 1999; Sessa et al. 1993; Tron et al. 2001; Tron et al. 2004). However, the dimerization of HD-ZIP transcription factors has not been described *in vivo*. Heterodimerization in the HD-ZIP IV family would have the potential to generate multiple diverse regulatory combinations that could compete for L1 box-like binding sites. In this section I investigate potential protein interactions with ATML1/PDF2. My approach was to identify HD-ZIP IV interactions *in planta,* using protein fusions of ATML1 and PDF2. Not only did I validate the homo and heterodimerization of ATML1 and PDF2 but I also have strong evidence that other HD-ZIP IV proteins form complexes with ATML1.

3.2.3.1 Identification of ATML1 protein interactions

The initial characterization of protein interactions with the ATML1 transcription factor was carried out using *pATML1-GFP:ATML1* transgenic plants generated before my arrival in the laboratory. The experiments included three biological replicates containing about 3 g of siliques and one sample of inflorescence tissue (up to 2 g) from wild-type and GFP-ATML1 expressing plants. I used Immunoprecipitation (IP) coupled with nLC-MS/MS (Nano Liquid chromatography-tandem mass spectrometry) for peptide identification to identify ATML1 complexes.

Briefly this technique involves three steps (Figure 3.12). The first is the immunoprecipitation of the GFP-tagged protein (GFP:ATML1) using columns containing magnetic monoclonal anti-GFP-linked beads that are used to select and isolate ATML1 complexes. Subsequently, samples are subjected to a sequence of treatments that includes washing, protein reduction, carboxylmethylation and trypsin digestion. Finally, the tryptic peptides are loaded onto a nano liquid chromatographic column and identified using mass spectrometry. The results from MS/MS were analysed using the program MaxQuant (Cox and Mann 2008). This program uses a database of MS/MS spectrograms of peptides that result from trypsin digestion. I

used the protein database provided by the MS facilities in Wageningen University. This database includes the UniProt protein sequence data for *Arabidopsis* and a selection of possible contaminants (such as human keratin). The predicted MS/MS spectra are then compared with the experimental spectra. The quantity of each peptide in the sample is measured by the peak area. The chosen output of this program is the LFQ (Label-Free Quantification) intensity of each protein, which combines the information for each peptide. Afterwards, I used the program Perseus for data treatment, for instance to mark the detection threshold of the nLC-MS/MS, to convert LFQ intensity to log10 in order to facilitate the analysis, and to estimate the t-test values between tagged lines and wild-type. The comparison of LFQ intensities between the *pATML1-GFP:ATML*1 line and wild-type corresponds to the protein abundance ratio and reveals the best candidates for ATML1 interactors *in vivo* (results are shown in HDZIPIV.xlsx file). These experiments were carried out during several visits to the Laboratory of Biochemistry and MS/MS facilities in Wageningen University.

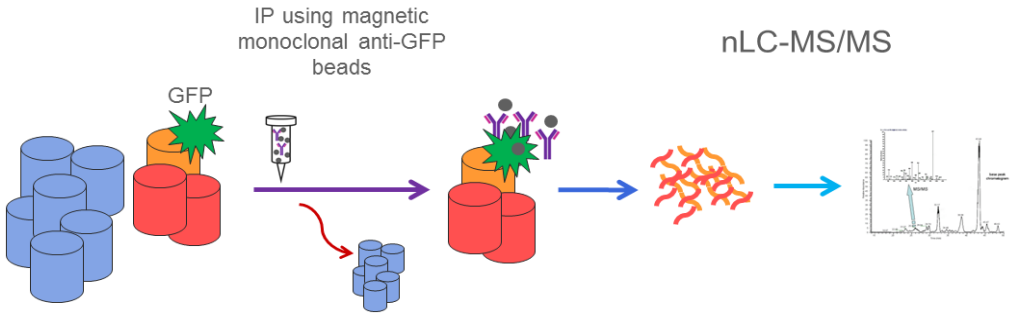


Figure 3.12 – Method of identification of protein complexes. The first step consists of the Immunoprecipitation of fusion proteins and their protein partners (orange and red respectively) using columns holding anti-GFP antibodies (purple) attached to magnetic beads (grey). The eluate is subjected to a series of chemical treatments including trypsin digestion. Finally, the identification of the peptides is achieved using nLC-MS/MS facilities.

In my analysis, the HD-ZIP IV transcription factors in the *Embryo* group (HDG1, 2, 5, 11, 12 and PDF2) were the best candidates to form protein complexes with ATML1 (Table 3.3). Additionally, ANL2 was also detected in the IPs, both in siliques and inflorescences. Since I used three biological replicates for the IP experiments in siliques I was able to test whether the results between GFP-ATML1 expressing plants and wild-type control were statistically significant (Table 3.3). In

order to account for protein sequence similarities I analysed the raw data for each peptide. I found that several peptides were shared between HD-ZIP IV members. It is worth mentioning that PDF2 and ATML1 have 16 peptides in common and 9 of these peptides are shared with HDG12. The study of unique peptides (peptides that are unique to a given protein) showed that the IP was very efficient since most of these peptides were undetectable in the wild-type samples (in both siliques and inflorescences).

Table 3.3 –HD-ZIP IV protein interactions with ATML1. For each HD-ZIP IV transcription factor, the average of LFQ intensity in the *pATML1-GFP:ATML1* transgenic lines and wild-type control (WT) is represented in the logarithmic scale. The ratio of protein abundance between GFP-ATML1 expressing lines and wild-type (Fold) is shown for two biological assays using siliques or inflorescences. Additionally, the ratio between unique peptides and total identified peptides number (UPep/IdPep) is indicated for each transcription factor.

Protein	UPep/ IdPep	Siliques				Inflorescences		
		GFP-Atml1	WT	Fold	t-test	GFP-Atml1	WT	Fold
ATML1	27/43	8.01	2.76	180014	0.0035	8.69	3.96	53613
PDF2	18/34	6.77	3.14	4224	0.0069	7.75	2.29	292266
ANL2	8/11	5.39	2.26	1344	0.0023	6.59	-	97081
HDG1	18/21	6.24	4.57	46.4	0.0024	7.15	2.55	40488
HDG2	9/18	5.87	-	18660	0.0047	6.85	3.62	1695
HDG5	23/24	6.46	2.81	4461	0.0019	7.82	1.8	904186
HDG11	13/14	5.61	3.56	111	0.0228	7.27	3.55	5313
HDG12	9/10	5.44	-	6962	0.0009	6.95	-	221513

To investigate the interaction between HD-ZIP IV proteins I analysed the data for unique peptides and ruled out some outliers in the silique dataset (measurements that identified a given unique peptide in only 1 out of the 3 biological samples). In pATML1-GFP:ATML1 lines, the concentration of ATML1 was found to be higher than that of GFP in both siliques and inflorescences datasets (5 and 6 times respectively), which is consistent with homodimerization between native and GFP fusion versions of ATML1. In siliques, the PDF2 concentration was calculated to be about 3 times lower than that of ATML1 and the concentration of the other six HD-ZIP IVs identified were between 12 (HDG2) and 25 (HDG11) times lower than that of ATML1. In the inflorescence dataset the general trend was similar. ATML1 concentration was about two times the concentration of PDF2 and 6 (HDG11) to 40 (ANL2) times higher than that of the other HD-ZIP IV proteins detected (see

HDZIPIV.xlsx file). Taking together, these observations may suggest that ATML1 dimers are present in higher-order complexes of HD-ZIP IV proteins in plants.

In addition to HD-ZIP IV transcription factors, one unique peptide of the MADS box transcription factor SEP1 (SEPALLATA1) was pulled down with GFP-ATML1 (Figure 3.13). This peptide is absent in two out of three wild-type controls in siliques. Additionally, this peptide was also identified in the inflorescences sample suggesting that the interaction between ATML1 and SEP1 is not seed specific.

A second experiment was carried out using the same experimental conditions using one biological sample of wild-type and pATML1-GFP:ATML1 siliques. In this independent experiment almost the same HD-ZIP IV transcription factors were identified, with the exception of HDG5. SEP1 was not detected.

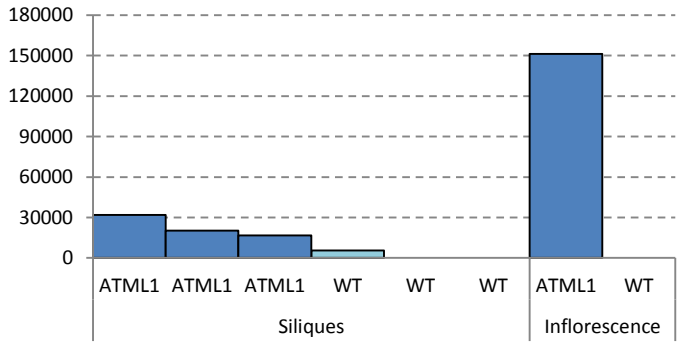


Figure 3.13 –SEP1 may interact with ATML1. This picture represents the ITQ Intensity of the SEP1 peptide for three biological samples of siliques and one biological sample of inflorescences of plants expressing GFP-ATML1, and wild-type.

Our results provide strong indications that HD-ZIP IV transcription factors form protein complexes. ATML1 has more affinity with PDF2 than with other HD-ZIP IV proteins. My results indicate that the proportion of each protein in ATML1 complexes changes according to the tissue studied although the same seven HD-ZIP IV proteins were detected in the inflorescence and silique data. My analysis indicates that HDG2 and HDG12 are more likely to interact with ATML1 in seeds/embryos, and that HDG11, 12 and 1 are more likely to interact with ATML1 in the apical meristem/inflorescence.

3.2.3.2 Confirmation of ATML1 and PDF2 transcription factor dimerization

From the IP experiments I strongly suspected that ATML1 and PDF2 interact physically. The validation of this interaction was carried out using the newly generated lines expressing GFP or FLAG fusion proteins of PDF2 or ATML1 driven by the *PDF1* promoter. These lines are described in previous sections. I crossed different lines expressing either ATML1 or PDF2 tagged proteins and obtained plants expressing FLAG-ATML1/GFP-ATML1, FLAG-PDF2/GFP-PDF2, FLAG-ATML1/GFP-PDF2 and GFP-ATML1/FLAG-GFP. With these lines we aimed to detect homo and heterodimers of ATML1/PDF2 using a combination of Co-Immunoprecipitation and Western blotting (Figure 3.12). These experiments were carried out in collaboration with Roberta Galletti and Gwyneth Ingram in the RDP laboratory, at the ENS of Lyon.

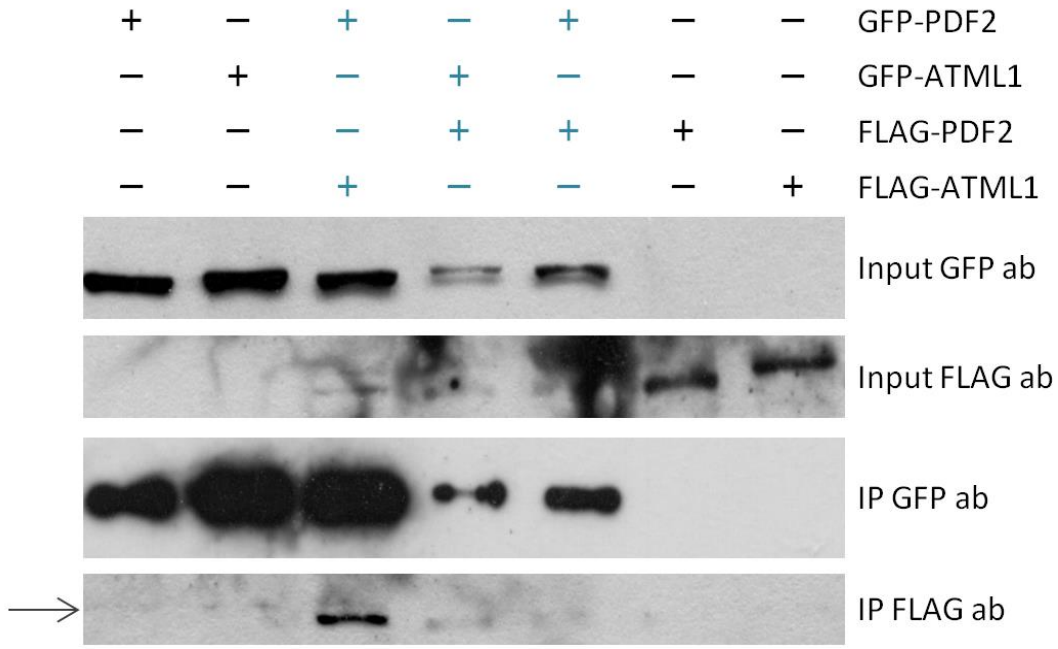


Figure 3.12 – ATML1 and PDF2 protein interactions in inflorescences. Western blot analysis was carried out on samples collected before (Input) and after the immunoprecipitation step. The immunoprecipitation was carried out using anti-GFP antibodies (IP). The detection of FLAG and GFP fusion proteins was carried out using the respective monoclonal antibodies. Three lines expressing different combinations of FLAG fusion and GFP fusion ATML1/PDF2 proteins were used to test homo and heterodimerization (blue), the controls consisted of lines expressing only one protein fusion (black). Immunoprecipitated FLAG-tagged proteins are clearly visible in lane 3, confirming heterodimerization of ATML1 and PDF2, with fainter bands visible in lanes 4 (heterodimerization of ATML1 and PDF2) and lane 5 (homodimerization of PDF2). A background band is present in the IP FLAG blot, situated slightly higher than the fusion proteins (arrow)

These experiments validated the identification of ATML1/PDF2 complexes obtained using the IP nLC-MS/MS approach, although the expression of tagged proteins tended to be considerably lower in double transgenic plants than in single transformed lines, making protein-protein interactions difficult to visualize. In particular, we identified GFP-ATML1/FLAG-PDF2, GFP-PDF2/FLAG-ATML1, and GFP-PDF2/FLAG-PDF2 confirming PDF2 homodimerization and ATML1/PDF2 heterodimerization, and showed the presence of full length fusion proteins for all ATML1/PDF2 FLAG and GFP constructs. Neither FLAG nor GFP tags seem to influence the dimerization of these two transcription factors indicating that they do not interfere with the function of ATML1/PDF2.

3.3 Discussion

The presence of a dimerization domain in HD-ZIP IV proteins and the similarity of their predicted binding sites, establish the conditions for multiple potential regulatory combinations of these transcription factors. Expression pattern diversity, associated with protein-protein interaction combinations could therefore provide a powerful tool for refining epidermis specification both during and after embryogenesis. Our results suggest that several HD-ZIP IV-encoding genes act together with *ATML1/PDF2* to specify epidermal identity during embryogenesis. The analysis of double and single knockout lines in association with the phenotypes of lines overexpressing ATML1 or PDF2 suggests that the dosage dependency of ATML1/PDF2 has a significant impact throughout plant development.

HD-ZIP IV transcription factors potentially involved in epidermal cell fate

The expression profiles of HD ZIP IV-encoding genes during embryogenesis vary both spatially and temporally. Overlapping expression of these genes is predicted in several different seed compartments and at different stages of embryo development. There seems to be a correlation between specific stages of embryogenesis and the number of genes activated. At the heart stage, the majority of HD-ZIP IV genes are expressed in the embryo (12/16). At this stage, epidermal identity is firmly established and presumably starts to be refined in different embryonic zones due to the elaboration of the shoot and root meristems and cotyledon primordia. The role of

HD-ZIP IV proteins in cell patterning could implicate the activation of regulators, or other direct targets involved in cell differentiation or cotyledon development. My results indicate that the dosage of ATML1, PDF2 and HDG5 significantly affects the number and orientation of cotyledon primordia initiated at this stage in embryogenesis. Further genetic analysis may uncover further players in this process, and in the process of meristem initiation.

Protein sequence similarities and previous genetic analysis have suggested the probability of high levels of functional redundancy between HD-ZIP IV transcription factors, supporting the possibility that other family members besides ATML1 and PDF2 are involved in epidermal cell fate specification. By grouping similar expression patterns, I defined two groups of HD-ZIP IV encoding genes: the *Endosperm* (*HDG3, 6, 8,* and *9*) and *Embryo* groups (*HDG1*, *2*, *5*, *11*, *10*, *ATML1* and *PDF2*).

Most genes in the *Endosperm* group were expressed specifically in the endosperm at early stages of embryogenesis. Intriguingly, however, HDG3 has been suggested to be involved in cotyledon development in cooperation with ATML1 and PDF2 (Nakamura et al. 2006). One possible mechanism could involve HDG3, together with other HD-ZIPs in the *Endosperm* group, activating the upstream signal used to initiate the expression of *ATML1*and *PDF2*. This signal could be generated in the endosperm and perceived by the outermost layer of the embryo. However, although my genetic analysis globally agreed with previously published results, I was totally unable to observed the phenotypes previously reported for seedlings in which the functions of *HDG3* and either *ATML1* or *PDF2* are lost (Nakamura et al. 2006). Since the *hdg3-1* allele that I used was identical to that used in the studies by Nakamura and colleagues, and my *atml1-3* and *pdf2-3* alleles appear to be as, or more, severe than those used in this previous study, it is difficult to explain this result. As discussed in the introduction, the *pdf2-1* allele may behave as a neomorphic or antimorphic allele, explaining the discrepancy in this background. However, this does not explain the difference in results obtained for *hdg3atml1* double mutant alleles.

In addition, I was unable to observe the *hdg5-1pdf2-1* flower phenotypes described by (Kamata et al. 2013). Again, this discrepancy may well be attributable to the *pdf2-1* allele (Kamata et al. 2013). The fact that HD-ZIP IV proteins have recently been described as being involved in the regulation of floral organ identity is intriguing in the light of the interaction that I observed between SEP1 and ATML1. SEP1, SEP2, and SEP3 transcription factors are functionally redundant and are required for the development of flower organs, in the triple mutant all flower organs develop as sepals (Pelaz et al. 2000). Interestingly, *pdf2-1hdg2-3* flowers display sepaloid petals (Kamata et al. 2013) and both genes have been shown to act together with *ATML1* (Abe et al. 2003; Peterson et al. 2013). These observations could suggest that HD-ZIP IV and SEP transcription factors are involved together in the regulation of flower organ development. The study of multiple mutant combinations between these gene families or expression levels of these genes in different mutant backgrounds will help to define their interactions and mechanism.

In order to study subtle embryonic epidermal defects I used the toluidine blue test described by (Tanaka et al. 2004; Xing et al. 2013). Most single and double mutant combinations tested showed no difference compared to wild-type, however seedling epidermal permeability defects were detected using single and double mutants of *hdg5-3* and *atml1-3* or *pdf2-3*. This result supports functional redundancy between these genes since the seedling phenotypes are stronger in the double mutants than in the respective single mutants, particularly in the case of *pdf2-3hdg5-3*. Defects in the permeability of seedlings could be related to defects in epidermal layer integrity as a consequence of disturbances to epidermal cell identity. This instability could be related to the cell wall composition, cuticle defects, or embryo cell disorganization originating from abnormal cell divisions. Interestingly, HDG5 is the least similar protein of the *Embryo* expression group to ATML1 and PDF2. The fact that *HDG5* nonetheless shows a similar enough expression pattern to ATML1 and PDF2 to be classed in the *Embryo* group, and in addition was identified as a potential interactor of ATML1 in my work, suggests that the relationship between these proteins requires further detailed analysis.

All genes clustered in the *Embryo* group are expressed in embryo and inflorescence and have overlapping expression domains with *ATML1/PDF2*. They are therefore the best candidates to be involved in epidermal specification. The results of my immunoprecipitation of GFP-ATML1 confirm that the proteins of the *Embryo* group and *ANL2* have the potential to work together in the specification of epidermis both during embryo development and in the inflorescence. However, these experiments only show binary protein-protein interactions with GFP-ATML1. The full identity of the complexes is unknown. There are several possibilities for regulatory combinations. The protein complexes could be dimers of GFP-ATML1/HD-ZIP IV or there could be higher order complexes involving combinations of different HD-ZIP IV transcription factors. The protein interactions ATML1/PDF2 and PDF2/PDF2 were validated using transgenic lines expressing combinations of FLAG and GFP tagged versions of ATML1 and PDF2. In further work it would be interesting to confirm interactions with other members of the protein family, and to use techniques such as Blue Native PAGE to ascertain the size of HD ZIP IV containing complexes.

To explore the hypothesis that HD-ZIP IV transcription factors are organized in protein complexes of more than two proteins, I used my nLC-MS/MS data to extrapolate the composition of the protein complexes involving ATML1-GFP; analysing the difference between two proteins in the same sample, with particular emphasis on GFP and each HD-ZIP IV transcription factor. The error associated with this comparative method includes the different length and sequence of the peptides, technical bias associated with efficiency of trypsin treatment, peptide affinity to the LC column and peptide identification efficiency by MS/MS. The fact that I am comparing proteins of the same family should reduce some of this inaccuracy. From the analysis of LFQ intensities corresponding to HD-ZIP IV proteins and GFP, in inflorescence and siliques datasets, I observed that ATML1 and PDF2 protein abundance is considerably higher than that of GFP and that the intensities of the other five HD-ZIP IV proteins (HDG1, 2, 5, 11, 12) are substantially lower than that of ATML1 and less than 5 times lower than GFP. From these observations it is possible to deduce that the complexes identified in these experiments are probably combinations of more than two HD-ZIP IV proteins. Interestingly the relative amount of ATML1/HDG changes between the silique and inflorescence dataset. For

instance, HDG11 is the least represented HDG interacting with ATML1 in silique samples and the most strongly represented in the inflorescences dataset. An opposite trend was observed for HDG2. It is therefore likely that the identity of the HD-ZIP IV complexes varies according to developmental stage and tissues studied.

ATML1 and PDF2 have a central role in epidermis specification

The best proof of the functional redundancy of *ATML1* and *PDF2* is the phenotype of the double mutant and the complementation of this lethal mutant with either ATML1 or PDF2 fusion proteins. The atml1-3pdf2-3 phenotype studied in this project is more severe than that described in the literature for *atml1-1pdf2-1*, since embryo development in *atml1-3pdf2-3* arrests at the globular stage with severe epidermal defects whereas *atml1-1pdf2-1* double mutants complete embryogenesis but are seedling lethal, producing embryos with a relatively normal root pole, which were capable of germination (Abe et al. 2003). Although the position of the T-DNA insertion in *pdf2-3* is similar to that in *pdf2-1*, the insertion in *atml1-3* is in the homeodomain encoding region of *ATML1*, whereas that in previously described *atml1-1* allele is more 3'. It is therefore possible that the *atml1-1* allele is not null, and that enough *ATML1* activity persists in this mutant to partially rescue the early embryo lethality which I observed in double mutants.

The lethal phenotype of *atml1pdf2* leaves several questions unanswered. How is seed development more generally affected by the loss of ATML1/PDF2? In the context of HD-ZIP IV redundancy, why are these two transcription factors so critical if other HD-ZIP IV proteins can act together with PDF2 and ATML1? In addition, because the loss of epidermis identity blocks embryo growth and *atml1-3pdf2-3* homozygous plants are never obtained, the effects of total loss of ATML1/PDF2 in other tissues or at other stages of plant development are also unknown. The solution for this problem may rely on the construction of inducible RNAi or artificial microRNA lines that could be used to disrupt *ATML1* and *PDF2* after germination and ideally at different stages of plant development. Alternatively the generation of sectors lacking the activity of both genes using a "PoPoff" strategy might be envisaged (Gilbertson 2003).

Dosage dependency of ATML1/PDF2

The number of functional *ATML1* and *PDF2* alleles has a significant impact on embryo and seedling development. The severity of development defects in embryogenesis is particularly strongly correlated with the number of functional gene copies of *PDF2* and *ATML1*, from weak phenotypes observed in single *pdf2-3* and *atml1-3* mutants to the total lethality observed in the *atml1-3pdf2-3* double mutant which arrests at the globular stage. This dosage dependency for *ATML1* and *PDF2* seems to be slightly less relevant in post-embryonic development since *pdf2-3*, *atml1-3*, *ppAa* and *aaPp* adult plants are globally similar in phenotype to wild-type plants. Thus, once epidermal cell fate is established, its maintenance may be less dependent on the number of functional *ATML1* and *PDF2* gene copies. At the other extreme, high concentrations of ATML1 and PDF2 disturb trichome branching, pavement cell lobbing, leaf epinasty, floral organ shape, and in severe cases leads to fertility loss. Additionally, the level of these two HD-ZIPs has a direct impact in the frequency of giant cells in sepals. The higher the quantity of ATML1 and PDF2, higher is the frequency of giant cells. In extreme cases the entire sepal is covered with giant cells.

At very early stages of embryogenesis, some embryos of *pdf2* and *atml1* single mutants show division problems in the embryo proper. These defects may be caused by identity problems, such as defective cell fate decisions between apical and basal identities. The changes in division plane orientations at early stages of development have consequences for apical/basal identity which will be discussed further in Chapter 5. Embryos that express only active copy of *PDF2* or *ATML1* (*aaPp* or *ppAa*) showed further cell division/organization problems in the epidermis, whereas the cell pattern of internal layers appeared comparable to the wild-type. Interestingly at the heart stage most embryos of these genotypes are viable. Together these results suggest that ATML1 and PDF2 function is more important, or at least less redundant, at early stages of development, when epidermis is being specified. It is possible that at this stage (Table 3.1) other HD-ZIP IVs compensate for the lack of either *ATML1* or *PDF2*, consistent with the observation that the majority of the genes encoding proteins of this class are expressed at the heart stage of embryogenesis.

It is important to point out that changes in the levels of ATML1/PDF2 may directly influence the activity of other HD-ZIP IVs which could compete for the same binding sites, be incorporated in ATML1/PDF2 containing complexes, or change their abundance because their genes are targets of ATML1/PDF2 (see Chapter 4). Therefore the different consequences of overexpression of *ATML1/PDF2* or knockout of *ATML1/PDF2* could be a reflection of a more global disturbance in the HD-ZIP IV balance.

Role of HD-ZIP IVs after germination

The post-germination phenotypes of *atml1* and *pdf2* single mutants and overexpressing lines indicate that there is also a dosage dependency for *ATML1/PDF2* and that the functional redundancy of *PDF2* and *ATML1* is not complete.

The sepal phenotypes of single mutants and *ATML1/PDF2* overexpressing lines are a textbook example of dosage dependency. Our results show that the number of giant cells produced on sepals is directly related to the abundance of ATML1/PDF2. In the L*er Arabidopsis* background, a mutant screen for phenotypes of sepals lacking giant cells resulted in the identification of *DEK1*, *ACR4* and two HD-ZIP IV genes *ATML1* and *HDG11* as being involved in giant cell identity (Roeder et al. 2012). Giant cells and small cells of the sepal have different division patterns and this process implies a tight coordination between cell division and cell identity (Roeder et al. 2012). Sepals of lines expressing *FLAG-ATML1* or *FLAG-PDF2* are almost covered with giant cells. This phenotype was also reported in lines over expressing *KRP1* (*KIP-RELATED PROTEIN 1)* and *LGO (LOSS OF GIANT CELLS FROM ORGANS*), cell cycle inhibitors, under the *ATML1* promoter, whilst the single *krp1* and *lgo* mutants show the opposite phenotype: lack of giant cells (Roeder et al. 2010; Roeder et al. 2012). The simplest explanation for this observation is that these two genes are direct targets of ATML1/PDF2/HDG11. Supporting this theory, the promoter of *KRP1* contains the L1box binding site of these transcription factors and the *LGO* promoter contains two truncated versions of the L1box (TAAATGY). The epidermal signalling pathway should be upstream of cell cycle inhibition, since the mutants of several genes of this pathway have identical phenotypes to *lgo* and *kpr1*. This

possibility is currently being tested in collaboration with A. Roeder. It is important to note that in the L*er* background, ERECTA, a receptor-like kinase which strongly influences growth, is not functional. This might influence the penetrance of mutant phenotypes. The *atml1-3* and *acr4-2* mutants used in this project are in the Col0 background and have a weaker sepal phenotype than that described in L*er*.

The leaf phenotypes (cauline and rosette leaves) of lines overexpressing FLAG fusion PDF2 or ATML1 proteins consist of epinastic leaf growth with mono or di-branched trichomes and less prominent lobes in the pavement cells. Interestingly, the double mutants (in the L*er* background) *atml1-2dek1-4*, *acr4-24dek1-4* and *hdg11-3dek1-4* showed hyponastic growth of cauline leaves (Roeder et al. 2012) whereas the *dek1-4* mutant shows a weak epinastic phenotype in the same organ. Roeder and colleagues therefore suggest that giant cells have a role in determining the curvature of sepals and leaves (Roeder et al. 2010; Roeder et al. 2012). Similarly, *hdg11-2* and *hdg11-2hdg12-2* trichomes show multiple branching (Nakamura et al. 2006), the opposite phenotype to that obtained in PDF2 and ATML1 over expressing lines. Trichome branching is also known to be strongly linked to cell cycle control and indeed several mutants that show loss of interdigitation of pavement cells also show trichome phenotypes, most frequently reduced size of branches, and leaf curling (reviewed in (Qian et al. 2009)).

The relationship between cell differentiation and cell cycle progression can be sequential, independent or coupled according to different cell fates (Jakoby and Schnittger 2004). In trichome development, initially there is a switch from mitosis to endoreduplication and simultaneous activation of several genes responsible for cell fate determination such as the HD-ZIP IV transcription factor-encoding *GL2* gene. Subsequently, cell expansion and branching are correlated with endoreduplication cycles (Grebe 2012; Hulskamp 2004; Schnittger and Hulskamp 2002). Endoreduplication is also linked to cell growth in epidermal pavement cells, and to the production of giant cells in several tissues. As a result mutants or overexpression lines of genes involved in the cell cycle often display phenotypes associated with the differentiation of such cell types (Borghi et al. 2010; Boudolf et al. 2004; Desvoyes et al. 2006; Gutzat et al. 2012; Roeder et al. 2010; Verkest et al. 2005), including

trichome branching defects and pavement cells phenotypes. Interestingly the loss of interdigitation is a phenotype of plants both overexpressing and underexpressing the calpain domain of *DEK1* (Johnson et al. 2008). In summary, these results suggest that the phenotypes of ATML1/PDF2 transgenic lines are likely related to defects in cell cycle control.

The very narrow leaves produced by plants overexpressing ATML1/PDF2 are also reminiscent of mutants with defects in abaxial/adaxial polarity. The study of mutant combinations of HD-ZIP III genes *REV*, *PHB* and *PHV* indicates that these genes are involved in the regulation of abaxial/adaxial polarity in leaves (Prigge et al. 2005). Could the overexpression of PDF2/ATML1 disturb the activity of HD-ZIP III proteins? One possibility is that ATML1/PDF2 could compete for the binding sites of HD-ZIP III proteins and thus change the regulation of their targets. This possibility could be tested by monitoring the expression of HD-ZIP III targets in ATML1/PDF2 overexpressing backgrounds. It would also be interesting to cross the dominant mutant *phb-1d* (McConnell and Barton 1998) with these transgenic lines.

The difference of phenotypes between FLAG and GFP fusion expressing lines is intriguing. It is possible that either the protein-protein interactions, or the protein-DNA interactions of different fusion proteins are altered, contributing to the phenotypes of transgenic lines. The adult phenotypes and sterility of FLAG fusion lines tend to be more pronounced than that of GFP lines. Interestingly, we also observed that overexpression of the transgene in FLAG lines tended to be higher than in GFP lines, which may, to some extent, explain differences in phenotypic strength. However both GFP and FLAG fusions proteins of ATML1 and PDF2 are sufficient to complement the *atml1pdf2* double mutant.

Our results show that the promoter of *PDF1* is activated, possibly by other *HD-ZIP IV* genes, in the absence of ATML1 and PDF2 at early stages of embryogenesis. Thus although the other *HD-ZIP IV* genes are not able to compensate the lack of *ATML1/PDF2* in epidermis specification, they can likely be expressed, together with *PDF1*, independently of ATML1 and PDF2. The study of seed development in the complemented *atml1pdf2* lines in the future will help in understanding the role of ATML1/PDF2 in apical/basal cell identity at early stages of embryogenesis. The

persistence of division problems in the embryo due to problems of cell identity would indicate that the *PDF1* promoter is not activated early enough to fully complement the early embryo phenotypes, whereas the absence of epidermal cell disorganization would indicate that the *PDF1* promoter is sufficient to drive L1 specific expression early in the embryo.

Chapter 4 – Targets of ATML1/PDF2

4.1 Introduction

The role of the transcription factors ATML1 and PDF2 in epidermal identity has been characterized phenotypically by the lethality of the double mutant, seedling defects and defects of embryogenesis both in single and double mutants (Chapter 3) (Abe et al. 2003). However the targets of these proteins remain unclear. In this chapter I have used two approaches to address this problem. Firstly I have used a theoretical approach consisting of predicting targets using bioinformatics tools. Additionally, I will try to compare *in silico* data with *in vivo* methods, namely ChIP and transcriptomic analysis of *pdf2-2* mutants and *PDF2* overexpressing lines, to identify direct targets of these genes and to validate the list of predicted targets.

The construction of a database of ATML1/PDF2 potential targets is based on the study of the *ATML1* (Takada and Jurgens 2007) and *ACR4* promoters. The presence of an L1 box (TAAATG(C/T)A) in the *ACR4* promoter and the fact that the expression of this gene is compromised in *atml1pdf2* double mutant suggest that *ACR4* is a potential target of ATML1/PDF2 (Abe et al. 2003; Tanaka et al. 2007). Part of the aim of this thesis is to analyse the relationship between *ACR4* and these two *HD-ZIP IV* genes. My *ACR4* promoter analysis is described in Chapter 5.

In the case of *ATML1*Takada and Jürgens identified a 101 bp long sequence from the *ATML1* promoter which was sufficient to drive expression of a GFP construct specifically in the L1 cell layer at different stages of embryogenesis. This sequence contains two regulatory regions that are both necessary for continuous epidermal expression in the embryo, since the mutation of either the L1 box (Abe et al. 2003; Abe et al. 2001) or WUSCHEL binding site (WUS1 TTAATGG) (Lohmann et al. 2001) causes gaps in L1 expression. Moreover, it was found that various combinations of promoter regions regulate different spatial and temporal components of *ATML1* expression during embryogenesis and the authors defend the existence of at least three specific regulation mechanisms for *ATML1* expression responsible for initial activation, maintenance and reactivation of expression at later stages of embryogenesis. (Takada and Jurgens 2007).

This study of the *ATML1* promoter demonstrates the complexity of regulation of *ATML1*. Since the WUS1 binding site is important for the regulation of *ATML1*, I hypothesise that some WOX (WUSCHEL RELATED HOMEOBOX) genes could be involved in the regulation of *ATML1* through interaction with WUS1. Some members of the WOX family are involved in embryo development and in some cases their expression pattern overlaps with that of *ATML1* (Haecker et al. 2004; Wu et al. 2007).

Several questions need to be answered regarding the complexity of *ATML1* regulation. Originally, a simple model of an autoregulatory loop had been proposed involving positive feedback of ATML1/PDF2 on their own promoters and positional cues from the cuticle or outer embryo wall, with a possible repressive loop in inner cells (Abe et al. 2003; Ingram 2004). However, the *ATML1* promoter study suggested a complex regulation where the spatial/temporal pattern of *ATML1* expression requires several different unknown mechanisms involving unknown regulators/signals. More than trying to find targets of ATML1/PDF2 I therefore also aimed to use bioinformatics approaches to test if there are other players that cooperate with these two genes either upstream or as partners to define and maintain epidermal identity. The HD-ZIP IV and WOX families are the strongest candidates for regulators.

4.2 Construction of a database of potential targets

The starting point for the construction of a database is to understand the needs of the user. My goal was to find the targets of two transcription factors using the predicted binding sites of these proteins. Therefore I needed to access the promoter sequences at a genomic level and analyse the presence of the motifs of interest. Additionally, since the aim of this project is to analyse the regulation of epidermal specification I aimed to use the same database to include other regulators so that I can analyse whether there is cooperation between transcription factors for the regulation of epidermal genes.

There are two postulates important for the choice of motifs. First, the binding site of ATML1/PDF2 is either an L1 box or a related sequence. Secondly, there is a

possibility that ATML1/PDF2 have partners, which could also regulate epidermal cell fate.

Regarding the first postulate, two forms of the basic L1 box (L1 box C (TAAATGCA) and L1 box T (TAAATGCA)) have been considered, and an additional three motifs which were found in the literature have also been included. The transcription factors PDF2 and ATML1 have been shown to bind *in vitro* to the L1 box motif (Abe et al. 2003; Abe et al. 2001). From the analysis of the HD-ZIP IV family made by Nakamura and colleagues (Nakamura et al. 2006), two sequences similar to the L1 box were identified. One of the motifs, here named HDZIPIV, (GCAT***TAAATGC***) was identified using three members of HD-ZIP IV family, namely HDG9, HDG7 and ATML1, and consists of a pseudo palindromic sequence. The other was identified using PDF2, so in this work it is named PDF2* (GCNN***TAAATGC***). Finally, another derivative of the L1 box was described as the binding site of ATML1/PDF2 according to the program ATHAMAP (Bulow et al. 2009). The sequence of this motif was named after its source (Table 4.1).

Table 4.1 – List of motifs used in the database of potential targets (Busch et al. 2010; Lohmann et al. 2001; Nakamura et al. 2006; Ulmasov et al. 1999)

Name	Sequence	
L1 box T	***TAAATG***TA	
L1 box C	***TAAATGC***A	
HDZIPIV	GCAT***TAAATGC***	**PDF2 like motifs**
PDF2*	GCNN***TAAATGC***	
ATHAMAP	G***TAAATGC***AC	
WUS1	***TTAATGG***	
WUS1*	***TTAAT***(C/***G***)(C/***G***)	
WUS2*	***CACGTG***	
WUS2	T***CACGTG***A	
ARE	TGTCTC	

As mentioned in the introduction of this chapter, the study of the *ATML1* promoter shows that both a WUS1 sequence and an L1 box are necessary for L1 expression in embryo development. This means that the L1 box is not sufficient for the entire L1 specific expression pattern (Takada and Jurgens 2007). However, the phenotype of

the double mutant *atml1pdf2* suggests that these genes are necessary for epidermal identity (Abe et al. 2003). Therefore, the L1 box and by extension the roles of ATML1 and PDF2 although necessary, may not be sufficient for L1 patterning in embryo. This supports the theory of a partner involved in epidermal cell fate regulation. The main candidates for this partner are WOX proteins that could bind to a WUS1 sequence. To explore this possibility further, I considered other potential WOX binding sites. Another description of WUS binding site, TTAAT(G/C)(G/C) (Lohmann et al. 2001), is less restrictive and here has been named WUS1*. An independent study confirmed that WUS binds both to the WUS1 sequence and surprisingly to the G-box element (Menkens et al. 1995). This study used EMSA (electrophoretic mobility shift assays) with purified WUS protein and probe sequences derived from genomic analysis (ChIP-chip data) of chromatin regions bound by WUS *in vivo* (Busch et al. 2010). The two versions of the G-box identified are here denominated WUS2 (TCACGTGA) and WUS2* (CACGTG) standing for WUS binding site 2.

In addition to potential HD-ZIP IV and WOX binding sites, I have also included the cognate AUXIN RESPONSE FACTOR (ARF) binding site, which is known as the (ARE) AUXIN RESPONSE ELEMENT (Ulmasov et al. 1999), since there are five AREs (TGTCTC) in the promoter of *ACR4*, the potential target of ATML1/PDF2. In addition, the analysis of the presence of this element constituted part of our ongoing collaboration with other members of the SIREN consortium, a European consortium interested in early embryogenesis, which financed this thesis.

In summary, the chosen motifs include 5 possible binding sites of ATML1/PDF2 found in the literature, 3 predicted binding sites of WUS and WOX's and the binding site of ARFs (Table 4.1).

In addition to choosing the motifs to analyse, it was necessary to identify the best bulk sequence database and the program most suitable to perform the motif matching. For *Arabidopsis Thaliana* TAIR10 is the most recent and complete genome database available (TAIR stands for *The Arabidopsis Information Resource*). Regarding upstream sequences there are six options covering different sizes and starting point (Table 4.2, Figure 4.1). The size options are 500, 1000 and 3000 bp of

upstream sequence. As for the type of sequence, one option is to consider sequences initiating at the starting point of translation, whilst the other considers sequences that start at the beginning of UTR region (for the genes where the 5' UTR information is not available the start of translation is used).

Table 4.2 –Representation of gene promoters that contain each motif using TAIR or RSAT upstream sequence databases (Thomas-Chollier et al. 2011).

	UTR START			Translation START			
Database	tair500	tair1000	tair3000	tair500s	tair1000s	tair3000s	RSAT
genome	33602	33602	33602	27416	27416	27416	27204
one motif	17903	25703	32959	14782	21139	26980	25659
L1box	2773	5105	11787	1943	3992	9663	7037
L1boxC	987	1940	5197	724	1522	4237	2894
L1boxT	1843	3375	7925	1252	2623	6499	4687
hdzipiv	9	23	63	5	14	40	26
PDF2	74	132	385	50	96	293	191
athamap	16	26	106	16	26	75	49
WUS1	3830	6591	14570	3065	5413	12066	9017
WUS2	249	477	1159	220	419	1027	720
WUS1*	9224	15301	26983	7308	12372	22051	18368
WUS2*	3253	5033	10508	2544	4161	8900	6697
ARF	7408	13299	26726	6827	11556	22224	18211

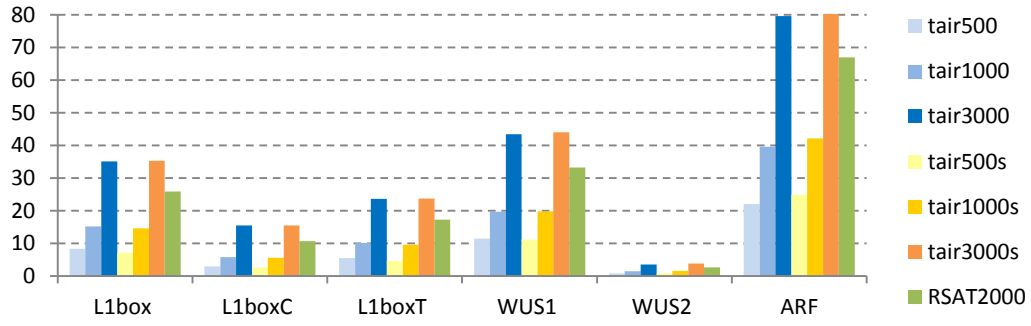


Figure 4.1 – Representation of the percentage of gene promoters that contain each motif using different TAIR and RSAT upstream sequence databases.

Several programs are available for matching known motifs to different databases. Most programs allow the user to submit their own database; the restriction being the size of the database both in terms of length and number of sequences. For genomic analysis, and specifically regulatory motif sequence analysis, I found the best

program to be RSAT, which stands for *Regulatory Sequence Analysis Tools* (Thomas-Chollier et al. 2011). Besides the ***DNA pattern*** function, here used to match the motifs in Table 4.1 to TAIR10 upstream sequence databases, RSAT also includes its own genomic analysis. This function is available for several genomes including *Arabidopsis thaliana*. The user can choose the length of the sequence. The starting point is the translation start. The program refers to NCBI (National Center for Biotechnology Information) which in return uses TAIR10. Therefore an additional database was analysed using a sequence length of 2000 bp. The result of genomic analysis of different databases is summarized in Table 4.2 and Figure 4.1.

From the comparison of different databases one can conclude that the number of motif matches increases with the length of upstream sequence, which is to be expected. However, the escalation tendency is not proportional to upstream sequence length, and the probability of matching any given motif is not equal. Concerning motif sequence, the number of occurrences is dependent both on the length and the proportion of T/A vs G/C. An example that illustrates this statement is the comparison of WUS2 and L1boxT, both of which have the same length but which have a difference in occurrence of approximately six fold. Additionally, there is a significant difference between L1boxC and L1boxT, which only differ in one base (Figure 4.1, Table 4.2).

For the purpose of simplification, the tair1000 and RSAT2000 databases were chosen to continue motif analysis. The information from the other databases has been kept and can be accessed using the ATG identity of each gene. The advantage of tair1000 is that it excludes the 3'UTR. Frequently regulatory motifs are not present in UTR region. In addition, the impact of intron sequences upstream of the start codon is prevented. Moreover the size of 1000 bp is very practical. Adding RSAT2000, there is a gain in coverage given the length and information from the UTR region. As a result, I have created a database that includes the data from both in order to take into account these different features. All the results that follow refer to this database that covers 30329 genes that contain at least one of the motifs described in Table 4.1.

The foundation of the database of potential targets is accomplished with the genomic display of motifs of interest. There are several features that can be added to the database to include more information and ease of access (Table 4.3). Firstly the description of each gene and its nomenclature were obtained using the feature *Gene Description* in TAIR. This information can be used to track gene families, query for keywords of interest or simply to find genes more efficiently. Secondly, to assist the search for new regulators, combinations of different motifs were made. The idea is to look for a combination with biological interest that may indicate the two transcription factors work together to regulate a group of genes.

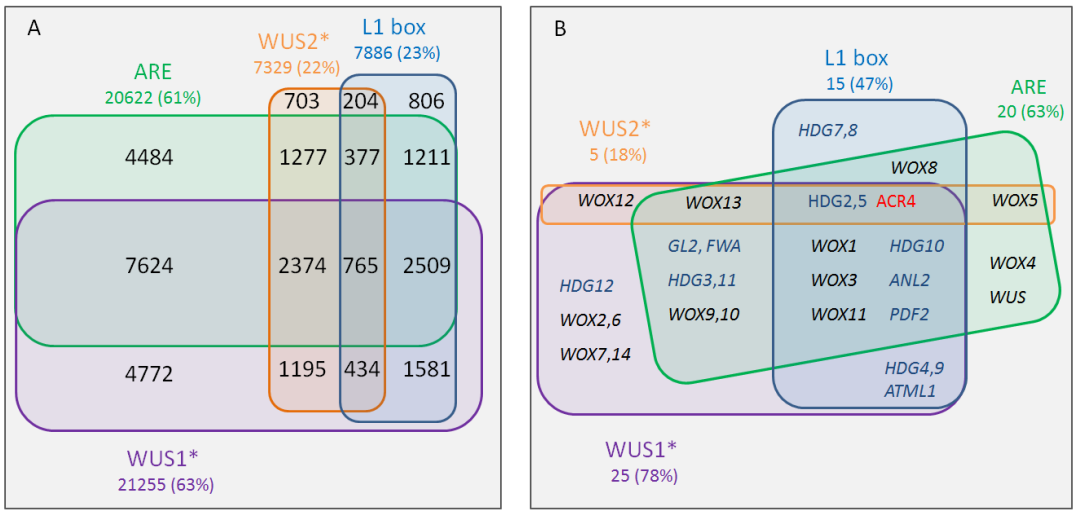
Table 4.3 – Summary of the information available in the database of potential targets of ATML1/PDF2. This table indicates the total number of genes that correspond to each feature. For each gene the database includes information of which motifs, or motif combinations, are present in the respective promoter, and if the gene is described by GO and PO terms of interest obtained using VirtualPlant 1.2 software (Katari et al. 2010).

Motifs		Motif combination		GO terms		PO terms	
L1box	7886	L1/no WUS	5262	Endomembrane system	3916	Shoot epidermis	102
L1boxC	3218	L1C/T	585	Organelle part	2489	Epidermis	278
L1boxT	5253	WUS ½	243	Response to stimulus	3498	Meristem	331
WUS1	10143	L1/WUS1	2512	Cell wall	556	Embryo development	50
WUS1*	21255	L1/WUS2	177	Lipid metabolic process	709	Seed development	83
WUS2	765	L1/WUS1/2	65	ATPase activity, coupled to movement of substances	163	Endosperm	36
WUS2*	7329	L1/WUS	2624	Developmental process	1765	Seed	104
ARF	20622	L1/ARF	4861	Regulation of transcription	1593		
Athamap	52	L1/ARF/WUS	1610	Signalling pathway	726		
Hdzipiv	35						
PDF2	223						

The evaluation of the datasets encompasses analysis of gene function annotations. There are several programs/tools available to run functional analysis. The program VirtualPlant 1.2 software platform (Katari et al. 2010) includes several tools to analyse, integrate, and see genomic data in a user friendly approach from a systems biology angle. The greatest advantage of this program over others is that it was developed specifically to study *Arabidopsis* data. A modest gene function analysis is

included in my database using information assembled by this software. This information includes some GO (Gene Ontology) and PO (Plant Ontology) terms (Table 4.3). I have optimized these with analysis of coexpression data for *ACR4*, *ATML1* and *PDF2* (Section 4.3), and transcriptional data (Section 4.4). The database is presented as a Microsoft Excel file to permit easy access for the user (TARGET DATABASE.xlsx file).

So far I have shown the number of gene promoters in the database that contain each motif (Figure 4.1, Table 4.3). However the representation of motif combinations is more realistic, it depicts better the distribution of potential DNA binding sites in the promoters at a genomic level and is a good tool to analyse the expected frequency of a given combination of DNA binding sites (Figure 4.2).



C

	All genes	*HD-ZIP IV*	*WOX*	*HD &ACR4*
no BS	9.8 %	6.3 % ↓	-	3.1 % ↓
1 BS	32 %	18.7 % ↓	40 % ↑	28.1 % ↓
2 BS	39 %	43.8 % ↑	33.3 % ↓	37.5 % ↑
3 BS	16.9 %	18.7 % ↑	26.7% ↑	21.9 % ↑
4 BS	2.3 %	12.5 % ↑	-	9.4 % ↑

Figure 4.2 – Representation of motif combinations. A) Illustration of the number of gene promoters that contain different combinations of L1box (blue box), WUS1* (purple box), WUS2* (orange box) and ARE (green box) binding sites. B) Illustration of motif combinations between the same DNA-binding sites present in the promoters of *WOX* (black), *HD-ZIP IV* (blue) and *ACR4* (red) genes. C) Comparison of proportion of gene promoters that have 0-4 different binding sites (BS) between HD-ZIP IV, WOX and the genomic level. The percentages refer to genomic level (33602), HD-ZIP IV (16), WOX (15), or both HD (31) gene promoters.

As the representation of all 11 motifs is impractical, I chose four motifs that represent the two potential binding sites of WOX transcription factors (WUS1*, WUS2*), the predicted binding site of ATML1/PDF2 (L1 box), and ARE. In order to gain more information I chose WUS1*/WUS2* and not WUS1/WUS2 since their sequences allow more matches. The number of promoters that contain each motif is represented by boxes, and the box area is correlated with the number of genes. In addition, I show the percentage of gene promoters that contains each motif at a genomic level (Figure 4.2A). While 85 % of all gene promoters contain ARE or WUS1*, only about 1/5 of the promoters contain L1 box, and the same proportion of promoters contain WUS2*. The largest combination of motifs containing the L1 box is the L1box/ARE/WUS1* combination and this could imply a relevant regulatory combination. However, in the group of genes that contain L1 boxes, the proportion of promoters that contain an ARE, WUS1* or both, is approximately the same as the genomic level. The same is observed for the group of promoters that contain WUS2* (Figure 4.2A). While only 2.3 % of the gene promoters contain the four binding sites, around 1/3 of gene promoters contain only one out of four binding sites and 39 % of genes contain 2 binding sites (Figure 4.2C).

As mentioned in Chapter 3, some HD-ZIP IV family members can form complexes with ATML1/PDF2 and are also thought to be involved in epidermal cell fate specification. From my ATML1 IP experiments (Chapter 3) seven HD-ZIP IV proteins are good candidates to form regulatory complexes with ATML1 (ATML1, PDF2, HDG1, HDG2, HDG5, HDG11, HDG12, and ANL2). The genes encoding these proteins are also present in the *ATML1*/*PDF2* coexpression lists (Section 4.3). These proteins could be involved in same pathways as *ATML1*/*PDF2* but could also be involved in different pathways. Additionally, it has been suggested that *ATML1*/*PDF2* are autoregulated (Abe et al. 2003). If this feature is common in other HD-ZIP IVs, it would create a complex regulatory network in this family. In addition, some *HD-ZIP IV* genes are potentially directly regulated by ATML1/PDF2. To explore similar features in the *WOX* family, and identify potential interactions between these two families of transcription factors, I analysed the presence of WUS binding sites, L1 box and ARE within the promoters of both families (Figure 4.2B). *ACR4* is also represented in this analysis as a likely target of ATML1/PDF2.

HDG1 is the only gene amongst the 16 *HD-ZIP IV* and 15 *WOX* genes which has none of the four motifs in its promoter. The number of promoters of *HD-ZIP IV* genes that contain the WUS1* binding site (81 %) or L1 box (62 %) is significantly higher than expected (Figure 4.2 A,B). This suggests that HD-ZIP IV transcription factors regulate their own expression and possibly act in cooperation with WOXs. In the WOX family, only the presence of the WUS1* binding site is significantly higher than expected (73 %) suggesting autoregulation within this family. In addition to the four *WOX* promoters that contain L1 boxes, the *WOX2* and *WOX9* promoters contain a shorter version of L1 box (TAAATGT) which indicates that HD-ZIP IV transcription factors could regulate *WOX* genes. Particularly, the expression pattern of *WOX2, WOX8, WOX3* and *WOX9* genes overlaps partially with that of *ATML1* in embryogenesis (Haecker et al. 2004; Takada and Jurgens 2007) raising the possibility that they could be regulated by ATML1/PDF2.

For both families the presence of WUS2* and AREs in the promoters is either comparable to the genomic level or inferior, suggesting that overall the regulation of *WOX* or *HD-ZIP IV* might be independent of auxin response and that the WUS2* binding site is less relevant than WUS1* (Figure 4.2B). Nevertheless, for *HD-ZIP IV* genes the proportion of promoters containing combinations of multiple motifs is higher than expected, and for *WOX* genes the proportion of promoters containing three binding sites is significantly higher than at the genomic level (Figure 4.2C). These results are consistent with the hypothesis that there is cooperation between these two classes of homeodomain transcription factors. Moreover, they might be regulated by different regulatory programs defined by specific combinations of DNA binding sites. Most likely there are multiple combinations of regulatory elements involved in the regulation of *HD-ZIP IV* genes, as previously demonstrated for *ATML1*. Interestingly, if extrapolated from genome-wide frequencies, the probability of *WOX* or *HD-ZIP IV* promoters containing the four regulatory elements included in my analysis is very low, this situation nonetheless exists for *HDG2* and *HDG5* as well as *ACR4*.

In summary, the main feature of the database that I have created is that by querying a given gene, one can obtain information about motifs of interest, combinations of

these motifs and functional annotations derived both from the TAIR description and GO or PO terms. In the context of this work I aimed to validate some of this data *in planta* so as to confirm the identification of novel targets.

4.3 Analysis of ATML1/PDF2/ACR4 coexpression data

The study of coexpression data of transcription factors can be very useful as a means of finding new targets and new regulators. In the context of this work I will base the classification of potential direct targets on the presence of an L1 box in their promoter. In terms of the identification of new regulators of epidermal cell fate, I searched for overrepresented motifs in the promoters of genes coexpressed with ATML1, PDF2 and ACR4. Finally I wanted to analyse combinations of regulatory motifs, with the aim of finding regulatory patterns (Section 4.5).

4.3.1 Datasets for coexpression

Coexpression data was obtained from ATTED II ver. 6.0 (Obayashi et al. 2009). This program uses publicly available micro-array data to evaluate the degree of similarity of expression between every pair of genes present on the arrays commonly used. In particular, ATTED II developed the *Mutual Rank* measure of gene coexpression, based on the geometric mean of the correlation of each pair of genes (Obayashi et al. 2009). The coexpression lists correspond to the best correlated gene pairs. It is possible to choose cut-offs at the top 300, 500, 1000 or 2000 correlated genes and also to access coexpression data in different conditions (tissue, abiotic and biotic stress, hormone and light). Furthermore, different features are available using the gene name as a query, namely gene and protein descriptions, GO terms, coexpression networks and expression profiles during development, stress and hormone response.

In this section I propose to analyse coexpression data for *ATML1*, *PDF2* and *ACR4* (described in COEXP.xlsx file) aiming to find potential regulators and novel ATML1/PDF2 targets. One approach consisted of analysing the coexpression gene network (20 genes) around one gene provided by ATTED II v6.0 (Obayashi et al. 2009). This integrates known information of protein-protein interactions (red dots) with coexpression (solid lines), that together will provide more information to comprehend gene function networks (Figure 4.3).

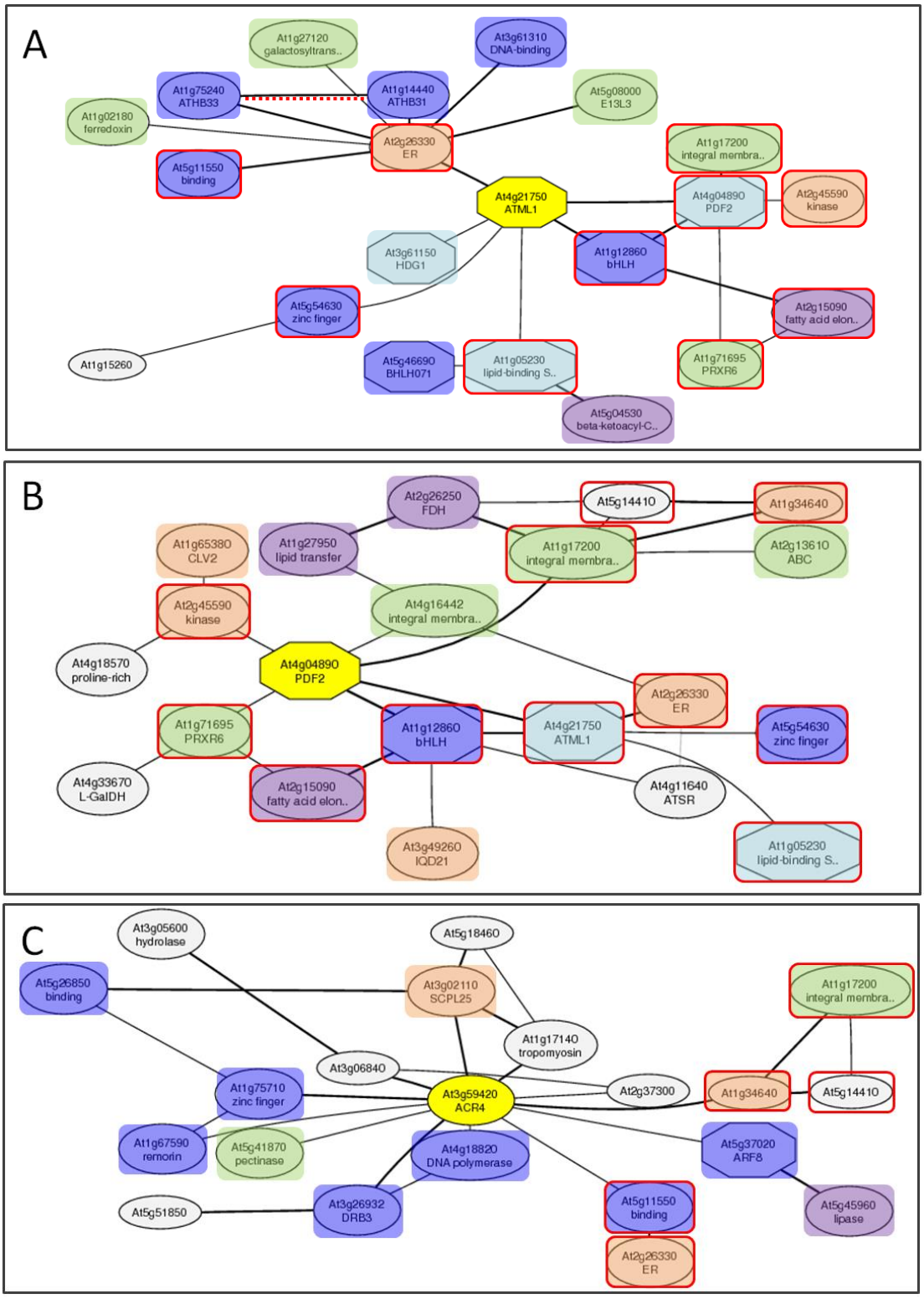


Figure 4.3 – Coexpression gene networks for three genes: ATML1 (A), PDF2 (B) and ACR4 (C). The core of each network is indicated in yellow. Genes are highlighted according to their protein function: regulation (blue and HD-ZIP IV in light blue), signalling (orange), cell wall and membrane associated proteins (green), and lipid process (violet). Red outline shows genes shared between networks. (Adapted from (Obayashi et al. 2009))

First of all I can classify most genes in the coexpression networks (Figure 4.3) into four main groups of genes encoding i) transcriptional regulators, ii) signalling proteins iii) lipid associated proteins and iv) proteins associated with cell wall and membranes. Unsurprisingly, around half of the genes present in the *ATML1* network are also present in the *PDF2* network (Figure 4.3) which supports the fact that these two genes act redundantly. Interestingly, two HD-ZIP IVs (*HDG1* and *HDG2*) and SCRM2 are present in the *PDF2/ATML1* networks. The other transcription factors include zinc finger, homeodomain or bHLH domain proteins. The proteins encoded by these genes are good candidates for co-regulators. The signalling group includes receptor kinases, and in particular ERECTA, which is present in all three networks.

Considering the expression pattern of *ATML1*/*PDF2*/*ACR4* and their close relationship, I decided to analyse the three lists of coexpressed genes together. Therefore I used the lists of 300 coexpressed genes and proceeded to analyse the connections between them (Figure 4.4).

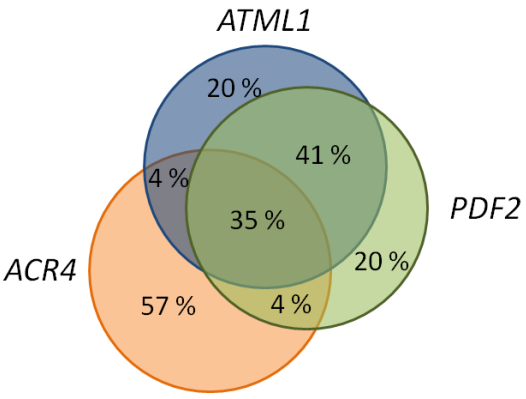


Figure 4.4 – Representation of overlap between coexpressed genes of *ACR4, ATML1* and *PDF2*. The data was taken from 300 coexpressed gene lists from ATTED II (Obayashi et al. 2009).

As expected a substantial number of genes are coexpressed with both ATML1 and PDF2 (around three quarters), which is again in accordance with the observed redundancy of these genes. Given that the ACR4 expression pattern overlaps with ATML1/PDF2, it is logical that their coexpressed gene lists overlap to a significant degree. Interestingly 1/5 of genes coexpressed with PDF2 or ATML1 are specific to ATML1 or PDF2, which indicates that the expression patterns and responses of these genes do not fully overlap. Nevertheless, the functional annotations of these groups

of genes did not reveal any biological function that could be specific to ATML1 or PDF2. The next step was to consider the rank of the genes according to their coexpression factor. The tendency is for top ranked genes to belong to more than one dataset. So, in general genes with a lower coexpression correlation appear only in one list. In total the three 300 gene coexpression lists associated with *ATML1, PDF2* and *ACR4* contain 542 genes. This list was named *CoexpA*. Additionally, I made a shorter list to restrict my selection to more significant results. Postulating that direct targets and possible regulators will score in the best positions I made an average of the rank of each gene in the three lists, and sorted the 100 best ranked genes. To this list I added 12 genes found relevant either by position or gene function. Nine further genes were strongly coexpressed with *ATML1*/*PDF2* but their coexpression factor with ACR4 was high. These were also included. Two of the remaining genes were coexpressed with *ATML1* and *PDF2* and could be involved in lipid metabolism. And the last gene is coexpressed with *ATML1* and *PDF2* and its promoter contains L1box, WUS1 and ARF. It is potential target of these proteins. This list of 112 genes was named *CoexpB*.

The next step in coexpression data analysis was to analyse predicted gene function. For the purpose of functional analysis of gene lists, VirtualPlant 1.2 (Katari et al. 2010) contains the tool BioMaps which uses text searching to analyse the overrepresentation (Fisher Exact Test, p-value cut off of 0.01) of keywords either from GO assignments by TAIR or functional classification by MIPS (Munich Information Center for Protein Sequence). The output is in the form of a table or network. This analysis was carried out for each list of genes coexpressed with *ACR4*, *ATML1* and *PDF2*, as well as for the lists *CoexpA*, *CoexpB* and *CoexpC* (which was constructed the same way as *CoexpA* but using data from the top 500 coexpressed gene pairs). With this information I can investigate target candidates easily, by analysing groups with the same function or association. Another type of analysis involved analysing isolated keywords that are overrepresented in one of the lists (Table 4.4).

Table 4.4 – Comparative functional analysis of different coexpression datasets using Biomap tool of VirtualPlant 2.1 (Katari et al. 2010). The lack of GO terms in coexpression lists is represented by (-).

GO terms	Genome		Observed %				
	Genes	%	*CoexpB*	*CoexpA*	*ACR4*	*ATML1*	*PDF2*
Lipid metabolic process	709	2.5	10.9	7.7	8.5	8.3	6.9
Very-long-chain fatty acid metabolic process	22	0.1	-	-	-	1.4	-
Organic acid metabolic process	664	2.4	9.1	5.6	-	6.6	6.2
Plasma membrane part	220	0.8	-	-	-	3.1	3.4
ATPase activity, coupled to transmembrane movement of substances	163	0.6	4.5	-	-	-	-
Chloroplast part	961	3.4	-	6.9	-	9.7	11
Cell differentiation	89	0.3	4.5	2.3	-	3.8	3.4
Developmental process	1765	6.3	15.5	12.9	12	16.3	13.4
Reproductive developmental process	796	2.9	-	6.3	7	7.3	-
Regulation of transcription	1593	5.7	15.5	13.8	12.7	14.9	13.4
Nucleic acid binding transcription factor activity	1678	6	18.2	15	15.5	16.3	14.1
Signalling pathway	726	2.6	-	-	-	6.6	-

The lists *CoexpA* and *CoexpC* essentially include the information of the other lists and therefore they will not be considered for the analysis of individual groups of GO terms. From the network generated from the *CoexpB* dataset, based on biological processes, there is a significant enrichment in lipid-related processes, regulation, plant development and cell differentiation. As for the molecular function, the most noteworthy overexpressed terms are transcription factors and hydrolases. This tendency is in general followed in all the other six lists. Interestingly, three groups of GO terms related to cell components, membranes or organelles are present in the *ATML1* and *PDF2* datasets but not in *ACR4* or *CoexpB*. Moreover, several GO terms associated with signalling, in particular receptor kinase activity are significant in all gene lists except *CoexpB*. The overrepresentation of genes involved in lipid metabolism and membrane/organelles could be related to cuticle biosynthesis, since lipid metabolism is deviated for cuticle deposition in epidermal cells. Moreover, a list of genes found to be important in cuticle formation was assembled in a recent review on epidermal cell fate and function (Javelle et al. 2011b) and 1/3 of these genes are present in my coexpression data. Furthermore, since cell fate specification and maintenance requires a strong network of signalling and specific transcriptional

regulation to establish a functional and robust epidermis, some important players might be amongst the receptor kinases and transcription factors identified in the coexpression data.

From the assembly of GO terms from the different datasets, 12 terms were selected as strongly over-represented in coexpression data (Table 4.4). The choice was made taking into consideration both the terms overrepresented in *CoexpB* and specific terms from the *ATML1/PDF2* coexpression lists. I also tried to cover different aspects of the GO classification (localization, development, and molecular function). Specific to *CoexpB* which includes the genes most significantly coexpressed with *ACR4, ATML1 and PDF2*, is a group of five genes involved in ATPase activity coupled to transmembrane movement of substances, which again could be related to biosynthesis of cuticle/cell wall. The fact that lipid and membrane related GO terms are overrepresented in these coexpression data is encouraging since I expect ATML1/PDF2 to regulate epidermal events that could include the formation of the cuticle layer. The maize homolog of these transcription factors, OCL1, is involved in the regulation of cuticle biosynthesis and lipid metabolism (Javelle et al. 2010). Additionally a small group of epidermal genes related to stomatal complex morphogenesis was found overrepresented in all lists (5 out of 9 genes), consistent with recent studies that indicate that ectopic expression of *ATML1* is sufficient to induce stomata cell fate in inner cells (Peterson et al. 2013; Takada et al. 2013). Compared to the genomic background level, in the coexpression lists GO terms related to development and transcriptional regulation were enriched more than two fold with high statistical significance. For the *Coexp500* list, the term "embryonic development" was also enriched underlining the importance of genes coexpressed with *ATML1/PDF2/ACR4* in embryogenesis.

4.3.2 Promoter analysis of coexpression data

Assuming that coexpression lists of the two transcription factors ATML1 and PDF2 contain several of their targets, it is expected that the predicted binding site of ATML1/PDF2 should be over-represented in the promoters of genes on these lists. The same inference works for other predicted regulators of epidermal cell fate.

Particularly, the proposed role of WOX proteins could be supported by the presence WUS binding sites in these sequences.

I therefore compared the coexpression lists, with the genomic level data to evaluate if there was a significant enrichment in specific elements or element combinations (Figure 4.5). For instance, since the L1 box is thought to be the binding site of ATML1/PDF2 and that targets of these genes are expected to be present in coexpression lists, I would predict that the percentage of genes containing an L1 box in their promoter should be greater in these lists than at the genomic level. In the following analysis I assume that an increase in percentage correlates to a significant enrichment of presence of a motif in the different datasets.

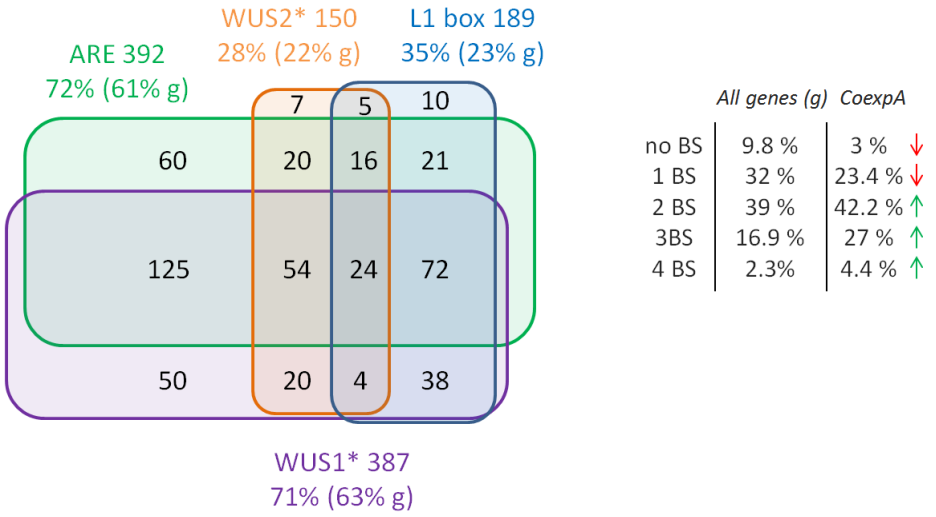


Figure 4.5 – Representation of motif combinations in promoters of genes coexpressed with *ATML1*, *PDF2* and *ACR4*. On the left is represented the number of *CoexpA* gene promoters (percentages refer to 542 gene promoters) that contain different combinations of L1box (blue box), WUS1* (purple box), WUS2* (orange box) and ARE (green box) binding sites. Moreover, the proportion of *CoexpA* gene promoters that have 0-4 different binding sites (BS) is shown on the right, and the red and green arrows represent the decrease or increase of the frequency of gene promoter with the same features compared to the genomic background (g) proportions corresponding to 33602 gene promoters (Figure 4.2 A,C).

The frequency of gene promoters containing each motif shown on Table 4.1 was higher in the *CoexpA* list and usually even higher in the *CoexpB* list compared to genomic background (expected frequency). In particular, in *CoexpA* and *CoexpB* lists, gene promoters containing the L1 box correspond respectively to 35 and 40 % while the expected frequency is 23 %. This increase in motif frequency could well be

correlated with the function of ATML1/PDF2 as transcription factors. The fact that they bind to L1box like motifs (Abe et al. 2001; Nakamura et al. 2006) supports the hypothesis that some direct targets of these genes are present in the coexpression lists.

The representation of different combinations of WUS1*, WUS2*, ARE and L1box localized in gene promoters of *CoexpA* (Figure 4.5) shows that the number of promoters that lack these four motifs, or contain only one of the named motifs is lower than expected (Figure 4.3C). Conversely the frequency of promoters containing more than 2 motifs is significantly higher than the expected frequency. The same tendency is seen for *CoexpB* with similar distribution. The analysis of each combination indicates that four motif combinations are particularly highly over represented in the lists *CoexpA* or *CoexpB* compared to the expected frequency while remaining motif combinations are either as expected, or decreased in frequency (Figure 4.1 and 4.5). One of the enriched combinations corresponds to promoters that contain all four motifs (5.3 % for *CoexpB*). Two combinations of three motifs are also enriched in the coexpression lists, namely L1box/ARE/WUS1* and WOX2*/ARE/WUS1* (17 and 8.9 % for *CoexpB* respectively). Finally promoters containing an L1box and WUS1* are also more common in both coexpression lists (10.7 % for *CoexpB*).

In summary, the analysis of HD-ZIP IV and WOX binding sites in promoters of genes coexpressed with *ATML1, PDF2* and *ACR4* supports the hypothesis that these transcription factors are important for epidermal cell fate specification. Combinations between these binding sites are more noteworthy in coexpression data compared to at the genomic level, in agreement with the hypothesis that WOXs and HD-ZIP IVs could act in cooperation.

4.4 An *in vivo* approach to finding new ATML1/PDF2 targets

ChIP (Chromatin Immunoprecipitation) is the gold standard for finding targets of transcription factors. The technique consists of studying the binding sites of a given transcription factor by enriching these regions using immunoprecipitation of the protein (and associated DNA) from plant tissue. When coupled with a powerful

sequencing method one can retrieve reliable information on all the direct targets of a transcription factor of interest in a given tissue. I used this approach in order to try and identify direct targets of ATML1 and PDF2. In order to provide complementary information I used a transcriptome based approach comparing *pdf2* mutant and PDF2 overexpressing lines with wild-type plants in order to analyse significant fluctuations in target expression. Using these two approaches I hoped to identify new targets of ATML1 and PDF2.

4.4.1 PDF2 and ATML1 bind to their own promoters and the promoter of *ACR4.*

Both ATML1 and PDF2 have been proved to be involved in establishment and maintenance of epidermal cell fate. These two proteins bind in vitro to L1 box binding site which are present in some L1 specific genes (Abe et al. 2003; Abe et al. 2001). These genes are therefore predicted to be potential targets of ATML1, PDF2 and other HD-ZIP IV transcription factors. However, the identity of these targets remains to be proven *in vivo*. Using techniques like ChIP not only should it be possible to confirm potential targets of ATML1/PDF2 but also to discover new targets and reveal new players or pathways involved in epidermal cell fate specification.

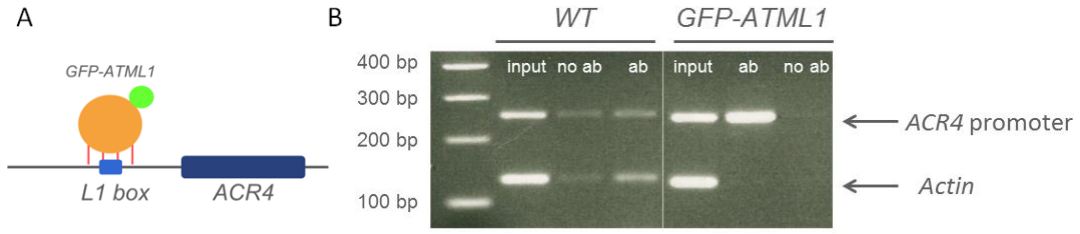


Figure 4.6 – ACR4 is a target of ATML1. A) Schematic representation of ATML1-GFP binding to ACR4 promoter L1box, B) ChIP results for ACR4 promoter. Two strains were tested: pATML1-GFP:ATML1 and Col0 wild-type. The ChIP controls are *ACTIN* (representing genetic background) and two negative controls: the input sample (DNA before the ChIP step) and a sample where no antibody was added.

The principle of ChIP is to detect *in vivo* binding sites of a specific tagged protein (Figure 4.6A). The first step is to fix all DNA-protein interactions. The DNA is then fragmented and the fragments to which the transcription factor/DNA binding protein of interest is bound are selected and purified. The selection step is usually made using specific antibodies that interact with a tag (for example GFP), or directly with

the DNA-binding protein of interest (Gendrel et al. 2005; Kuo and Allis 1999). This technique can be used to prove *in vivo* a particular interaction or in a genomic approach to discover targets of the transcription factor.

In this project I used this technique on lines expressing GFP tagged ATML1 and PDF2 driven by different promoters (Section 2.6). Plants already existed which expressed GFP:ATML1 under *pFDH,* and *pATML1* (Gifford et al. 2005), and I was involved in generate plants expressing GFP:ATML1 and GFP:PDF2 under *pPDF1*. The chosen protocol was adapted for inflorescence and seedling tissue. I successfully prove that *ACR4* is a direct target of ATML1 using both seedling and inflorescence tissue (Figure 4.6B). The enrichment of *ACR4* promoter PCR product compared to *ACTIN* was observed in pATML1-GFP:ATML1 but not in Col0 or input (sample taken before ChIP step) and no antibody (no ab) controls. Before sequencing ChIP samples I wanted to analyse if other potential targets were also enriched. Therefore I tested, unsuccessfully, *PDF1*, *FDH*, and *PDF2* promoter fragments containing an L1box. Because of the unreproducibility of ChIP results using this material I initiated experiments using using lines expressing GFP:ATML1 and GFP:PDF2 under the *PDF1* promoter depicted in Figure 4.7.

Using transgenic plants expressing *pPDF1-GFP:ATML1* (79) or *pPDF1-GFP:PDF2* (83) and wild-type plants as control I performed ChIP experiments using inflorescences and siliques (Figure 4.7). As described before I used two internal controls: a sample collected before the ChIP step (orange) and another sample where no antibody was added at the immunoprecipitation step (blue). The most significant result from these experiments is the enrichment of at least 2 fold of promoter fragments from *ACR4*, *ATML1* and *PDF2* in ChIP of silique samples expressing *GFP:PDF2* (green), which indicates that PDF2 should regulate *ATML1*, *ACR4* and its own expression during embryogenesis or (more-likely) during seed-coat development. The same tendency can be observed in one of the samples of inflorescences (83-7) also expressing *GFP:PDF2*, but the enrichment is lower. The experiments using lines expressing GFP:ATML1 show only a very weak enrichment, insufficient to assure that ATML1 regulates *ACR4* and its own expression. The analysis of the levels of *PDF1* suggests that this gene is also regulated by PDF2 in

inflorescences, and the fact that no enrichment was observed in siliques may be due to the fact that *PDF1* is not expressed in the developing seed coat. It should also be noted that quantifying enrichment at the *PDF1* promoter is complicated by the fact that the transgenes are driven by the *PDF1* promoter. This explains the high background levels in the transgenic lines compared to wild-type lines using primers in this promoter. These levels suggest that, despite the fact that these lines underwent antibiotic selection for single insertion sites, there are several (up to 8 or 9) copies of the transgene in each transgenic line.

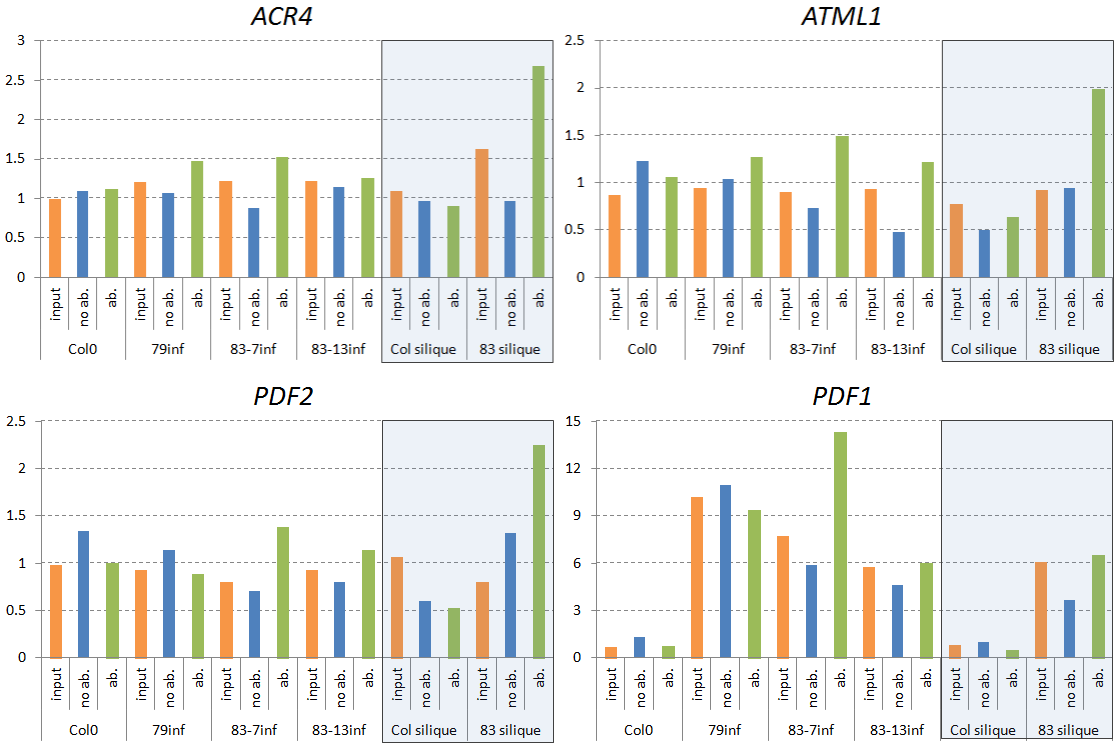


Figure 4.7 – ChIP experiments using plants expressing GFP fusions of ATML1 and PDF2. For each biological sample of wild-type or transgenic lines expressing GFP-ATML1 (79) or GFP-PDF2 (83) driven by *PDF1* promoter I analyzed three samples: input (before ChIP step shown in orange), ChIP sample (ab, green) and a negative control where no antibody was added (no ab, blue). In addition two plant materials were used: inflorescences and siliques (grey box). I analyzed the DNA enrichment of the promoter regions of *ACR4*, *ATML1*, *PDF2*, and *PDF1*, comparing wild-type with transgenic lines and ChIP sample with other controls. Q-PCR is normalized to *EIF2* expression.

The general lack of success of ChIP experiments could be explained by competition to bind L1 boxes both from the endogenous protein and other HD-ZIP IV TFs. The same explanation may also explain the non-reproducibility of the results.

Introduction of transgenes into mutant backgrounds has recently been completed and ChIP experiments in complemented mutant backgrounds are now underway.

In conclusion, my ChIP results suggest that ATML1 and PDF2 are autoregulated and that both bind directly to the promoter of *ACR4*. However, technical improvements need to be completed before ChIP-SEQ approaches can be envisaged.

4.4.2 A transcriptomic analysis to uncover potential targets of ATML1/PDF2

Investigation of gene expression at the whole transcriptome level has many applications. The idea of a transcriptomic approach is to compare a biological system in different states. In plants researchers may be interested in comparing different tissues, different conditions (abiotic/biotic stress, temperature, light, etc.) and different genetic backgrounds. The result of transcriptomic approaches can be the discovery of new pathways or gene clusters responsible for a given trait/response. The basis of a transcriptomic approach is to choose the best material or conditions to address the biological question.

Our biological question was the identification of the targets of ATML1/PDF2. A transcriptomic approach was applied to different backgrounds as a method of finding both direct and indirect targets of PDF2. By analysing expression levels in the three lines wild-type/*pdf2*/*pPDF1-FLAG:PDF2*, I hoped to identify genes that are over/under expressed in the mutant or overexpressor line compared to the wild-type system. Moreover I tried to identify gene clusters with similar function which were co-regulated in this experiment.

To study the transcriptional role of PDF2 as gene repressor or activator I analysed independently the genes upregulated and downregulated in Col0 vs *pdf2* and Col0 vs *pPDF1-FLAG:PDF2*. The tissues used were inflorescence tips pooled from several plants grown under identical conditions. Unfortunately, for economic reasons, I was not able to carry out biological replicates. I was particularly interested in finding genes both upregulated in *pPDF1-FLAG:PDF2* and downregulated in *pdf2*, or both downregulated in *pPDF1-FLAG:PDF2* or upregulated in *pdf2*, as I would consider such genes as likely to be direct targets of PDF2 regulation. The data analysed in this

section can be found in PDF2-RNAseq.xlsx file. The RNA samples of the three backgrounds were sequenced by GATC Biotech and the raw data was uploaded to Galaxy software (Section 2.5) that includes several programs to process the transcript assembly, estimate RPKM (reads per kilobase per million) and test for differential expression and regulation in RNA-Seq data. The chosen output is an Excel file displaying the RPKM of each sample, the fold change between two samples (wild-type was used as reference), p-value and q-value for each gene. This process was carried out twice for Col0 vs *pdf2* and Col0 vs *pPDF1-FLAG:PDF2*. I included two different cut-offs of the data to build a restrictive dataset and an informative dataset. The cut-off of RPKM and p-value (lower than 0.05) was the same for both datasets. For the RPKM at least one of the samples was required to have a RPKM over 5. The selection of the two datasets was made with the constraint of fold change. The first restriction was log2 of fold change superior to 0.3 which is equivalent to a 1.23 ratio and the respective dataset was named *SEQ1*. While the second restriction was a 2 fold change and therefore the dataset was named *SEQ2*. Since I analysed *pdf2* and *pPDF1-FLAG:PDF2* separately I obtained four lists *SEQ1-M* and *SEQ2-M* using Col0 vs *pdf2* data, or *SEQ1-O* and *SEQ2-O* using Col0 vs *pPDF1-FLAG:PDF2* data. As expected *PDF2* is present in all four lists, confirming that *PDF2* is overexpressed (2.8 times) in *pPDF1-FLAG:PDF2* expressing plants, consistent with Q-PCR results (Chapter 3).

The next step was to analyse groups of genes that are upregulated or down regulated in each list. These genes are potential direct or indirect targets of PDF2. Therefore I created 8 lists isolating the genes upregulated (up) and downregulated (down) from the above lists: *SEQ1-Mup, SEQ1-Mdown, SEQ2-Mup, SEQ2-Mdown, SEQ1-Oup, SEQ1-Odown, SEQ2-Oup* and *SEQ2-Odown*. It is expected that genes that are transcriptionally activated by PDF2 are upregulated in plants expressing *pPDF1-FLAG:PDF2* and downregulated in *pdf2*. On the contrary the genes that are repressed by PDF2 activity should be upregulated in *pdf2* and downregulated in *pPDF1-FLAG:PDF2* overexpressing lines. I grouped the genes that fulfil these requirements by overlapping *SEQ1-O* and S*EQ1-M*. The list *SEQ1-PAP* corresponds to the genes that are potentially activated by PDF2 and *SEQ1-PRP* corresponds to the genes that are potentially repressed by PDF2. When I overlapped *SEQ1-O* and *SEQ1-M* I

verified around 20 % of the genes were present in both lists. In this group of genes 25% correspond to *SEQ1-PAP* (200 genes), 4 % correspond to *SEQ1-PRP* (34 genes) and 63 % correspond to genes that are downregulated in *PDF2* overexpressing lines and *pdf2* (this list was named *SEQ1-OMdown*). If the number of genes potentially activated by PDF2 is higher to those potentially repressed, the data support a global activator role for PDF2. This is specifically significant when I restrict the expression difference to 2 fold.

To find gene clusters regulated by PDF2, I used the program Biomaps from VirtualPlant 1.2 and analysed the different lists of genes. As mentioned before, this tool shows the overrepresentation of GO terms in a gene list compared to a genomic level. For this analysis I included all GO terms for each list in the same table in order to find patterns in this functional analysis. I considered that *SEQ1-Mdown, SEQ2-Mdown, SEQ1-Oup, SEQ2-Oup* and *SEQ1-PAP* were related to a transcription activation role of PDF2, whereas *SEQ1-Mup, SEQ2-Mup, SEQ1-Odown, SEQ2-Odown* and *SEQ1-PRP* were related to transcription repression role of PDF2. I used the lists *SEQ2-M* and *SEQ2-O* to perform a functional analysis of potential direct and indirect targets, free from the association of up or downregulation and to gain information about the significance of functional analysis. In addition I included the list *SEQ1-OMdown*. The lists *SEQ2-Odown* and *SEQ1-PRP* were each enriched with one GO term related to cell wall and endomembrane system respectively, and *SEQ2-Mup* (23 genes) showed no functional enrichment. Overall I found three groups of GO terms associated with an activator or repressor role of PDF2 and a group of GO terms that was not specific to either role of PDF2 and in which most terms were enriched in *SEQ1-OMdown* list.

I found that genes potentially activated by PDF2 are enriched for the terms related to stress response, "transport", biological regulation, cell wall, "biopolymer modification", "apoplast" and cell death. Interestingly, the groups of genes defined by stress response, apoplast and cell death are also present in the 2 fold data (*SEQ2-M* or *SEQ2-O*). Genes likely to be repressed by PDF2, include genes involved in cell growth and genes encoding constituent parts of organelles or the nucleolus. However, none of these GO terms was overrepresented in the 2 fold data. Moreover,

genes associated with GO terms related to "lipid metabolic processes" and "extracellular regions" are found to be downregulated in both the mutant and the overexpressor. However the significance of these terms is arguable since neither is present in 2 fold data analysis and about half the genes considered are present in the group of genes downregulated in both mutant/overexpressor (*SEQ1-OMdown*).

The most convincing link with PDF2 activity is stress response. The epidermis is the first barrier against many abiotic and biotic stresses. From my analysis of potential target function, stress response is the more significant term and is not only specific for genes potentially activated by PDF2 but is also significantly overrepresented with low p-values and a difference to genomic background of at least 3 fold. Additionally a possible link between PDF2 regulation and epidermal signalling is supported by a fairly significant over representation of genes associated with phosphorylation/ receptor activity, transport and cell wall/endomembrane functions needed for signal transduction. It is must be remembered however, that a proportion of the genes associated with signalling events overlap with those associated with stress response. Similar overlaps are found for other overrepresented terms.

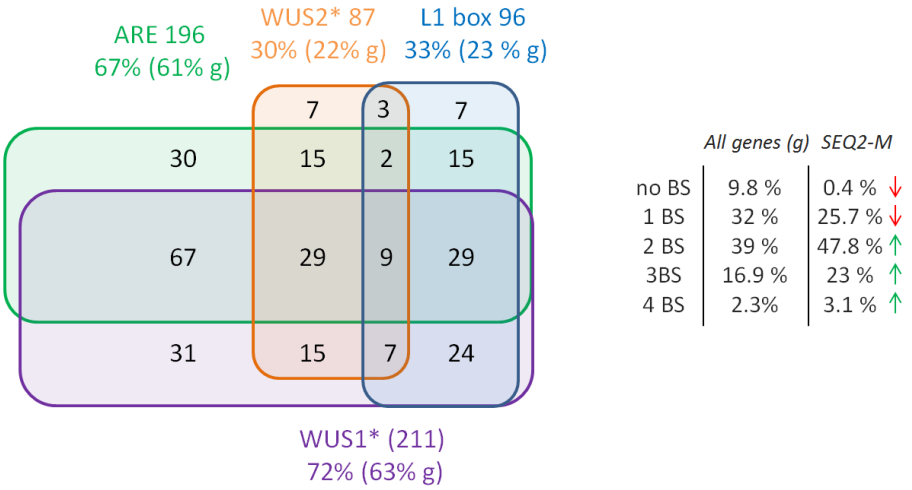


Figure 4.8 – Representation of motif combinations in promoters of genes up or down regulated in *pdf2*. On the left is represented the number of *SEQ2-M* gene promoters (percentages refer to 291) that contain different combinations of L1box (blue box), WUS1* (purple box), WUS2* (orange box) and ARE (green box) binding sites. Moreover, the proportion of *SEQ2-M* gene promoters that have 0-4 different binding sites (BS) is shown on the right, and the red and green arrows represent the decrease or increase of the frequency of gene promoter with the same features compared to the genomic background (g) percentages corresponding to 33602 gene promoters (Figure 4.2 A,C).

Interestingly a relationship between PDF2 regulation and cuticle/lipid biosynthesis is not strongly supported. Some hints from functional analysis regarding lipid metabolism and organelle constitution could, in fact, suggest a repressive role for PDF2 in the deviation of lipid metabolism to cuticle biosynthesis. However, these functional associations have low significance.

In the context of this project it was important to understand if WOX and HD-ZIP IV binding sites were overrepresented in the promoters of genes up or downregulated in *pdf2* (Figure 4.8) or lines expressing *pPDF1-FLAG:PDF2*. Therefore I analysed the presence of the motifs WUS1*, WUS2*, ARE and L1 box in the different *SEQ1* and *SEQ2* lists. In general the frequency of either motif or motif combination was higher than the expected frequency (Figure 4.3) with the exception of the ARE binding site. In particular I show the results for the *SEQ2-M* list in Figure 4.8 which indicates that the proportion of gene promoters containing at least 2 motifs is higher than expected. In particular, the combinations including an L1 box show higher frequency than in the genomic background. Interestingly the frequency of WUS2* and respective combinations is higher than the expected frequency, which could be related to the fact that WUS2* is also the G-box binding site which is related to stress response. Once again these results suggest that PDF2 acts in cooperation with other regulators.

4.5 Network of epidermal cell fate regulation

The aim of this chapter is to expand the knowledge of how to specify epidermal identity. In this section I combine my findings using motif analysis, transcriptomic and coexpression data. I will start with a motif analysis aimed at discovering new regulators of epidermal specification that are either regulated by *ATML1/PDF2* or participate in association with these genes. Later, I will present a list of potential targets of ATML1/PDF2.

4.5.1 Exploring regulatory regions to find new regulators

The role of ATML1/PDF2 in epidermal specification is irrefutable. However, the range of their regulatory functions is unclear. The question of the exact role of ATML1/PDF2 as key/master regulators at the core of epidermal specification, or indeed as important pieces of a complex network, remains unclear. In this section I

propose to analyse the promoter regions of *ATML1*/*PDF2*/*ACR4* coexpressed genes and up and downregulated genes in *pdf2*/*pPDF1-FLAG:PDF2* backgrounds to find regulatory regions overrepresented that could help identify other regulators of epidermal cell fate.

The study of regulatory regions consisted of analysing motifs of a determined length (6 to 8 bp) that are overrepresented in gene promoters sharing the same trait: either coexpressed with a given gene or up/down regulated in *pdf2*/*pPDF1-FLAG:PDF2*. The analysis was made using the program RSAT (Thomas-Chollier et al. 2011) and applying two methods to calculate background frequencies: Markov model order 2 and a predefined background based on upstream sequences of *Arabidopsis thaliana*. The data used to describe and analyse overrepresented motifs is indicated in REGULATORS.xlsx file.

Once overrepresented motifs are identified, the next step is to compare them to known motifs described in different databases: AtCoecis, PLACE, AGRIS, ATTED II and DATF (Higo et al. 1999; Obayashi et al. 2009; Vandepoele et al. 2009; Yilmaz et al. 2011). The results for 6 bp motif analysis are summarized in Table 4.5. The evaluation of motif match is qualitative and encompasses a parameter for quality of match (overrepresentation significance, combination of motifs, motif length) and two other elements that indicate which data was used to find the overrepresented motif, namely coexpression lists (*CoexpA, CoexpB*, *ATML1*, *PDF2* or *ACR4* coexpression lists of 300 genes) or transcriptomic analysis (*SEQ2-M, SEQ2-O, SEQ1-PAP,* or *SEQ1-PRP*).

The most important finding from this analysis is that both the L1 box and two WUS binding sites are overrepresented in both coexpression and transcriptomic data supporting the theory that WOX's and HD-ZIP IVs could act together in specification of epidermal cell fate. Different combinations of regulatory elements could be associated with regulation of specific groups of genes. To investigate this theory I analysed overrepresentation of GO terms in lists of genes containing each of four motifs (L1 box, WUS1, WUS2, ARE) and combinations of these motifs with L1 box (all at a genomic level) using VirtualPlant 1.2. Most combinations show overrepresentation of genes described by GO terms such as "transcription factors",

"development process", "biological regulation" and "catalytic activity". Additionally, if L1box and WUS2 (CACGTG) are combined, the terms "response to stimulus" and "membrane" are significantly overrepresented. Less significant is the appearance of "lipid metabolism" in the list L1/ARE/WUS. The analysis of promoters of genes potentially activated by PDF2 in relation to these three types of motifs is consistent with coexpression data (Section 4.3), since the frequency of these motifs both individually and combined, is higher than expected at genomic level.

Table 4.5 – Regulatory motifs overrepresented in coexpression and transcriptomic data. The motifs were scored (*** higher score) for matching quality (M), overrepresentation in coexpression data (C) and overrepresentation in transcriptomic data (T).

Name	Sequence	TF	M	C	T	Description/reference
Homeobox						
L1 box	taaatcya	HD-ZIP IV	***	***	**	Binding site of HD-ZIP IV (Abe et al. 2001; Nakamura et al. 2006)
WUS1	ttaatss	WOX	***	***	*	Binding site of WUS (Lohmann et al. 2001)
WUS2/ G-box	cacgtg	WOX/ bZIP/ GBF	***	***	***	Binding site of WUS and stress response element (Busch et al. 2010; Hudson and Quail 2003; Menkens et al. 1995; Terzaghi and Cashmore 1995)
ATHB1	caatwattg	HD-ZIP	**	**	***	HD-ZIP I, HD-ZIP II binding site (Harris et al. 2011; Johannesson et al. 2001; Sessa et al. 1993)
ATHB2	caatsattg	HD-ZIP	**	*	***	
ATHB6	caattatta	Homeo domain	**	*	***	ABA signal pathway (Himmelbach et al. 2002)
Response to ABA /water stress						
ABRE	yacgtggc	bZIP	***	-	***	ABA responsive element (Choi et al. 2000)
ABRE-like	bacgtgkm		**	-	***	Dehydration (Shinozaki and Yamaguchi-Shinozaki 2000)
GCC-box	gccgcc	AP2-EREBP	***	-	***	Dehydration (Shinozaki and Yamaguchi-Shinozaki 2000)
DPBF1&2	acacnng	bZIP	**	**	-	ABA response and embryo-specification (Kim et al. 1997)
AtMYC2	cacatg	MYC	***	***	-	Involved in dehydration response (Abe et al. 1997; Simpson et al. 2003)
RY repeat	catgca		***	-	***	Seed-specific activity of the napA promoter regulated by ABA (Ezcurra et al. 1999)
Light response						
Box II/G box	acgtggc		**	-	***	Light response/ABA response (Block et al. 1990; Nakashima et al. 2006; Terzaghi and Cashmore 1995)
EVENINGAT	aaaatatct		***	-	**	Circadian control (Harmer et al. 2000)
SORLIP4	gtatgatgg		*	**	**	Motifs found *in silico* by comparison of induced (SORLIP) and repressed (SORLREP) phyA-responsive promoters (Hudson and Quail 2003)
SORLIP5	gagtgag		**	**	-	
SORLREP1	ttwtactagt		**	*	**	
SORLREP2	ataaaacgt		*	-	**	
SORLREP3	tgtatatat		**	**	-	
SORLREP4	ctcctaatt		*	*	*	
SORLREP5	ttgcatgact		*	*	*	
0thers						
W-box	ttgact/ tgacy	WRKY	***	-	***	Plant defense (Nishiuchi et al. 2004; Yu et al. 2001)
IRO2OS	cacgtgg	bHLH	***	-	***	Fe deficiency stress (Ogo et al. 2006)
MYB3	taactaac	MYB	**	***	*	Environmental stress (Chen et al. 2002)
MYB4	amcwamc		**	***	*	
HSEs	agaannttct	HSF	*	**	-	Heat stress (Nover et al. 2001)

In addition to the L1 box and WUS binding sites three potential binding sites for homeodomain proteins (ATHB1, ATHB2 and ATHB6) described in the literature to be involved in hormone response, match some motifs overrepresented in both coexpression datasets and transcriptomic analysis.

Interestingly, a significant number of regulatory motifs associated with different stress and hormonal responses are overrepresented both in transcriptomic and coexpression data (Table 4.5). In particular, elements involved in response to environmental stress such as water stress, light response, heat shock and abscisic acid (ABA) signalling are well represented among the matching regulatory motifs. These results taken together with the functional annotation of transcriptomic results, provide some evidence that there is a real connection between stress response and ATML1/PDF2 regulatory activity. Light stress responsive genes are also significantly overrepresented in the list of genes coexpressed either with *ATML1*, *PDF2* or *ACR4*. Thus, regulators of stress response either cooperate with ATML1/PDF2 in activation of stress genes, or are directly regulated by these two proteins.

I found binding sites of several classes of transcription factors that could correspond to regulators that either act in cooperation with ATML1/PDF2 or are direct targets of these two transcription factors regulating groups of genes involved in epidermal cell fate. The most relevant transcription factors probably belong to the homeodomain protein family and, interestingly, the most relevant functional annotation seems to be stress response.

4.5.2 Identification of potential targets of HD-ZIP IV proteins

This section summarizes the information obtained to identify the potential targets of ATML1/PDF2. Since my ChIP-SEQ experiments did not go as planned, I use the information of transcriptomic analysis of *pdf2*/*pPDF1-GFP:PDF2* in inflorescences, together with coexpression data for *ATML1/PDF2/ACR4* available in ATTED II (Obayashi et al. 2009). Moreover I selected some potential targets to validate by Q-PCR.

The selection of potential targets took into consideration both transcriptomic and coexpression data (Table 4.6). The first criterion is the presence of the proposed ATML1/PDF2 binding site, the L1 box. In order to gain more information, I used the shortest version of the L1 box TAAATCY. Regarding the transcriptomic data, I used 2 fold lists as a threshold (*SEQ2-M* and *SEQ2-P* lists) with two main exceptions: genes coexpressed with ATML1/PDF2 and genes in the *SEQ1-PAP* list. Additionally I checked if the genes had an expression pattern similar to *PDF2* in inflorescences using the Arabidopsis eFP browser (Winter et al. 2007). When coexpression data was used to identify potential targets, I checked for the presence of L1 box in gene promoters, and analysed whether the expression pattern was similar to that of ATML1/PDF2 in inflorescence and embryo. I preferentially selected genes that either belonged to ATML1/PDF2 coexpression network or had good coexpression coefficients relative to both ATML1 and PDF2 (TARGET DATABASE.xlsx file).

I found several candidates for transcription factor family binding motifs overrepresented in transcriptomic and coexpression data. In particular, homeodomain proteins, WRKY, bZIP, and bHLH transcription factors were identified. Even though the regulation mechanism is unknown, it is possible that either these regulators cooperate with ATML1/PDF2 for the activation/repression of a group of genes, or act downstream of ATML1/PDF2 and therefore are direct targets of these two proteins. The hypothesis that these regulators could be direct targets of ATML1/PDF2 implies that genes necessary for the specification of epidermis are regulated indirectly by ATML1/PDF2 and therefore need no L1 box in their promoter. For these reason I have included some potential target genes that do not contain an L1 box when relevant to a particular gene family or functional category (Table 4.6).

In addition to transcription factors, I grouped potential target genes into six groups. The first group corresponds to genes involved in stress and hormone response. It must be noted that some genes classed in the “response” cluster can also belong to one of the other groups, but in order to avoid duplicates I decided to include them in response.

Table 4.6 –Potential targets of ATML1/PDF2. The genes in blue issue from coexpression data, whereas genes in black issue from transcriptomic data and genes in grey are present in both datasets (*no L1 box TAAATCY present in promoter region).

Class	Genes
Epidermis	*PDF2 ATML1 FDH PDF1 ACR4 LTP1*
Transcription factors	
Homeodomain	*HDG2 HDG5 ANL2 HDG11* HB7 HB21* HB28 HB34 HAT3 HDG12 HDG7 WOX9* HB12* HAT14*
NAC	*NAC019* NAC036 ATNAC6 NAC3* NAC001* NAC046 NAC103 NAC102**
WRKY	*WRKY18 WRKY46 WRKY6 WRKY48 WRKY54 WRKY60 WRKY33*
bZIP	*SCR HY5* ATBZIP34 ATBZIP61*
bHLH	*SPCH SCRM2 SCRM AT4G20970 AT3G20640 AT1G62975**
Other TFs	*SPL4 SPL9 NGA1 NGA2 FIL YAB3 .AT1G75710 AT5G54630 MYB106 MYB62* MYB31**
Response	
Water stress	*COR47 ERD10 ERD6 ERD7* ERD4* ERD14 ERD15* RAB18* DDF1 SRO5*
Hormones	*ARF18 ARF1 ARF8 ARF6* IAA8 IAA2 SAG12 NIT2 PIN1 JAZ1 OPR3 HVA22D* HVA22E PEN1 COR414-TM1*
Chemical stress	*ERF2 ETR2 AT5G61590 FSD1 WIN1*
Biotic stress/defence	*MLP43 RLP37 RLP33 RLP32* PME41 HR4 NAI1*
Temperature	*GOLS3 HS90.1**
Kinase/receptor	*RLK TMM PK1B ER ERL2 RLK4 CRK2*
Calcium binding	*IQD21 IQD22 IQD17 APRR2 MAP18 PBP1 TCH3 CRCK1 ATPLC4 CML38 AT1G21540*
Cuticle/lipid	*CER6 LTPG1 CER5 BDG1 GPAT8 CER3 CER8 ABCG32 PEL3 PLDGAMMA3 LOX3 KCS17 CER4 WBC19 LACS9 LACS3 LTP4 LTP3*
Membrane/cell wall	*EPF2 scpl25 AT1G34640 PAP29 AT3G05470 PSB29 PDCB3 ABCG11 NAT12 ATCSLC8 FLA13 FLA15 FLA2* BGLU23 TCH4 XTH9* XTH16 AT2G20870 AT4G20830 SRC2*
Other	*NRPD1B BCAT-1 AT1G17200 ICR1 DRB3 DRB5 QWRF1 QWRF2 ATSR OFP15 AT5G14410 AT5G13400 AT4G18570 AT1G12330 AT3G61550 SVB RALFL20 RALFL30 RALFL15 RALFL19*

The group of response genes possibly activated by PDF2 comprises five subgroups; namely hormone response, water stress, temperature stress, chemical stress and biotic stress. The most significant subgroup corresponds to the early-responsive to dehydration (ERD) genes. Although, these genes have different biochemical functions and cell localization, they are characterized by a fast response to dehydration (Alves et al. 2011). In our transcriptomic data 13 out of 16 *ERD* genes are significantly down regulated in *pdf2* or upregulated in *pPDF1-GFP:PDF2*. Six of these genes have similar spatial expression patterns to *PDF2* and show at least a 2 fold reduction in expression in *pdf2* mutants. A significant group of potential targets of ATML1/PDF2 (direct or indirect) seem to be related to signalling events. Not only did I find genes involved in several different types of response signalling, I also

found a group of genes related to calcium signalling, and receptor kinases that could have important roles in establishing and maintaining epidermal identity.

Finally, I selected a group of genes associated with membrane and cell wall functions and another of genes involved in cuticle or lipid biosynthesis. However, with exception of the lipoxygenase *LOX3* (repressed by PDF2) and the phospholipase-encoding *PLDGAMMA3* (activated by PDF2), the expression variation of other members of these groups is inferior to 2 fold in my transcriptome data. Although several genes involved in cuticle and lipid biosynthesis have an L1 box in their promoter and are present in the coexpression lists of *ATML1/PDF2*, when it comes to transcriptomic data their expression either is unchanged or not significantly up/down regulated. The best example is the gene *FDH* involved in lipid biosynthesis (Yephremov et al. 1999) which is expressed specifically in the L1 layer with an expression pattern similar to *PDF2/ATML1* and is in the top 5 of genes coexpressed with *ATML1/PDF2* according to ATTED II. In transcriptomic analysis *FDH* has approximately the same expression level in *pdf2*, wild-type or *pPDF1-FLAG:PDF2*. Either the transcript level of these genes is highly controlled, being compensated by ATML1 or other HD ZIP IVs, or the pathways behind cuticle deposition are independent of PDF2/ATML1 regulation.

4.6 Discussion

In this chapter, I was interested in discovering the targets of two HD-ZIP IV transcription factors: ATML1 and PDF2. Although single mutants seem largely aphenotypic (suggesting redundancy) the double mutant of these genes is lethal. By comparison of the promoter sequence of epidermal genes, an 8 bp motif named the L1 box had previously been identified and proven *in vitro* to be bound by these two proteins (Abe et al. 2003; Abe et al. 2001). Sequences similar to the L1 box were also found in DNA-binding site selection for ATML1, PDF2 and other HD-ZIP IV proteins (Nakamura et al. 2006). Assuming that the L1 box is the binding site of HD-ZIP IV proteins, and taken together with the facts that several *HD-ZIP IV* genes are expressed specifically in epidermis, and that the L1 box is present in the promoters of most *HD-ZIP IV*-encoding genes and several L1 specific genes, I can propose that genes containing an L1 box are potential targets of HD-ZIP IV proteins.

Finding biological targets of PDF2/ATML1: a tricky proposition

Intuitively, one might expect that genes involved in cuticle biogenesis, stomatal formation and trichome specification or morphology would be good candidates for HD-ZIP IV regulation since these traits are specific to the epidermis, and considerable evidence exists linking the activity of HD-ZIP IV proteins to the formation of these structures (discussed in the general introduction). Our analysis of *ATML1/PDF2* coexpression data supports a role of these two genes in cuticle biosynthesis regulation. Several genes involved in cuticle formation and transport, or lipid metabolism, contain an L1 box in their promoters and have good coexpression coefficients relative to *ATML1* and *PDF2*. However, the expression level of these genes is steady both in *pdf2* and *PDF2* overexpressor lines compared to wild-type conditions. One might expect a direct target of PDF2 to be either up or down regulated in presence of a mutant line or overexpressor of the respective transcription factor. However, other HD-ZIP IV proteins likely act together with PDF2 to regulate these genes, and may compensate for the loss of *PDF2* in mutant backgrounds (see below). Furthermore, by definition, gene coexpression is a similarity of gene expression patterns, which in our biological system could generate confusion between genes which are really HD-ZIP IV regulated, and genes which simply have an epidermal expression pattern. In other words, coexpression relationships between PDF2/ATML1 and genes involved in cuticle, stomata, or trichome might reflect their epidermis specificity more than direct regulation.

One of the challenges in finding the direct targets of ATML1/PDF2 is to overcome the functional redundancy between these two proteins and other HD-ZIP IV family members. The most direct *in vivo* method to identify targets of a transcription factor is ChIP-SEQ. Using ChIP I confirmed that *ACR4* is likely a direct target of ATML1/PDF2 using *pATML1-GFP:ATML1* (inflorescences and seedlings) and *pPDF1-GFP:ATML1* (siliques and inflorescences) transgenic lines. Additionally, the ChIP results suggest that *ATML1* and *PDF2* regulate themselves and each other. However I could not reproducibly confirm other potential targets such as *FDH*, and *PDF1*. The lack of reproducibility in ChIP experiments may be caused by competition between GFP tagged ATML1/PDF2, endogenous ATML1/PDF2 and

other HD-ZIP IVs. In order to avoid this problem one could use the same transgenic lines in mutant and double mutant backgrounds. This material has been generated for some transgenic lines (Chapter 3) and will be used in future experiments.

As an alternative to ChIP-SEQ, I used a preliminary transcriptomic analysis of inflorescences of *pdf2*/wild-type/*pPDF1-FLAG:PDF2* to identify PDF2 targets indirectly. The expression of the *PDF1* promoter is very strong in inflorescences (Chapter 3). However I will have lost information on potential embryo specific PDF2 targets by using this approach. In addition, the fact that there is significant redundancy between *HD-ZIP IV* genes may again impede the identification of PDF2 targets using this approach. Genetic and phenotypic analysis suggests that *ATML1* globally compensates for the lack of *PDF2* in *pdf2* mutants. I have also shown that *ATML1* is downregulated in *pPDF1-FLAG:PDF2* plants, consistent with PDF2 repressing *ATML1* (Chapter 3). However, it is possible that the roles of hetero and homodimeric combinations between ATML1 and PDF2 in the regulation of epidermal cell fate are subtly different, especially given the slight differences in binding site selection shown by Nakamura and colleagues (Nakamura et al. 2006). The mutant *pdf2* should only contain ATML1/ATML1 complexes and complexes between ATML1 and other HD ZIP IV proteins. In contrast, the overexpressor lines should additionally contain abnormally high numbers of PDF2/PDF2 complexes as well as PDF2/ATML1 combinations, and combinations of PDF2 with other HDZIP IVs, possibly at the expense of the formation of ATML1/ATML1 combinations. Thus specific targets of PDF2 containing complexes could be identified by transcriptome analysis of *pdf2* mutants, even if the phenotypes of these plants show no major defects. In theory, these targets should show converse regulation in *PDF2* overexpressing plants. Using the same logic, specific targets of ATML1/ATML1 complexes may also be identified in this analysis. For example, in theory, genes specifically activated by ATML1 homodimers might be upregulated in the *pdf2* mutant and downregulated in *pPDF1-FLAG:PDF2* plants. However common targets of ATML1 and PDF2 may well not be identified in our analysis.

With this in mind, a viable future approach to target identification might be the use of inducible lines either overexpressing or repressing the expression of both *ATML1*

and *PDF2*. Our results and the results of (Peterson et al. 2013; Takada et al. 2013) suggest that the best strategy for over expression studies would be to use an inducible epidermis-specific promoter for this type of study.

Overlap between response to stress and epidermal specification pathways

In this chapter I have used different approaches to study the network involved in epidermal cell fate and several connections with stress responses emerged. However it is unclear how stress response is involved in epidermal identity. An answer could be to invert the question and ask how epidermal identity is involved in stress response. The epidermis protects plants from abiotic and biotic stress. Its structure is optimized for defence. Three epidermis specific structures actively protect the plant from environmental stress: the cuticle, stomata and trichomes. Response mechanisms and epidermal cell fate specification share the same type of mechanism: signal perception, activation of signalling cascades and regulation. Both systems have a complex network of signalling cascades. The mechanisms are inevitably different but there may well be some overlap.

Several members of the HD-ZIP I and HD-ZIP II families (Figure 4.9) are involved in responses to abiotic stress (particularly light and dehydration) and response to phytohormones (ABA and auxin) to coordinate plant development (Harris et al. 2011). Interestingly, these transcription factors participate in homo or hetero dimerization within their families. They have different expression patterns and apparently different functions but they compete for the same binding site CAATNATTG, and like other HD-ZIP families they are highly redundant, implying a complex system of expression modulation (Ciarbelli et al. 2008; Harris et al. 2011; Henriksson et al. 2005). Recently it was shown that *HDG2* binds to a similar binding site as class I HD-ZIP *in vitro* (Peterson et al. 2013) and it is likely that other HD-ZIP IV proteins compete for similar binding sites. Our data also suggest links between ATML1/PDF2 and the class I and II HD-ZIP families. The binding site CAATNATTG was found to be overrepresented both in our promoter analysis of the *ATML1*/*PDF2* coexpression datasets but also, more significantly in *PDF2* transcriptomic data. It is thus formally possible that PDF2 (and by extension ATML1) bind directly to these sequences.

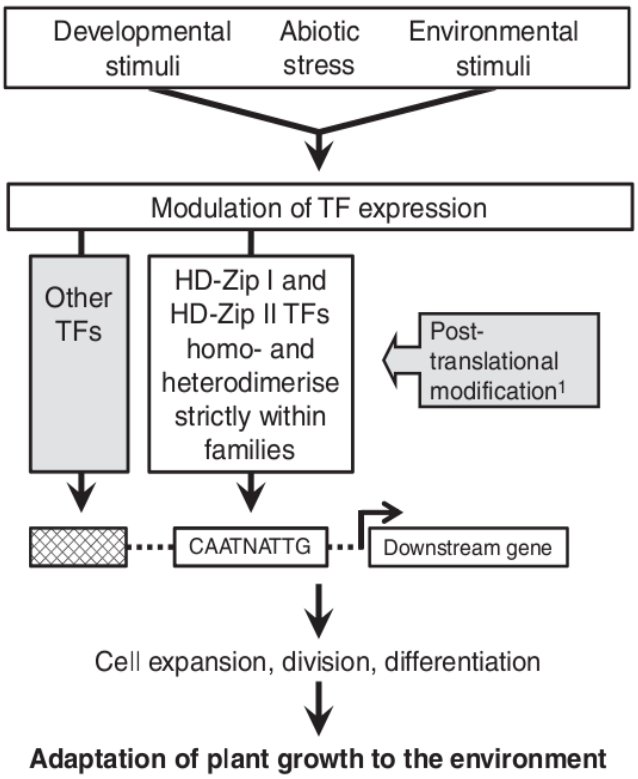


Figure 4.9 – HD-ZIP I and HD-ZIP II families are involved in a developmental and regulatory network. Developmental and environmental stimuli together with abiotic stress modulate TF expression. In cooperation with other TFs HD-ZIP I and HDZIP II would compete for CAATNATTG binding sites of promoters of genes that are involved in cell expansion, division or differentiation to promote adaptation of plant growth to the environment. Dashed lines represent a promoter. [1] Post-translational modification affects dimerization or DNA binding. (Adapted from (Harris et al. 2011))

My results suggest that some genes of HD-ZIP I (HB7, HB12, and HB21) and HD-ZIP II (HAT3, HAT14) families are potential targets of ATML1 and PDF2. HB7, HB14 and HB21 transcription is activated in the presence of ABA or water deficit treatments (Henriksson et al. 2005; Olsson et al. 2004). Ciarbelli and colleagues report that HAT3 is induced by low red/far red light and is probably under the control of the phytochrome system (Ciarbelli et al. 2008). More recently Turchi and colleagues have produced compelling results showing that *HAT3* is involved, together with *PHB, PHV* and *REV* in apical embryo development and patterning along the adaxial-abaxial axis of the embryo, and has an expression pattern overlapping with that of *ATML1* and *PDF2* in the embryo (Turchi et al. 2013). Together these results suggest that ATML1 and PDF2 could participate in the regulatory networks of other HD-ZIP transcription factors.

Within the HD-ZIP IV family, there are also reported links to stress responses. For example, *HDG11* has been shown to be involved in response to drought and salt stress (Cao et al. 2009; Yu et al. 2008). Several traits related to drought tolerance are enhanced by the activation of *HDG11* expression *in Arabidopsis* and ectopic overexpression of this gene in tall fescue results in improved tolerance to drought and salt stress (Cao et al. 2009; Yu et al. 2008).

Although dimerization of HD-ZIP family members has been described to be specific for each class of these transcription factors, it is possible that cooperation between proteins of different classes is mediated by the presence of associated binding sites within target gene promoters, as suggested by our data, or even by competition for the same binding sites. Harris and colleagues have proposed a network for HD-ZIP I and HD-ZIP II protein activity (Figure 4.9) in which these proteins control cell division, and growth in response to environmental signals (Harris et al. 2011). Interestingly, a recent paper from Adrienne Roeder has identified one role of *ATML1* as being to regulate cell cycle control (Roeder et al. 2012), presumably in response to developmental signals, again reinforcing similarities between the roles of these gene families.

Several genes coexpressed with *ATML1*/*PDF2*/*ACR4* or genes significantly up or downregulated in *pdf2* and *PDF2* overexpressor line are involved in hormone and stress response. Additionally the promoter analysis of these gene lists also supports an overlap between stress response and epidermal cell fate. Several binding sites associated mainly with light stress, water stress, ABA signalling (plant response to stress), and auxin response are found to be overrepresented in these groups of genes. The genes encoding the calcium binding proteins *CML38*, *PBP1* and *TCH3* (wounding response) are potential targets of PDF2 and are known to be activated by auxin. In addition PBP1 and TCH3 interact with PID (involved in auxin transport), linking stress, hormone and calcium responses (Benjamins et al. 2003; Benjamins et al. 2001).

Interestingly, there is strong enrichment for genes expressed during early responses to dehydration in the potential targets of PDF2 (13 out of 16 ERD genes). It could be argued that these genes are indirectly regulated due to defects in the epidermis of

pdf2 mutants, however most are downregulated in *pdf2* mutants, countering this argument. This result then strongly supports the idea that PDF2 could play an active role in stress responses by co-operating with other stress responsive transcription factors such as members of other HD-ZIP families.

Given the defensive role of the epidermis, one possible explanation for the observed link between HD-ZIP IVs and the expression of stress-responsive genes, is that this gene family could initially have diverged from other *HD-ZIPs* in order to render the outer cell layer, of early multicellular plants, special responsiveness to environmental stresses. This role might have become subsumed into a developmental role as the epidermal cell layer became more and more specialized during plant evolutionary history.

Other regulators involved in epidermal cell fate

Clearly ATML1 and PDF2 regulate key genes required for the specification of L1 cell fate. Interestingly however, our data suggest that proteins binding to WUS binding sites are also important for epidermal cell fate. From our promoter analysis the two sequences suggested to be the binding site of WOX proteins are overrepresented in promoters of both PDF2 potential targets and *ATML1*/*PDF2* coexpressed genes. Additionally analysis of functional annotations associated with combinations between L1box and WUS binding sites shows overrepresentation in GO terms such as transcriptional regulation, development, response and lipid metabolism. Taken together these results suggest that HD-ZIP IV and some WOXs cooperate in regulation of development and perhaps epidermal cell fate. From the WOX family the best candidates are *PRS*, *WOX2*, and *WOX9* since their expression domains overlap with *ATML1*/*PDF2* during embryogenesis. Of these genes, *WOX9* is potentially regulated by *PDF2* according to transcriptomic data.

Epidermal cell fate specification requires a network of signalling and regulation. To enlarge our knowledge of this network I decided to look for evidence of other regulators involved in epidermal specification. Given the importance of *ATML1* and *PDF2* in epidermal cell specification, I hypothesised that some of these regulators would be coexpressed with these genes or possibly even direct or indirect targets of

ATML1 and PDF2. From transcriptomic and coexpression results (including promoter analysis of these data) there are six transcription factor families that could be involved in specifying epidermal cell fate or in pathways used in epidermal specification/differentiation (Tables 4.5 and 4.6).

As previously discussed, the homeodomain family of transcription factors appears important for epidermal cell fate since family members and predicted binding sites appear in transcriptome analysis and co-expression analysis. The HD-ZIP, ZF-HD and WOX super families appear particularly important. The other five TF families potentially implicated epidermal cell fate are bZIP, WRKY, bHLH, MYB and NAC regulators, and interestingly all these families of transcription factors are related to abiotic stress response (reviewed in (Lindemose et al. 2013)). The role of bHLH transcription factors in epidermal cell fate has been recently reviewed by Zhao and colleagues (Zhao et al. 2012). In particular, five bHLH transcription factors, SCRM, SCRM2, MUTE, FAMA are involved in the regulation of stomata differentiation, and additionally SCRM and SCRM2 are also involved in cold response (Kanaoka et al. 2008; Lai et al. 2005; MacAlister et al. 2007; Ohashi-Ito and Bergmann 2006; Pillitteri et al. 2007). WRKY transcription factors have been described to be involved in ABA signalling, plant defence, trichome development and senescence (Eulgem et al. 2000; Tripathi et al. 2012). bZIP transcription factors are known to regulate pathogen defence, light and stress signalling, and in particular ABA responsive genes (Fujita et al. 2011; Jakoby et al. 2002). MYB transcription factors in *Arabidopsis* are involved in regulation of cell fate and cell identity, plant development, regulation of primary or secondary metabolism and abiotic/biotic stress response (reviewed in (Dubos et al. 2010)). The role of NAC transcription factors in the regulation of biotic and abiotic stress is reviewed in (Nakashima et al. 2012). Therefore, proposed roles for members of these protein families in epidermal specification do not seem unreasonable, and should provide directions for future research.

In summary, ATML1 and PDF2 appear to act together with other regulators in epidermal cell fate specification. The identity of these regulators is unknown, but our data suggests that several transcription factors involved in stress response, ABA response and development could be good candidates. A specific link between HD-

ZIP IV and WOX genes seems to be relevant for epidermal cell fate, supported by promoter analysis of *ATML1*, genes coexpressed with HD-ZIP IV encoding genes, and transcriptome analysis of the *pdf2* mutant and *PDF2* overexpressing plants.

Chapter 5 – A feedback loop model for epidermal cell fate

5.1 Introduction

5.2 Results

5.2.1 In seedlings, ACR4 mediated signalling positively regulates the expression of ATML1 and PDF2, which in turn repress the expression of ACR4

5.2.2 ODE model of the feedback loop core

5.2.3 Transcriptional regulation of ACR4

5.2.3.1 Epidermal pattern of ACR4 promoter is achieved by complex regulation

5.2.3.2 ACR4 promoter expression pattern in the SAM

5.2.3.3 A WUS binding site containing region and ARE-containing regions regulate expression of ACR4 in the root.

5.2.4 Genetic analysis of components of the epidermal feedback loop.

5.2.4.1 ACR4, ATML1 and PDF2 act in the same genetic pathway during embryo development.

5.2.5 ACR4 and WOX genes interact genetically during embryogenesis and after germination

5.3 Discussion

5.1 Introduction

Epidermal cell fate is established during early stages of plant embryogenesis. A number of genetic backgrounds compromised in epidermal cell fate specification are embryo lethal, suggesting that the specification of the L1 layer is critical for plant development. The aim of this section is to define a model for the establishment and maintenance of the epidermal cell fate.

The transcription factors ATML1/PDF2 are involved in the regulation of epidermal cell fate, presumably by regulating the expression of L1-box containing genes (Abe et al. 2003). The loss of these two *HD-ZIP IV* genes impedes epidermis specification in early stages of embryogenesis leading to embryo lethality, suggesting that they play a central role in epidermal cell fate. Other HD-ZIP IV transcription factors are predicted to be involved in the establishment and maintenance of epidermal cell fate either in cooperation with ATML1/PDF2 or through independent pathways. The analysis of *ATML1* promoter expression in different stages of embryogenesis shows that the expression of this gene is spatially and temporally dynamic, and additionally that the presence of both an L1box and a predicted WUS binding site is necessary to ensure the correct epidermal expression pattern of this promoter (Takada and Jurgens 2007).

The presence of an L1 box in the promoters of both *ATML1* and *PDF2*, and the fact that elimination of this element reduces expression of the *ATML1* promoter during embryogenesis, has led to the hypothesis that these genes undergo positive autoregulation (Abe et al. 2003; Abe et al. 2001; Takada and Jurgens 2007). A similar argument was used to suggest that *PDF1* was a likely target of ATML1 (Abe et al. 2003; Abe et al. 2001). However, results in the previous chapters have shown that, at least in seedlings, ATML1 and PDF2 repress their own expression. Nonetheless, we cannot exclude the possibility that a different regulatory relationship exists during early embryogenesis. As discussed previously, the fact that I was able to complement the *atml1-3pdf2-3* double mutant with *ATML1* or *PDF2* driven by the *PDF1* promoter strongly suggests that other HD ZIP IVs (or other transcription factors) positively regulate this promoter during early embryogenesis. Thus the

situation appears much more complex than has previously been assumed in the literature.

The findings of Takada and Jurgens suggest that in addition to HD ZIP IV proteins, WOX proteins could be implicated in the specification of epidermal cell fate in embryogenesis (Takada and Jurgens 2007). The analysis of *ATML1* promoter in the embryo indicates that a sequence of 101 bp, containing L1 box and WUS binding site is sufficient to mimic the expression of the *ATML1* full length promoter. Mutation or deletion of either the L1 box or the WUS binding site affects the GFP expression levels and patterns of this "minimal" promoter, but neither modification totally abolishes L1 expression in embryos (Takada and Jurgens 2007). These results suggest that the WUS binding site is important for the L1 expression pattern of *ATML1*, but that other transcription factors, including HD-ZIP IV proteins, are involved in the regulation of *ATML1*, especially at early stages of development. *WUS* is expressed in the embryo from the dermatogen stage onwards in the subepidermal SAM precursor cells (Mayer et al. 1998). Recently, however, it has been proposed that WUS can move between cells in inflorescences (Yadav et al. 2011) and it is possible that the same mechanism takes place in embryogenesis, allowing WUS to regulate *ATML1*. Additionally other proteins belonging to the WUS-related family (WOX) of homeodomain transcription factors could bind to similar binding sites to WUS. The expression patterns of *WOX2* and *WOX9* (Haecker et al. 2004) overlap in some regions with the epidermis, and therefore could act upstream of *ATML1* or in cooperation with *ATML1*.

The receptor kinase ACR4 is expressed specifically in the epidermal layer in the embryo and apical meristem (Gifford et al. 2003). *acr4* mutants display some defects in cell organization of sepal margins and ovule integuments. In addition, other phenotypes are found in seed development, namely aborted ovules, epidermal outgrowths and round misshapen seeds (Gifford et al. 2003; Watanabe et al. 2004). In particular, the seedlings and seedcoats of *acr4* mutants were more permeable than wild-type to the hydrophobic dye toluidine blue, which reflects defects in the epidermis (Watanabe et al. 2004). ACR4 is predicted to act together with the receptor

kinase ALE2 since these proteins reciprocally phosphorylate each other and the *acr4* mutant is reported to be epistatic to *ale2* (Tanaka et al. 2007).

In addition to being implicated in the development of the epidermal cell layer of the shoot, ACR4 also plays an important role in root patterning. ACR4 has been shown to control cell division during the initiation of lateral root growth and cell proliferation in the root columella region, together with the receptor kinase CLV1. In this context ACR4 is thought to perceive the peptide CLE40 (De Smet et al. 2008; Stahl et al. 2013; Stahl et al. 2009). The *ACR4* promoter (~1.9 kb) drives expression in the main root tip and lateral root tip in similar patterns. Expression in lateral roots is initiated at very early stages, just after the two asymmetric divisions which give rise to lateral root precursors in the pericycle (De Smet et al. 2008; Gifford et al. 2003). In the main root tip the *ACR4* promoter drives expression in the central cells of the quiescent centre (QC), columella initial cell layers (close to the QC), the lateral root cap (LRC) and in epidermal cell files as they emerged from the LRC (De Smet et al. 2008; Gifford et al. 2003; Stahl et al. 2009).

Previously published results suggest that *ACR4* may be a target of ATML1/PDF2, since *ACR4* expression was shown to be lost in *atml1-1pdf2-1* double mutant seedlings (Abe et al. 2003). However, the lack of an epidermal layer of the *atml1-1pdf2-1* seedlings makes this data very hard to interpret, since *ACR4* is specifically expressed in the epidermis. The relationship between *ACR4* and *ATML1/PDF2* therefore needs to be clarified, since the published evidence is insufficient either to prove that *ACR4* is a direct target of these transcription factors, or to confirm how they might regulate its expression. My ChIP experiments prove that ATML1 and PDF2 bind *in vivo* to the *ACR4* promoter, which contains an L1 box. However it is still unclear whether they can repress or activate *ACR4*. Recently *ATML1* and *ACR4* have been implied in the control of giant cell identity in sepals. Single mutants of either gene display a reduced number of giant cells compared to the wild-type (Roeder et al. 2012). The similarity of acr4 and atml1 phenotypes point to ATML1 and ACR4 acting in the same pathway, however the mechanism is unclear. Additional evidence from previous member of the Ingram group, from microarray data in inflorescences, indicates that ACR4 could positively regulate *ATML1/PDF2*

expression (Gwyneth Ingram personal communication). In the context of the shoot epidermis, the existing published data, together with this observation, can be summarized in the following scheme (Figure 5.1).

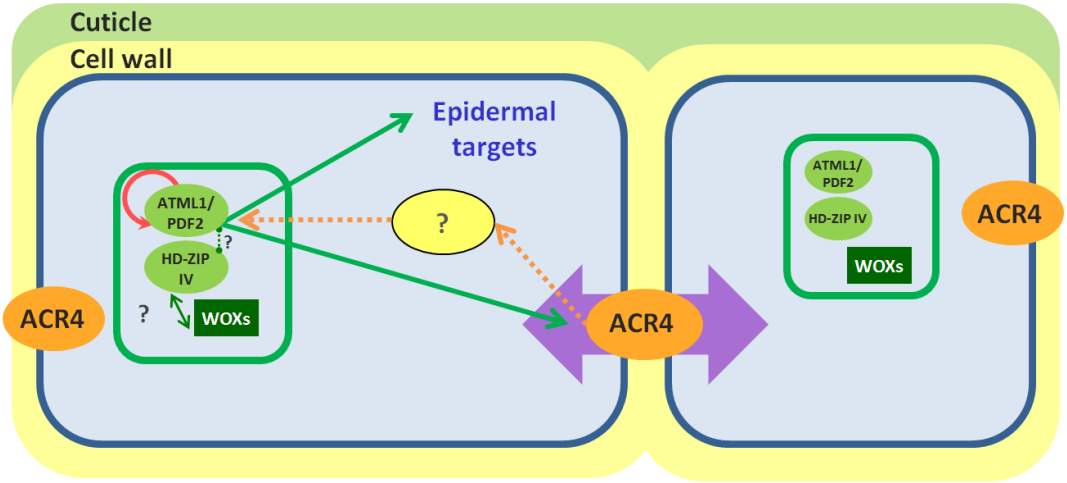


Figure 5.1 – Current model for epidermal cell fate. The transcription factors ATML1/PDF2 (green circle) autoregulate themselves (red arrow) and are responsible for the regulation of genes involved in epidermal cell fate and of the receptor kinase encoding *ACR4* gene (green arrows). ACR4 (orange circle) is involved in positional signalling (purple double arrow) necessary for the maintenance of epidermal cell fate identity. The ACR4 signalling pathway activates expression of the ATML1/PDF2 transcription factors (orange dashed arrows and yellow circle). However this mechanism is unknown. WOX genes (dark green) could be involved in the regulation of *ATML1* and potentially other HD-ZIP IV genes in embryogenesis (dark blue arrow). Other HD-ZIP IV protein complexes could act in cooperation with ATML1/PDF2 to promote epidermal cell fate (dashed green link).

In this chapter I aimed to test aspects of this model, including the proposed interactions between its various elements. I used a combination of genetic analysis and mathematical modelling to explore various regulatory scenarios, and propose a simple feedback loop model as a backbone for the specification and maintenance of epidermal cell identity in *Arabidopsis.*

5.2 Results

5.2.1 In seedlings, ACR4 mediated signalling positively regulates the expression of *ATML1* and *PDF2*, which in turn repress the expression of *ACR4*

The interaction between the receptor kinase encoding *ACR4* gene and the HD-ZIP IV transcription factors is complex and controversial. In order to clarify the situation I

measured the transcript levels of *ATML1*, *PDF2*, and *ACR4* in seedlings of mutants and transgenic lines. The levels of *ATML1* and *PDF2* were consistently and significantly downregulated in *acr4-2* mutant seedlings compared to wild-type plants (Figure 5.2A). This observation was also made in inflorescence tips (data not shown) and confirms previously obtained micro-array data from the laboratory. Thus I was able to validate the hypothesis that the expression of *PDF2* and *ATML1* is positively regulated by ACR4-mediated signalling.

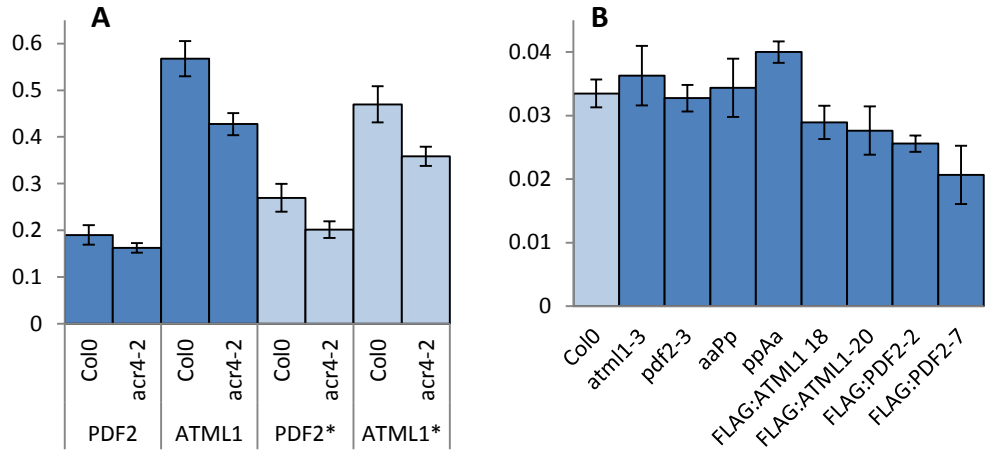


Figure 5.2 – Expression analysis of *ACR4*/*PDF2*/*ATML1* epidermal genes in seedlings. (A) Transcription levels of *ATML1* and *PDF2* in *acr4-2* and wild-type backgrounds using two pairs of primers for each transcript. (B) Transcription levels of *ACR4* in different mutant backgrounds, namely *atml1-3*, *pdf2-3* and *atml1-3 pdf2-3 (aaPp* and *ppAa*) mutants and lines expressing FLAG-PDF2 and FLAG-ATML1 fusion proteins under the *PDF1* promoter. Q-PCR is normalized to *EIF2* expression.

However, contrary to data published in the literature, I found that the expression of *ACR4* in different genetic backgrounds is consistent with repression of *ACR4* by the transcription factors ATML1/PDF2, at least in seedlings (Figure 5.2B) and inflorescence tips (data not shown). Thus, the expression of *ACR4* is significantly lower in lines overexpressing *ATML1* or *PDF2* and significantly higher in the segregating seedlings of plants which were heterozygous for *atml1-3* and homozygous for *pdf2-3*. However, in the single mutants of *atml1-3* and *pdf2-3* the expression level of *ACR4* is indistinguishable from that in wild-type seedlings (Figure 5.2B). Therefore three of the four copies of *ATML1/PDF2* need to be absent before *ACR4* expression is significantly affected. This is consistent with the global functional redundancy between *ATML1* and *PDF2*.

5.2.2 ODE model of the feedback loop core

The literature supports a positive regulation of *ACR4* and *ATML1/PDF2* by ATML1/PDF2 via the L1box, the strongest argument being the loss of *ACR4*, *ATML1* and *PDF2* expression in *atml1-1pdf2-1* seedlings (Abe et al. 2003; Tanaka et al. 2007). In contrast, my results (Chapter 3) strongly suggest that ATML1 and PDF2 actually repress *ACR4*, *ATML1* and *PDF2* expression in seedlings, a result which I believe is more strongly supported genetically than other published assertions. This disparity is counterintuitive. However we are not in a position to exclude the possibility that both situations exist, but at different times during development. The fact that I have confirmed that, as previously proposed, ACR4 can activate the expression of *ATML1/PDF2* via an unknown signalling pathway, permits me to propose two feedback loops, one positive, and one negative. The properties of these two feedback loops were tested in collaboration with Etienne Farcot (Virtual Plants laboratory in Montpellier, presently at the University of Nottingham).

Briefly, two models were tested to describe the different interactions between *ATML1/PDF2* and *ACR4* (Figure 5.3). Both models considered that ACR4 (*A*) is the receptor of an undetermined ligand L (in the absence of knowledge regarding this ligand it is represented by a time dependent variable $\lambda(t)$) responsible for the activation of a signalling cascade that activates expression of *ATML1* and *PDF2* which, for the purposes of these models are considered to be fully functionally redundant and are consequently defined by the same kinetic equations and described by a unique variable (*P*). In both models the transcription factors ATML1 and PDF2 directly regulate expression of *ACR4*. The difference between the two models is the regulatory role of the ATML1 and PDF2 proteins. The simplest model M1 considers that ATML1/PDF2 activate *ACR4* gene expression whereas the M2 model considers that ATML1/PDF2 repress the expression of *ACR4* and in addition repress their own gene expression (Figure 5.3). Additionally the second model introduces the homo or heterodimerization of ATML1/PDF2 described by the association constant (γ). The dynamics of ACR4 signalling and ATML1/PDF2 transcriptional regulation are simplified and follow the Hill equation (Hill 1913). The models are described by a system of two ODEs (Ordinary Differential Equations) representing *ACR4*

expression and *ATML1/PDF2* expression. In the M1 model the transcriptional activation by ATML1/PDF2 corresponds to a Hill coefficient of 1 whereas in the M2 model the transcription repression by ATML1/PDF2 dimers corresponds to a Hill coefficient of -2 (Figure 5.3). In both models ACR4 and ATML1/PDF2 are subject to a linear degradation.

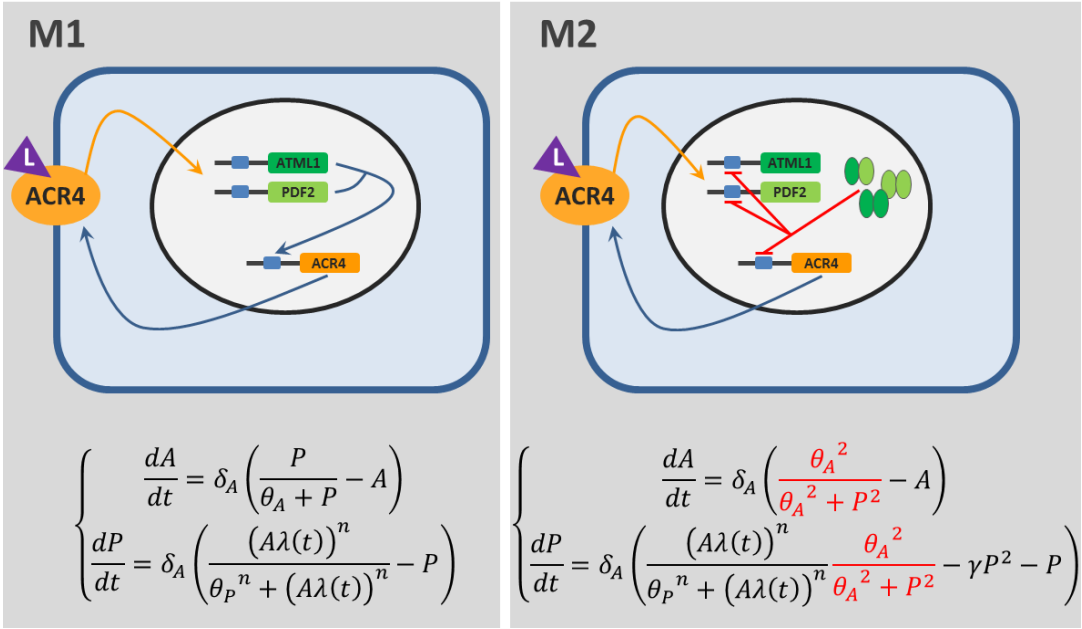


Figure 5.3 – ODE models for the ACR4, ATML1 and PDF2 feedback loop. (M1) Transcriptional activation of *ACR4* (orange, *A*) mediated by ATML1/PDF2 (green, *P*) transcription factors via the L1 box (blue square). (M2) ATML1/PDF2 transcription factors repress the transcription of *ACR4*, *ATML1* and PDF2 (red). In both models, the binding of a ligand (purple) to the receptor kinase ACR4 activates a signalling cascade (orange arrow) that activates *ATML1/PDF2*. The kinetics are defined by the Hill function, where θ_A corresponds to *ACR4* transcription regulation by ATML1/PDF2, θ_P corresponds to the activation of these two genes by ACR4 signalling and ligand binding to ACR4 ($\lambda(t)$) is time dependent. The degradation of ACR4 and the transcription factors is represented by a linear decay (δ_A and δ_P). In the model M2, the association constant γ defines the dynamics of the homo and heterodimerization of ATML1/PDF2.

The lack of kinetics knowledge of the different interactions between ACR4 and ATML1/PDF2 is an obstacle for the quantitative determination of the coefficients described in M1 and M2. Therefore, numerical simulations were made to examine the qualitative behaviour of these two normalized systems. Both models, show asymptotic dynamics in a system with a fixed high input value of ligand over time which results in positive values of *A* and *P* variables (Figure 5.4).

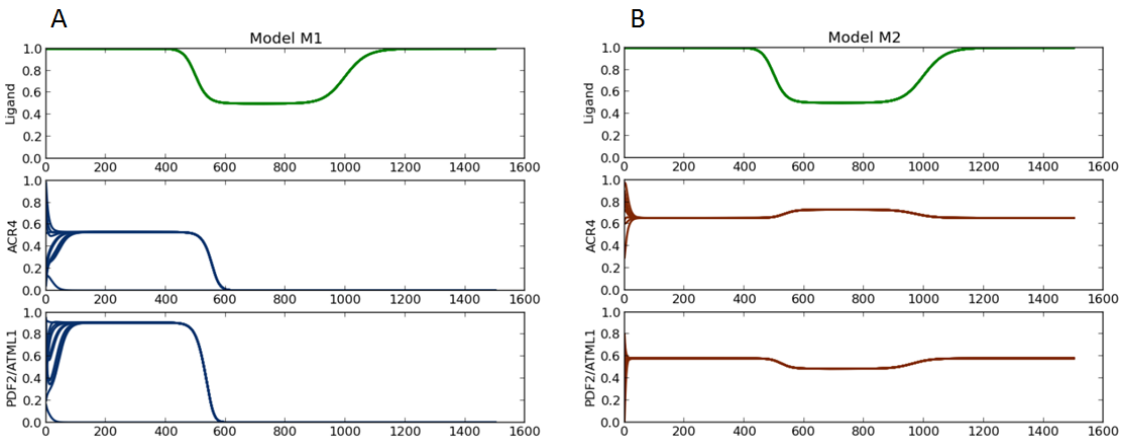


Figure 5.4 – Signalling perturbation on ACR4 feedback loop models. M1 and M2 systems correspond to (A) and (B) respectively. The transient decrease of ligand concentration between instant 500 to 1000 (top graphic, green), results in different profiles of ACR4 (middle graphic) and ATML1/PDF2 (bottom graphic) over time for 10 different initial conditions. While in M1 both ACR4 and ATML1/PDF2 variables decrease and reach (0,0) steady state, the M2 model is more robust and recovers from the system perturbation. θ_A =0.3, θ_P =0.8, n=4, δ_A=δ_P=1 and γ=0.5

The model M1 corresponds to a positive feedback loop: the increase of ACR4 signalling drives production of ATML1/PDF2. Consequently ACR4 signalling increases until this loop reaches a steady state (dependent on degradation of the components). However, we found that perturbations in ligand concentration frequently induced an irreversible switch of the system. Thus a decrease of the ligand concentration induces the steady state A=P=0 (Figure 5.4A). Subsequent increases in the ligand concentration do not lead to a recovery in either A or P. This behaviour of M1, leading to two steady states, was shown for several simulations characterized by a strong induction of ATML1/PDF2 production by the ACR4 signalling pathway, and thus corresponding to a value of the Hill constant (n) which is reasonably high, without being biologically improbable. Thus, the M1 model showed a hysteretic behaviour typical of bistable systems.

The model M2 converged to a single steady state where both variables A and P are positive. Unlike M1 model, the M2 system is more robust with one steady state that can diverge reasonably with changes of the ligand concentration but recovers to the initial steady state when the perturbation is reversed (Figure 5.4B).

Introducing the dimerization of ATML1/PDF2 into model M1 (steepness parameter *n*=2), or removing it from model M2 (steepness parameter *n*=1) had no impact in the global behaviour of the respective system.

The epidermis specification pathway is required at early stages of embryogenesis to define the identity of the L1 layer. Additionally, this pathway is present in shoot apical meristem and organ primordia, where it is presumably required to maintain epidermis identity. The fact that M2 evolves towards a unique and robust steady state is more consistent with a solution to the biological problem of maintaining epidermal identity than a model that generates bistability. Consistent with this, our transcriptional analysis of the behaviour of the system in seedlings (where epidermal identity is maintained but not established *de novo*) are consistent with the predictions of M2 but not M1.

5.2.3 Transcriptional regulation of *ACR4*

In order to explore the regulation of *ACR4* further, and in particular to ascertain whether regulatory elements other than the L1 box might be important for the regulation of this gene, I carried out a limited analysis of the *ACR4* promoter. Because of time constraints, and on the basis of ongoing collaborations in the laboratory, I tested different versions of ACR4 promoter (Figure 5.5) which were designed to focus on the possible roles of three types of regulatory regions: the L1 box, WUS boxes and AREs (Auxin response elements).

The *ACR4* promoter region which I analysed had previously been shown to be sufficient to provide the full expression pattern of ACR4 (Gifford et al. 2003).

This region is 1068 bp in length, starts in the *ACR4* 5' UTR (untranslated region) and ends after 3' UTR of the upstream gene (AT3G59510). I produced eight pACR4-GFP:GUS constructs using the plasmid pKGWFS7 (Karimi et al. 2002) which allows expression pattern determination using GUS staining or GFP imaging. The constructs P1-P6 are 5' deletion versions of the *ACR4* promoter (Figure 5.5B) and are designed in such a way that three out of four AREs (ARE 4 being in the 5'UTR region of *ACR4*), and 2 out of three WUS boxes (WUS1* overlaps with the L1 box) are

sequentially removed from the 5' end of the promoter (Figure 5.5B). Elements are as previously described in chapter 4 (Table 4.1).

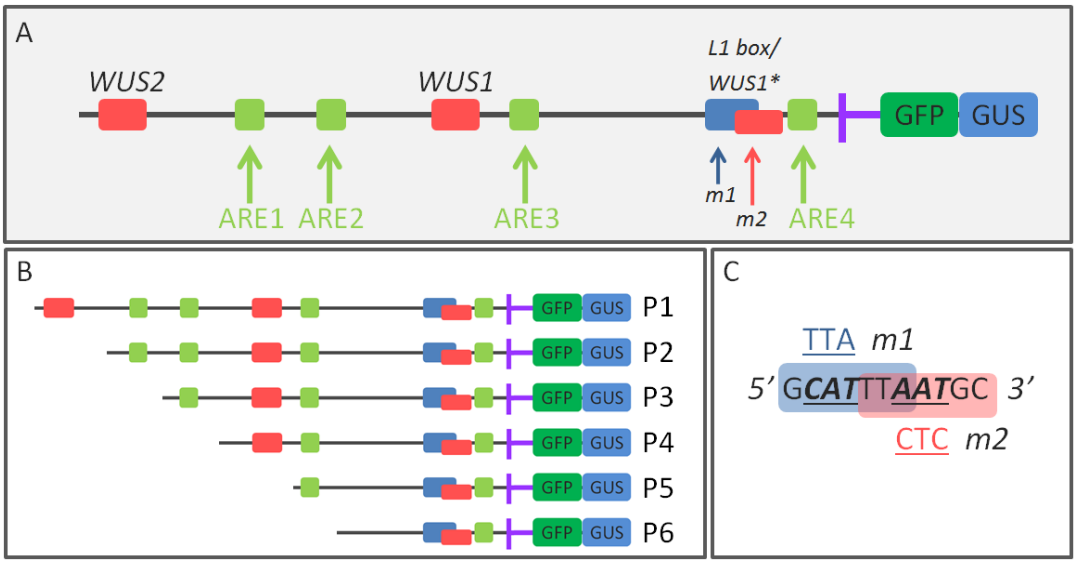


Figure 5.5 – *ACR4* promoter construct strategy. The *ACR4* promoter drives the expression of GFP and GUS (A, B). This promoter contains eight potential regulatory elements of relevance, namely 4 AREs (green), three WUS binding sites (two WUS1, and one WUS2 represented in red) and one L1 box (blue). Eight versions of *ACR4* promoter were made consisting of six deletion constructs (P1-P6) (B) and two mutant versions of the full length promoter (C) disrupting either the L1 box alone (m1) or both the WUS1* and the L1 box (m2).

Additionally, two mutations were made in the L1 box region (Figure 5.5C) in the full length promoter P1. The first mutation (m1) is localized in the L1 box and the second mutation affects both the end of the L1 box and the interior of the WUS1*. In reality I made two versions of each promoter construct, one group ending at the start of the *ACR4* 5' UTR and the second group including a small portion of the 5' UTR. I found that in each case these two versions have similar expression patterns, and as a result, transformants were generally considered as one group. The transformation of M1 and M2 constructs containing the 5' UTR fragment was unsuccessful due to a problem with culture conditions, and was not repeated since results were expected to be similar to those obtained with the slightly shorter versions. GUS staining experiments using seedlings of several (6-10) independent lines of each of the 14 transformed constructs allowed the selection of representative lines that were used to examine the GFP signal in embryos and inflorescences.

5.2.3.1 Epidermal pattern of *ACR4* promoter is achieved by complex regulation

The full length *ACR4* promoter expression in embryogenesis is restricted to the epidermis (P1 in Figure 5.6) consistent with the *ACR4* expression pattern described in the literature (Gifford et al. 2003; Tanaka et al. 2002). The GUS signal of *pACR4-GFP:GUS* expressing lines was difficult to analyse during embryogenesis and the proportion of GUS stained embryos varied between siliques, probably due to problems of penetration of the staining reagents. Nevertheless the GUS staining results were globally consistent with the GFP signal analysis (Figure A1 in appendix).

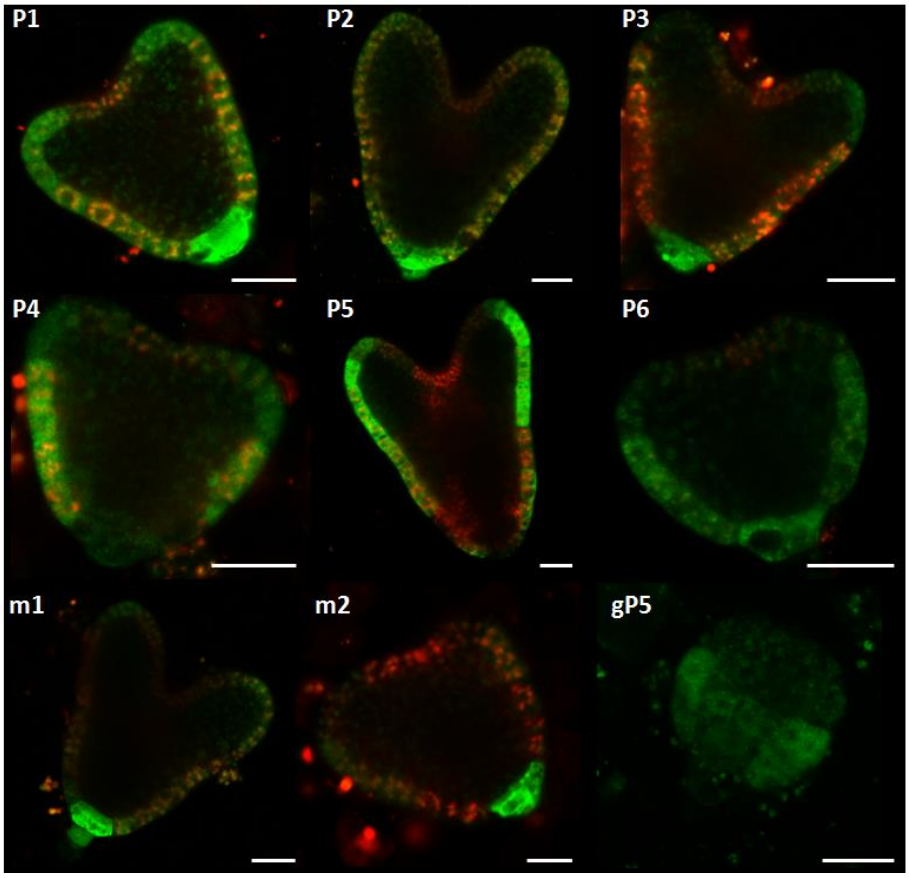


Figure 5.6 – Expression pattern of *ACR4* promoter constructs in embryogenesis. Expression pattern of P1-P6, m1, and m2 constructs (shown in Figure 5.5) in the heart stage of embryogenesis. gP5 corresponds to the 3D reconstruction of a globular embryo expressing P5 construct. Scale bar is 20 µm.

The expression of the full length promoter, P1, was difficult to detect at early stages of embryogenesis (consistent with the results of *in situ* hybridizations (Gifford et al. 2003), GFP expression could be observed from the 4-8 cell embryo onwards). I

therefore concentrated my analysis at the globular stage and heart stage, which were examined for representative lines of each *ACR4* promoter construct (Figure 5.6 and Figure A2 in the appendix).

In globular and heart stage embryos the signal of P1 is restricted to the L1 layer. The expression of P2 (which has lost a WUS2 binding site) in globular stage embryos is globally comparable to P1 and slightly stronger in the basal pole. Later, at the heart stage the expression of P2 in the L1 is weaker than in P1 suggesting that regulation via WUS2 could be positive with respect to epidermal expression. In globular stage embryos the expression of P3 is difficult to detect. In early heart stage embryos the expression of P3 can be detected in the basal pole but it is significantly weaker in the L1 compared to P1 and P2. Interestingly, P4 and P5 lines totally lacked expression in the root pole, and also lacked expression in apical regions of the embryos at both the globular and heart stages. These lines drove expression in a band of L1 cells between the apical and basal poles (P4, P5 and gP5 in Figure 5.6). P6 showed the same expression profile as P4/P5 with exception of some embryos which at the heart stage also showed weak expression in the basal pole.

The expression pattern of P6 (around 0.3 kb) suggests that the presence of the L1 box and the overlapping WUS1* is not sufficient to drive the whole epidermal expression pattern of *ACR4* during embryogenesis. However the expression profile of m1 and m2 in which most of the L1 expression of the *ACR4* promoter is eliminated from the embryo imply that an intact L1 box and WUS1* are necessary for the L1 expression pattern.

Finally the deletion analysis suggests that *ACR4* expression in the basal pole is strongly dependent upon the presence of ARE containing regions rather than the identified WOX binding site or L1 containing zones since neither the lack of WUS2 (P2) or mutation of L1box/WUS1* (m2) has any major effect on the basal pole expression.

5.2.3.2 *ACR4* promoter expression pattern in the SAM

The characterization of the *ACR4* promoter in the shoot apical meristem was based on two independent representative lines for each construct. Again the expression

patterns shown by the promoters regions upstream of the 5'UTR (8 constructs) was similar to that of the promoters containing 5' UTR regions (6 constructs).

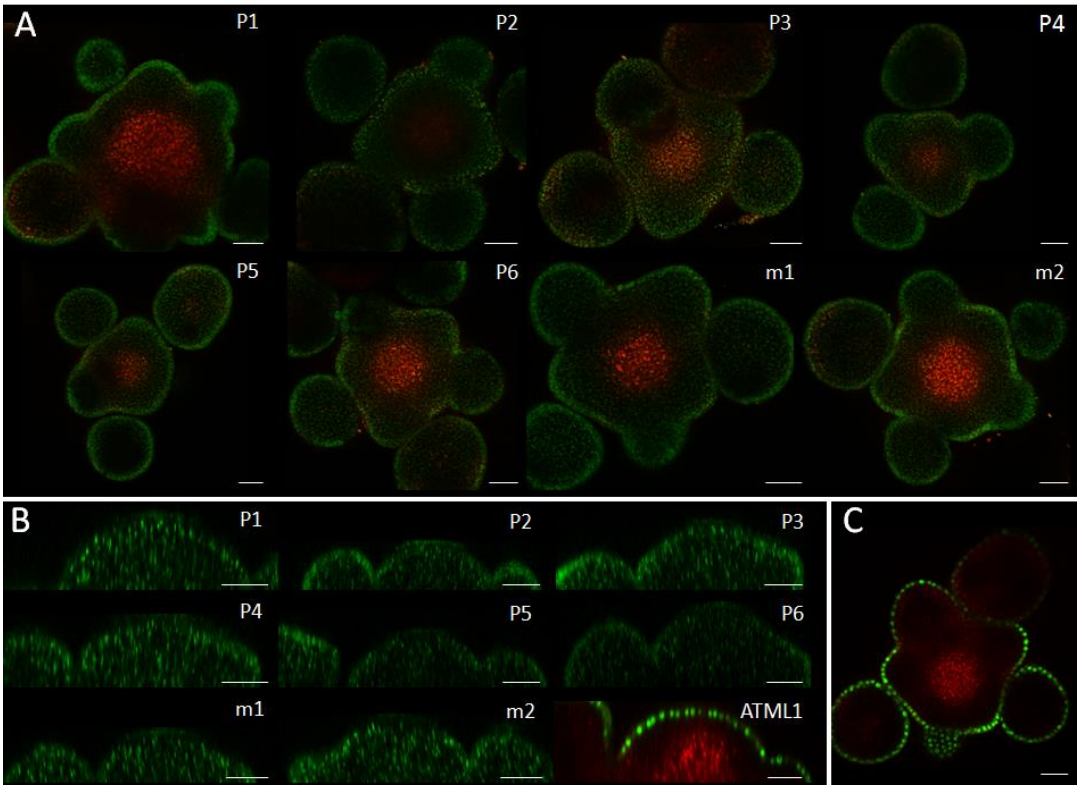


Figure 5.7 – *ACR4* promoter expression in the SAM. A) Horizontal sections of meristems expressing P1-P6, m1 and m2 constructs; B) Side view of the meristems expressing P1-P6, m1, m2, and pPDF1-GFP:ATML1 constructs. C) Expression of pPDF1-GFP:ATML1 in the SAM. Scale bar is 25 µm

The GFP signal from the reporter vector used is unfortunately rather weak, a fact of which I only became aware relatively late in the project. In addition, according to the eFP Browser, and my Q-PCR results, the expression levels of ACR4 (275) are around ten times lower than, for example, those of *PDF1, PDF2 or ATML1* in the apical meristem (Vandepoele et al. 2009). A high laser intensity therefore had to be used during confocal microscopy on meristems. In order to be able to compare results between lines I always carried out GFP signal imaging in the same experimental conditions. The analysis includes four independent lines each of P1-P5 constructs, three independent lines of P6 and two lines each of m1 or m2. I found that at least one line of P1-P6, m1 or m2 showed some L1 specific expression (Figure 5.7). This observation suggests that the small fragment P6 is sufficient to

give L1 expression in the SAM and that the mutation of the L1 box and WUS box1' is not sufficient to eliminate all L1 specific expression in the meristem.

For most versions of *ACR4* promoter the GFP expression level in the SAM varied between independent lines, with the exception of P4, P5 and m2 which showed the same expression level in all lines. For instance, the m1 mutation of the L1 box affected the gene expression of one line but not the other. Moreover, in general the expression driven by the P3 constructs is weaker than that of other constructs. The stronger lines of each construct were studied in more detail. Interestingly, the L1 expression pattern of the m2 construction is weaker at the top of the apical meristem increasing towards the organ boundaries and stronger in the buds. This trend can also be seen, although less strongly, in P2, P4, and P5. These fluctuations may indicate that as in embryos, uniform expression of *ACR4* in the meristematic L1 depends on the interplay between a range of different factors, and is, in central meristematic zones, at least partially dependent upon the L1 box/ WUS1* sequence.

5.2.3.3 A WUS binding site containing region and ARE-containing regions regulate expression of *ACR4* in the root.

The root expression pattern of the full length *ACR4* promoter P1 (Figure 5.7A) was consistent with its description in the literature (De Smet et al. 2008; Gifford et al. 2003). The expression of P1-P6, m1 and m2 constructs in seedlings was investigated by GUS staining (Figure 5.8B).

Using GUS staining, all constructs (P1-P6, m1, and m2) are expressed in the shoot apical meristem of seedlings including leaf primordia, and in the tips of the lateral roots of seedlings (Figure 5.7A). Epidermal expression in the primary root was observed only in the seedlings expressing the full length promoter, P1 (Figure 5.7). This suggests that the WUS2 containing region could be necessary for this expression pattern. Interestingly, the expression of m1 and m2 in this region was also lost, suggesting regulator binding to the L1 box may also be necessary for expression in this domain. The expression pattern in the primary root tip is undetectable in most P3 lines (Figure 5.7) and a similar trend was observed in seedlings containing the P4-P6 constructs. Interestingly the expression level of P3 at early stages of lateral root development was significantly decreased and restricted to the tip of the lateral root

primordia where the auxin peak is localized (Tanaka et al. 2006). The expression of P4-P6 promoters was only detected at later stages of lateral root development.

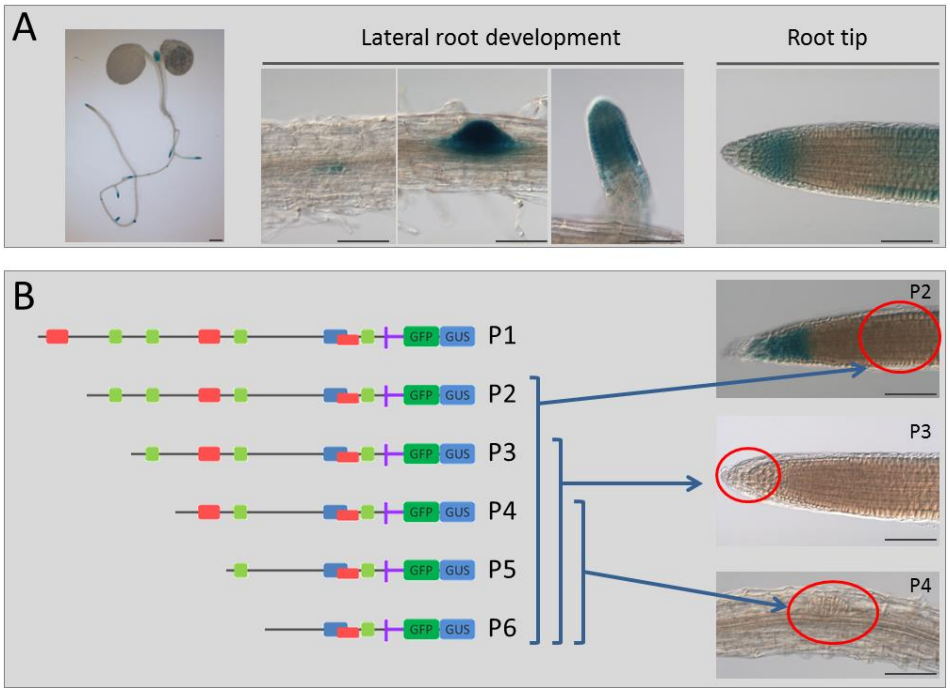


Figure 5.7 – GUS expression patterns of *ACR4* promoters in the root. A) The expression of full length promoter P1 is initiated at early stages of lateral root development and is restricted to the lateral root tip. Additionally, P1 is expressed in root epidermal cell files and main root tip. B) The epidermal expression pattern in the primary root is lost in P2-P6 constructs, the primary root tip expression is absent in the P3-P6 constructs and the expression of P4-P5 in the lateral root only starts at later stages of lateral root development. Scale bar is 50 µm with exception of the top left picture where the scale bar is 500 µm.

5.2.4 Genetic analysis of components of the epidermal feedback loop.

To confirm the genetic relationships implied between the various players of the epidermal feedback loop model I analysed the double mutant combinations between mutant lines of *ACR4* and *ATML1* or *PDF2*. Because of the possible redundancy of *ATML1* and *PDF2* with other HD-ZIP IV encoding genes I also tested mutant combinations between mutant alleles of most of these genes and *acr4-2* mutant.

The genetic interactions between genes encoding HD-ZIP IV transcription factors and the *ACR4* were determined by crossing each mutant of the *HD-ZIP IV* family (Table 3.2) with the *acr4-2* mutant. Double mutants between *acr4-2* and *hdg1-3* were never identified; which is likely to be due to lack of genetic recombination since

HDG1 (AT3G61150) and *ACR4* (AT3G59420) are localized in the same region of the genome. Crosses between *acr4-2* and *hdg11-2* mutants were difficult to obtain in both directions. The F2 generation of the *acr4-2 hdg11-2* cross was obtained with difficulty and could not be examined further due to time constraints. Finally *hdg12-2acr4-2* and *anl2-2acr4-2* double mutants had not been obtained at the time of writing, and segregation ratios in F2 plants suggest that there may be problems of either transmission or lethality associated with these mutant combinations. Again I was unable to follow this further. Most of the double mutants obtained show no exacerbation of the acr4-2 phenotype, with the exception of *acr4-2pdf2-3* and *acr4-2atml1-3* which showed developmental problems described below (Figure 5.8-5.12).

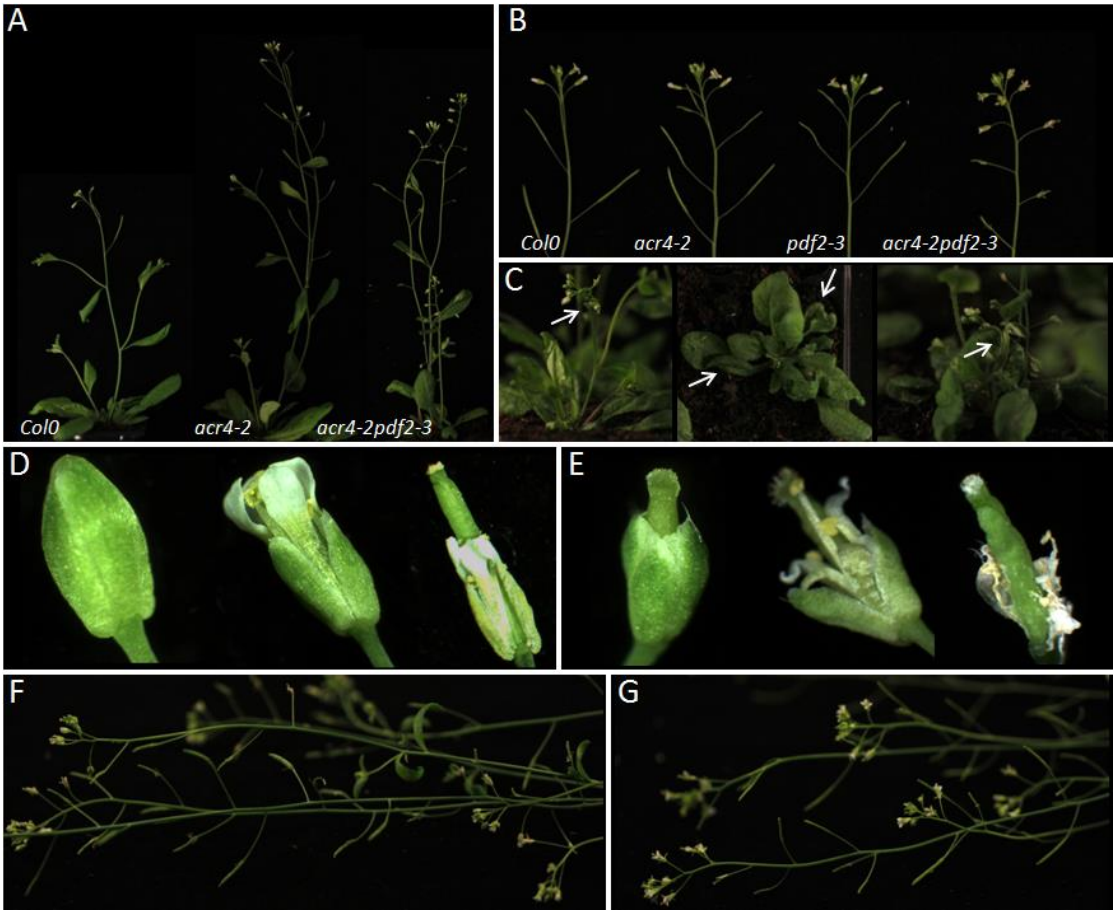


Figure 5.8 – Characterization of the *acr4-2pdf2-3* double mutant. Comparison of plants (A) and inflorescences (B) of *acr4-2pdf2-3* double mutant with respective single mutants and wild-type; (C) the double mutant can display misshapen leaves and stem fusions (white arrows). Furthermore, compared to the wild-type (D) *acr4-2pdf2-3* flower organs are misshapen and can be fused together (E). The *acr4-2pdf2-3* double mutant plants expressing of FLAG-PDF2 (G) complement the flower phenotypes and sterility of *acr4-2pdf2-3* plants (F).

Both *acr4-2pdf2-3* and *acr4-2atml1-3* double mutants show fertility problems which are more severe than those seen in the *acr4-2* single mutant. However both double mutants are fully viable and can be obtained with ease from parental plants segregating one or both mutant alleles. In general the phenotypes observed of *acr4-2 atml1-3* double mutants were weaker than those of *acr4-2pdf2-3* (Figures 5.8 to 5.12). For instance *acr4-2pdf2-3* is completely sterile due to severe ovule fusion and malformations whereas *acr4-2atml1-3* is partially fertile producing a limited number of viable seeds (Figure 5.8, 5.9 M, O and 5.10). To address the fertility problem I crossed *acr4-2pdf2-3* plants with wild-type plants using *acr4-2pdf2-3* as female or male. Only the crosses using *acr4-2pdf2-3* pollen produced seed, proving that the main contributor to *acr4-2pdf2-3* sterility is the ovule phenotype.

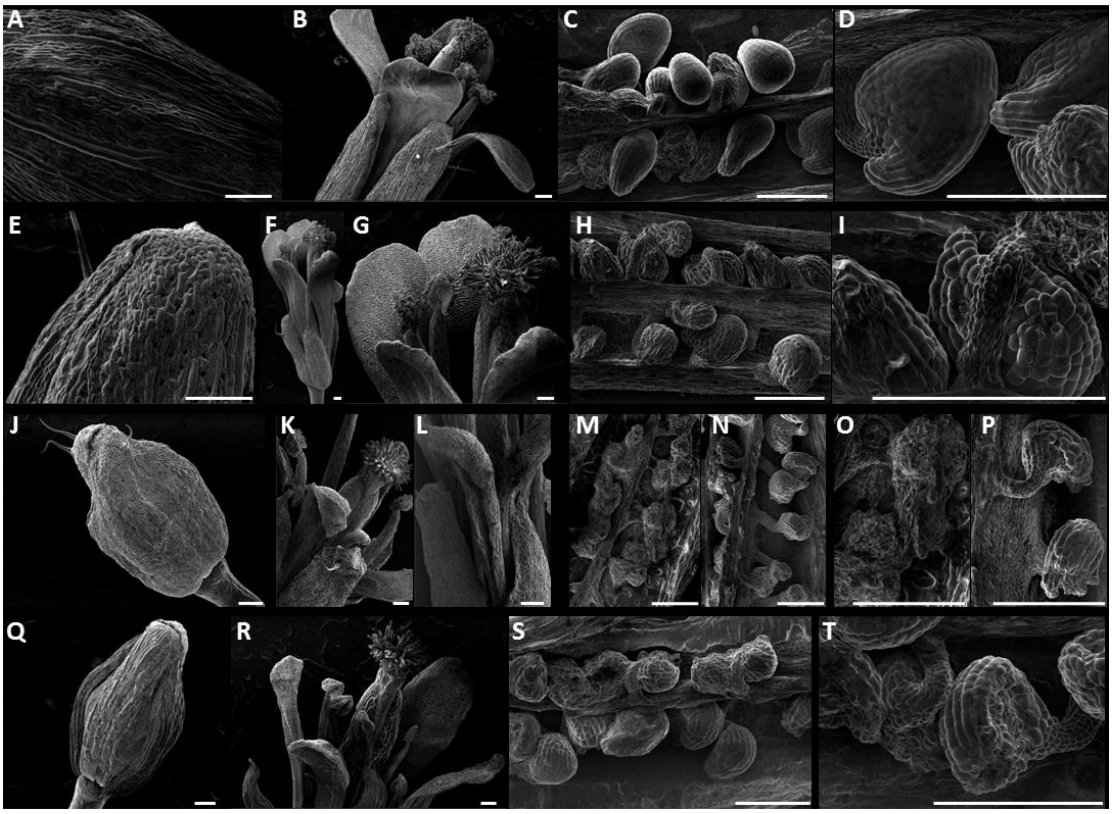


Figure 5.9 – Flower development of *acr4-2pdf2-3* and *acr4-2atml1-3* double mutants. The different stages of flower development are represented from the left to right, namely flower buds (A, E, J, and Q), flowers (B, F, G, K, L and R) and ovules (C, D, H, I, M-P, S, and T) for different genetic backgrounds: (A-D) represent the wild-type, (E-I) *acr4-2*, (J, K, M and O) are *acr4-2pdf2-3,* (L, N and P) *acr4-2atml1-3*, and (Q-T) correspond to *acr4-2pdf2-3* flowers expressing FLAG-PDF2 under the *PDF1* promoter. Scale bar is 100 µm.

The post-germination phenotype of *acr4-2pdf2-3* includes sporadic fusions of floral organs and stems and misshaped floral organs (Figures 5.8, 5.9). The single mutant *acr4-2*, *pdf2-3* and *atml1-3* showed fewer giant cells present in the sepals compared to the wild-type, while the double mutant *acr4-2pdf2-3* and *acr4-2atml1-3* have almost no giant cells in the sepals (Figure 5.9).

5.2.4.1 *ACR4*, *ATML1* and *PDF2* act in the same genetic pathway during embryo development.

Self seeds from *acr4-2atml1-3* double homozygous plants showed a germination rate that was lower than either wild-type or *acr4-2* single mutants. Most of the seedlings that germinated produced either cotyledons which are in non-opposite positions or only produced one cotyledon. Similar seedling phenotypes were found in the progeny plants homozygous for *pdf2-3* and heterozygous for *acr4-2* (or vice versa) and these misshapen seedlings were later genotyped as double homozygous (Figure 5.10, 5.11A). A few seeds were recovered from *acr4-2pdf2-3* double homozygous plants grown in short day conditions but did not germinate. The seeds harvested from plants homozygous for *acr4*-2 and heterozygous for *pdf2-3* (*ccPp*) are similar to the *acr4-2* seeds but there is an increased frequency of seed abortion in this background. The seed abortion phenotypes and germination problems associated with *acr4-2* mutants heterozygous for either *pdf2-3* or *atml1-3* do not appear to be due to the presence of either *acr4-2pdf2-3* double homozygous or *acr4-2atml1-3* double homozygous embryos, since the frequency of double mutant plants in the self-progeny of these backgrounds does not vary significantly from that predicted by mendelian genetics. Defects are thus likely due to exacerbation of the sporophytic ovule phenotypes associated with loss of function of *ACR4*. In order to avoid the maternal sporophytic seed phenotype associated *acr4* mutation, I therefore used the progeny of plants homozygous for *pdf2-3* and heterozygous for *acr4-2* (*ppCc*) to study embryo development and obtain double homozygous plants. (Figure 5.10, 5.12)

I predicted that the post germination phenotypes of the double mutant *acr4-2pdf2-3* or *acr4-2atml1-3* might be related to the disruption of the epidermal maintenance pathway. To study the impact of these genes in epidermal specification I studied

different stages of seed development and tested seedling permeability to toluidine blue (Figure 5.11B and 5.12). Like the single mutants *acr4-2*, *pdf2-3* or *atml1-3* the seedlings of the double mutant *acr4-2pdf2-3* (*ppCc*) or *acr4-2atml1-3* are significantly (t-test inferior to 2 %) more permeable to toluidine blue than the wild-type. However, the difference between double mutants and respective single mutants, particularly *acr4-2* is not significant (t-test superior to 2 %).

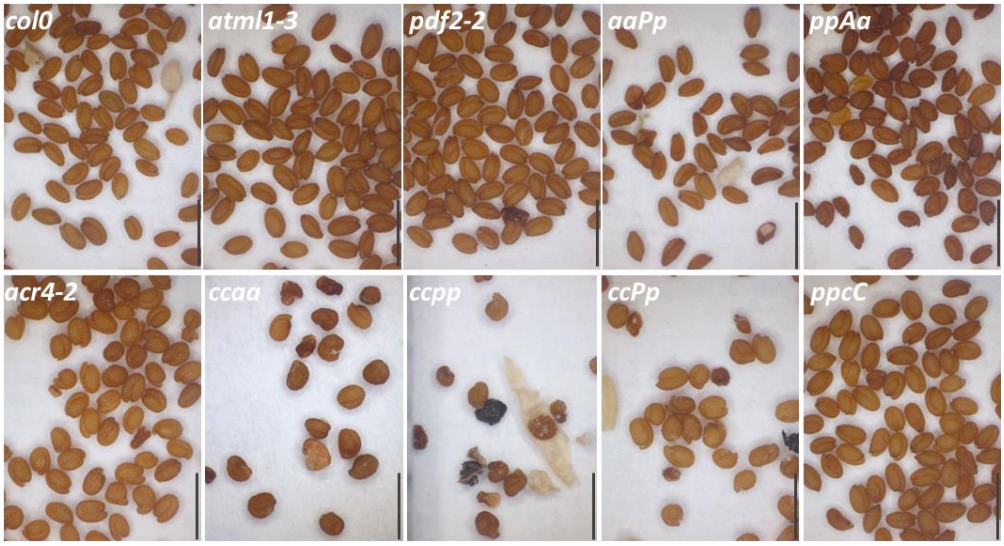


Figure 5.10 – Seed phenotypes of epidermal double mutants. The homozygous for one gene (xx) and heterozygous for the other gene (Yy) are represented as *xxYy*, were "*c*", "*p*" and "*a*" corresponds to *acr4-2*, *pdf2-3* and *atml1-3* respectively. The scale bar is 1 mm.

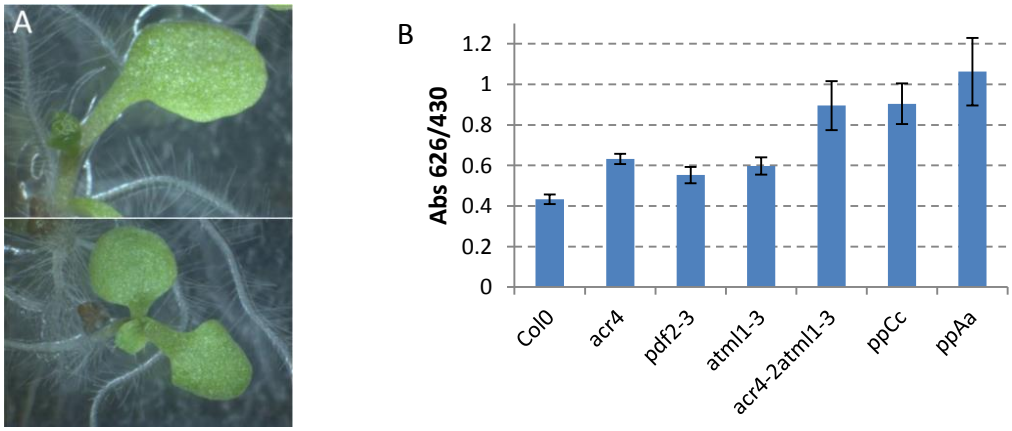


Figure 5.11 – Seedling phenotype of acr4-2 pdf2-3 or *acr4-2 pdf2-3* double mutants. A) Seedling phenotype of *acr4-2pdf2-3* or *acr4-2atml1-3*. B) Seedling permeability to toluidine blue, representation of toluidine blue/chlorophyll ratio (Abs626/430) for double mutants, respective single mutants and wild-type.

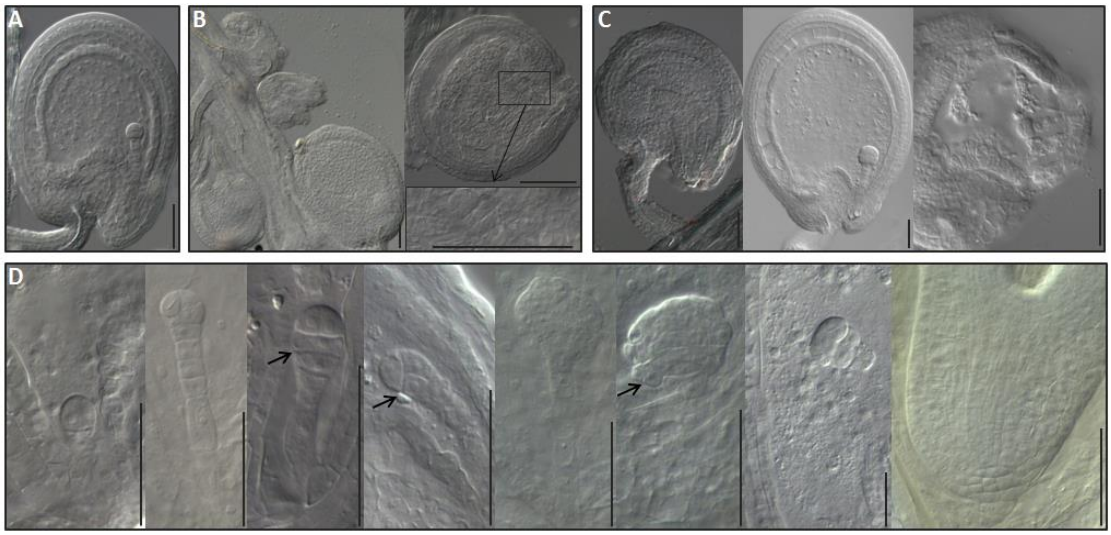


Figure 5.12 – Embryo phenotype of potential *acr4-2pdf2-3* and *acr4-2atml1-3* double mutants. (A) and (B) show respectively wild-type and *acr4-2* seed phenotypes. (C) From left to right are represented seeds *of acr4-2atml1-3* double mutant, and seeds of two *acr4-2pdf2-3* lines segregating either *acr4-2* (*ppCc*) or *pdf2-3* (*ccPp*). (D) *ppCc* embryos can show several types of phenotypic abnormality related to suspensor length, abnormal cell divisions, epidermal defects and basal polar defects. However all the phenotypes shown are present at very low frequencies (<1%) and are occasionally observed in *acr4-2* or *pdf2-3* single mutants, Most *acr4-2pdf2-3* mutant embryos develop normally. The scale bar is 50 µm.

In siliques from *ppCc* plants a significant proportion of ovules were not fertilized and the proportion of seeds with interesting embryo phenotypes could therefore not be directly correlated to the *acr4-2pdf2-3* double homozygous genotype. Abnormal embryo phenotypes in basal zones included abnormally short or long suspensors and abnormal division planes in the basal pole of the embryo. In severe cases the basal defects affect the viability of the embryo, which arrests at the globular stage. The frequency of such defects remained very low. The epidermal defects occasionally found in potential *ppcc* embryos are similar to those found in most *ppAa* embryos including abnormal division planes in the embryo epidermis that creates an apparent double epidermal layer in some regions (Figure 5.12). However the frequency of the defects again remained low, suggesting that most *ppcc* embryos are viable and show minimal disruption to the developing epidermal layer.

The expression of *pPDF1-FLAG:PDF2* complemented some phenotypes of *acr4-2pdf2-3* namely cotyledon disposition, ovule fusion, flower organ persistence and most importantly plant sterility (Figures 5.8 and 5.9). The flower and sepal

phenotypes associated with overexpression of *PDF2* were largely dominant. However, embryo phenotypes resulting from division problems were still observed very sporadically. Interestingly the expression of *pPDF1-FLAG:PDF2* in *acr4-2pdf2-3* double mutants only restores ovule phenotypes to a state similar to that observed in *acr4-2* single mutants. This suggests that the ovule phenotypes in *acr4-2* single mutant are likely not attributable to lower expression levels of *PDF2/ATML1*, but likely to the loss of expression of independently regulated genes.

Our results suggest that the combination of *acr4-2* with either *atml1-3* or *pdf2-3* has a significant impact in flower development when compared to the respective single mutants but this difference is considerably less evident in the developing epidermis during embryogenesis. These results are globally consistent with *ACR4, PDF2 and ATML1* acting in the same pathway during embryonic epidermal development.

5.2.5 *ACR4* and *WOX* genes interact genetically during embryogenesis and after germination

Promoter analysis of *ACR4* suggested that WOX proteins might play an important role in regulating the expression of this gene. In order to study the genetic interactions between *ACR4* and WOX transcription factors-encoding genes, I produced different mutant combinations using a candidate approach. The choice of *WOX* genes was based on expression pattern analysis of this gene family in the embryo and SAM, concentrating on cases where an overlap between *WOX* and *ACR4* expression was predicted.

The *wox1prs*, *prs*, *wox2-2*, *wox2-1wox8-1* (WS background) mutants have already been described in the literature (Breuninger et al. 2008; Haecker et al. 2004; Vandenbussche et al. 2009). These mutants were crossed with *acr4-2* (Col0 background). *WOX1* and *PRS* are required for normal leaf development and in particular for adaxial/abaxial polarity and leaf margin patterning (Nakata et al. 2012; Vandenbussche et al. 2009). Lateral sepal growth is repressed in the *prs* mutant whereas the adaxial and abaxial sepals are similar to wild-type with exception of sepal margins (Matsumoto and Okada 2001). The *wox1prs* double mutant displays

narrow petals and sepals and narrow, thickened leaves with increased adaxial tissue and leaf margin curling (Nakata et al. 2012; Vandenbussche et al. 2009).

The presence of the *acr4-2* mutation causes an exacerbation of the flower phenotype of both *wox1prs* and *prs*. Globally, floral organs are more misshapen, asymmetric and narrow in combinations of these mutants with *acr4-2* (Figures 5.13, 5.14). Moreover, *acr4-2wox1prs* rosettes show irregular leaf edges and increased leaf curling compared to *wox1prs* (Figures 5.13, 5.14). This phenotype is less evident in the *acr4-2prs* double mutant.

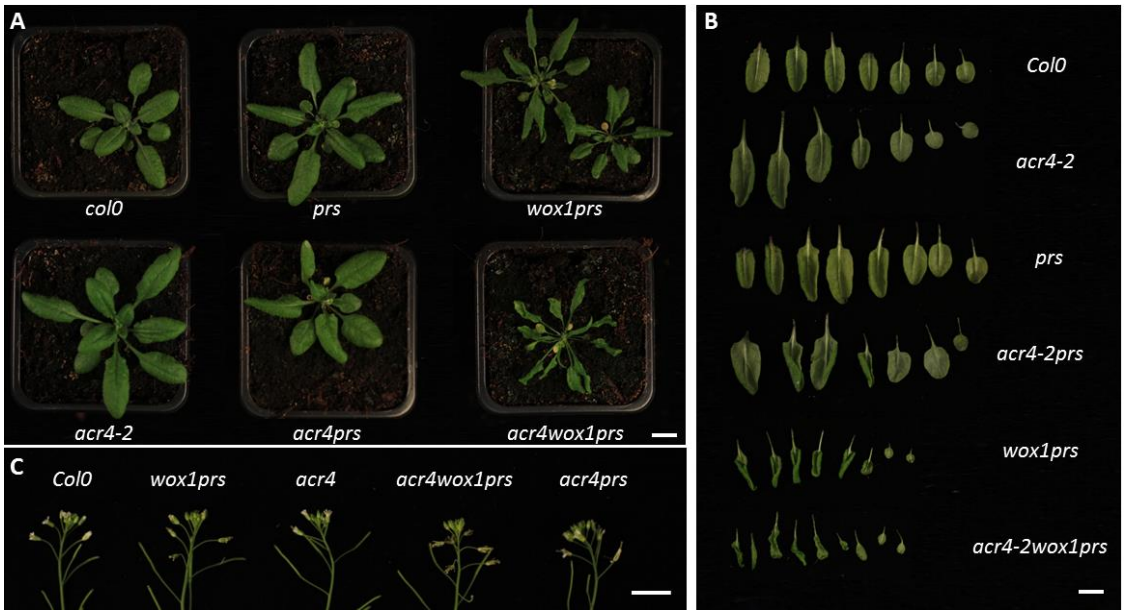


Figure 5.13 – Post-germination phenotype of *acr4-2wox1prs* and *acr4-2prs* mutants. Representation of rosette (A), abaxial side of rosette leaves (B) and inflorescence (C) phenotypes of the different mutant backgrounds. Scale bar is 1 cm.

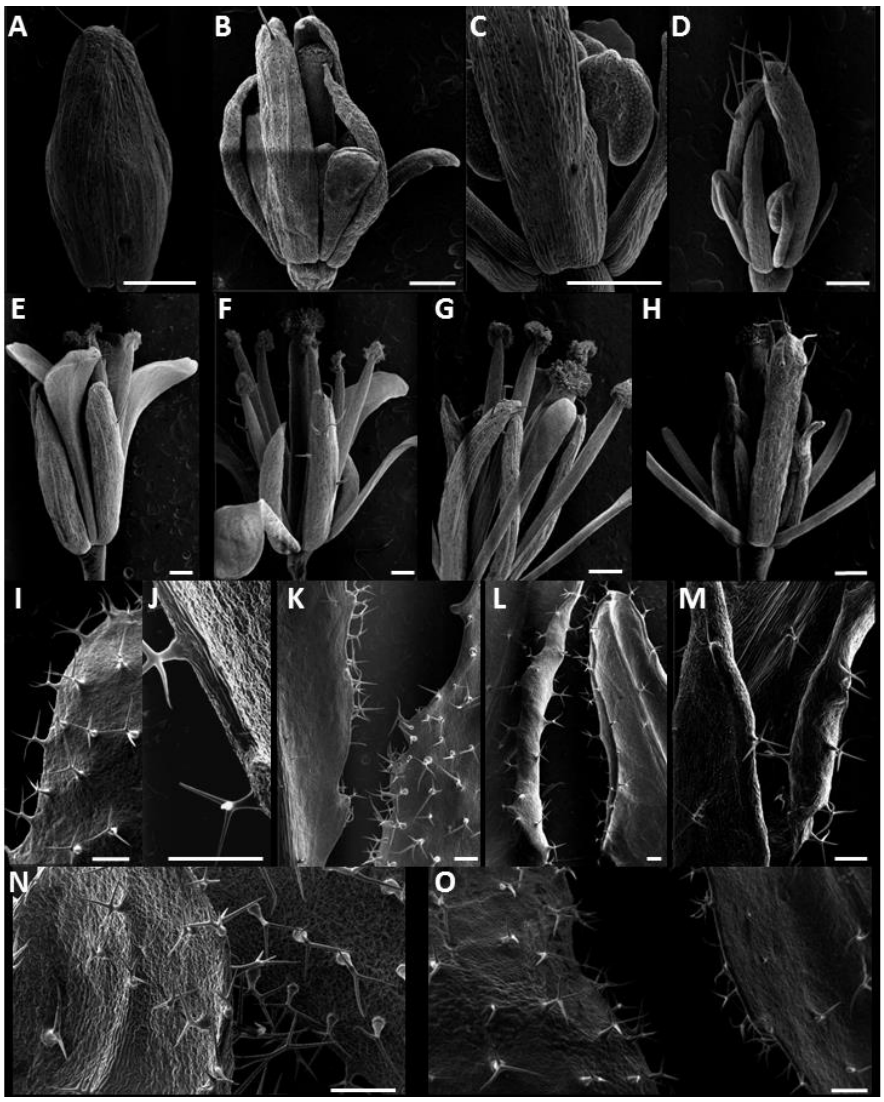


Figure 5.14 –The genetic interaction between *ACR4* and *WOX1*/*PRS*. Representation of flower bud (A-D), flower (E-H) and cauline leaves (I-O) for different genetic backgrounds: *prs* (A, E, I, J), *acr4-2prs* (B, F, K), *wox1prs* (C, G, L), *acr4-2prswox1* (D, H, M), *acr4-2* (N) and Col0(O). The scale bar is 200 µm.

Mutations in none of the four *WOX* genes selected exacerbate the *acr4-2* seed shape phenotype. The phenotypes already described for *wox2wox8* (Breuninger et al. 2008) were also present in *acr4-2wox2wox8* triple mutant, namely the abnormal division planes identified at early stages of embryo development, in particular the division planes observed in the apical part of 16 cell stage embryos when the epidermis is defined (Figure 5.15B). Interestingly around 10 % of *acr4-2wox2-1wox8-1* embryos showed additional epidermal defects in the basal part of embryos at early globular stage onwards (Figure 5.15D). In addition, a few *acr4-2wox2-1wox8-1* embryos showed defects in the root pole (Figure 5.15D). At the heart stage a small number of

acr4-2wox2-1wox8-1 embryos were found to have arrested at the globular stage (Figure 5.15E).

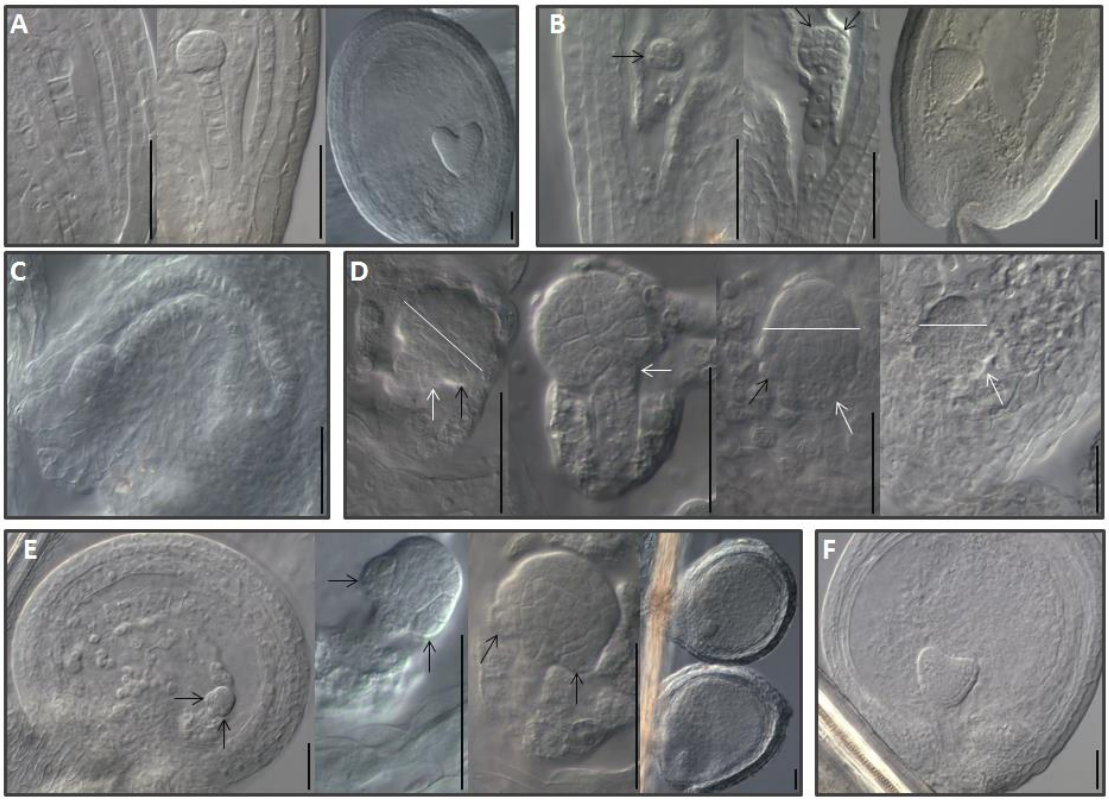


Figure 5.15 – Embryo phenotypes of *acr4-2wox2-1wox8-1*. Compared to the wild-type embryogenesis (A), the *wox2-1wox8-1* mutant showed abnormal divisions at early stages of embryo development (B). In addition to these phenotypes, the triple mutant *acr4-2wox2-1wox8-1* showed substantial problems in the basal regions, typified by abnormal cell divisions of the basal epidermis (C, D, E). These abnormalities are observed in at least 10 % of the embryos. The *acr4-2* seed phenotype (F) is also present in this triple mutant. The scale bar is 50 µm.

In general, our results suggest that *WOX* genes and *ACR4* may work together in epidermis specification and cell identity. However, the mechanisms underlying these interactions need to be studied in greater detail.

5.3 Discussion

I have provided extensive genetic evidence to support the existence of a feedback loop which acts both during and after embryogenesis to allow epidermal maintenance in *Arabidopsis* embryos. The core of this model is the activation of expression of the *ATML1/PDF2* transcription factor encoding genes via the ACR4 signalling pathway and the transcriptional regulation of *ACR4* by ATML1/PDF2 which additionally

regulated their own transcription. In this chapter I aimed to understand the connections between these elements and try to expand the model by identifying novel interactions.

Feedback loop model to define epidermis specification pathway: a top-down approach.

The identification of ATML1/PDF2 as key transcription factors for epidermal cell fate identity and the role of the L1 specific receptor kinase ACR4 in cell layer organization were the first steps required to define a pathway for epidermal cell fate specification. The interactions between these genes have been the subject of considerable speculation, and a primitive positive feedback model had previously been proposed (Abe et al. 2003; Tanaka et al. 2007). However, this model was based on transcriptional analysis of seedlings with a severely compromised/absent epidermis. The absence of expression of epidermally expressed genes such as *ACR4* in this context is neither surprising, not particularly meaningful. My analysis of the transcription levels of *ACR4*, *PDF2* and *ATML1* in different mutant backgrounds and lines overexpressing *PDF2* (*pPDF2-FLAG:PDF2*) is consistent with a negative feedback loop in seedling tissue. We compared the properties of the negative feedback model supported by my results, and the previously proposed positive feedback loop using mathematical simulations. Our simulations suggest that our negative feedback model would entrain a robust steady state of epidermal identity that fits with the robust nature of epidermal cell identity observed in post-embryonic growth. In contrast, simulations of the positive feedback loop show that it tends to generate a bistable system that is relatively sensitive to system perturbations leading to the risk of a catastrophic and permanent loss of epidermal identity in situations where ligand concentrations vary.

Interestingly, although loss of epidermal identity is not often observed in post embryonic growth, it represents a key step in early embryogenesis. The genes *ATML1*, *ACR4* and *PDF2* are all expressed throughout the 8 cell embryo in *Arabidopsis.* The cells of the wild-type 8 cell embryo divide periclinally to the surface of the embryo to generate an “inside” and an “outside” population of cells in the dermatogen embryo. Critically, the “inside” cells rapidly lose expression of all

three genes. This could be explained by a total loss of input (Ligand) to the system, as would, for example, be predicted if the input came from the cells surrounding the developing embryo. This hypothesis is plausible for both positive and negative feedback loop model. However it might also suggest that the cells of the early embryo are more sensitive to changes in ligand concentration, possibly indicating that the situation during early embryogenesis resembles more closely our positive feedback model than our negative feedback model. The relative sensitivity of early embryogenesis to the number of active copies of *ATML1/PDF2* could also point to a situation similar to our positive feedback model. However, the measurement of transcript levels in young embryos, necessary to test this hypothesis, is technically beyond the scope of my project.

The importance of *ACR4* in the epidermal signalling pathway is poorly understood. The phenotype of the single *acr4-2* mutant is relatively weak when compared to the severe phenotypes caused by loss of the maize homolog *cr4* or the embryo lethal phenotype of *atml1-3pdf2-3*. Functional redundancy between ACR4 and four other CR4-related proteins (CRRs) could explain to some extent the weak phenotype of *acr4-2*. In particular the expression pattern of *CRR3* partially overlaps with *ACR4*, especially during embryogenesis (Le et al. 2010; Winter et al. 2007). However unpublished work from the group suggests that this proposed redundancy does not exist and that quintuple *acr4crr1crr2crr3crr4* mutants closely resemble *acr4* single mutants (Gwyneth Ingram personal communication). Interestingly ACR4 has been reported to work together with other receptor kinases including ALE2 and CLV1, potentially forming protein dimers (Gifford et al. 2005; Stahl et al. 2013; Tanaka et al. 2007). *ale2-1* mutants display ovule fusion phenotypes (Tanaka et al. 2007) that are comparable to those seen in *acr4-2pdf2-3* double mutants. In fact, *ale2-1* single mutant recently grown in the laboratory were almost indistinguishable from *acr4-2pdf2-3* double mutants. However, *ale2-1* mutant remain viable, which still begs the question of whether other receptor kinases, potentially those related to CLV1, could show strong functional redundancy with ACR4 and ALE2.

The combination of *acr4*-2, with *ale1-4*, *zou-4* or *gso1gso2* is lethal ((Tanaka et al. 2007; Watanabe et al. 2004; Xing et al. 2013) and unpublished results), while *ale1-4*

pdf2-3 or *ale1-4atml1-3* showed very strong seedling phenotypes, very high toluidine blue permeability and a significant proportion of seed abortion (San-Bento et al., submitted, results produced by Gwyneth Ingram). The triple mutant *pdf2-3gso1gso2* is also embryo lethal. The similarity between the genetic interactions of *ACR4, ATML1* and *PDF2* with *ALE1, ZOU,* or *GSO1/GSO2* further supports the hypothesis that ACR4 acts in the same pathway as ATML1/PDF2. Moreover, these results suggest that ACR4/PDF2/ATML1 pathway is independent of the ALE1/GSO1/GSO2 pathway and more importantly that the disruption of both pathways is fatal for the embryo. Single *ale1* mutants, and double *gso1/gso1* mutants, show weak and strong defects respectively in seedling cuticle permeability, but neither plays an apparent role in epidermal specification (Tanaka et al. 2001; Tanaka et al. 2007; Tsuwamoto et al. 2008). In addition the expression of *ACR4, ATML1* and *PDF2* in these mutant backgrounds is not reduced (San-Bento et al., submitted, results produced by Gwyneth Ingram, (Tanaka et al. 2007)). The fact that cuticle phenotypes associated with the loss of *ALE1, GSO1* and *GSO2* are restricted to embryo development (Tanaka et al. 2001; Tanaka et al. 2007; Tsuwamoto et al. 2008), has led us to propose that the *ALE1/GSO1/GSO2* pathway could influence embryo epidermal specification indirectly, by affecting accumulation of the ACR4 ligand during early embryogenesis. Until the identity of this ligand is known, it will, however, be difficult to test this hypothesis.

Our transcript analysis suggests that ATML1 and PDF2 negatively regulate *ACR4*. To support this theory, I have shown that ATML1 and PDF2 bind to the *ACR4* promoter *in vivo* (Chapter 4). However, if ATML1 and PDF2 repress *ACR4*, how is *ACR4* activated? Are the regulators of ACR4 also responsible for the regulation of other genes involved in epidermis specification? As suggested above one possibility is that ATML1 and PDF2 act to positively influence *ACR4* expression in early embryos. Other HD-ZIP IVs may also act positively on the *ACR4* promoter. Alternatively as suggested by the analysis of the *ATML1* promoter (Takada and Jurgens 2007) other transcription factor families could also be involved.

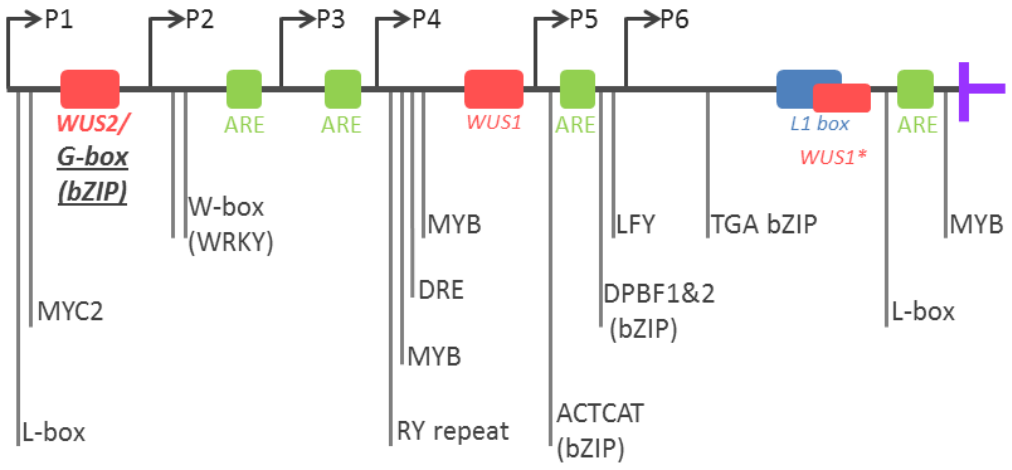


Figure 5.16 – Binding sites in the *ACR4* promoter. The localization of P1-P6 *ACR4* promoter fragments is represented by arrows. The binding sites of WUS and potentially other WOX proteins is shown in red, AREs are green, the L1 box is blue and the other binding sites are shown in grey. The localization of transcription factor binding sites is based on AtCoecis *ACR4* promoter analysis considering the 1 kb region upstream of the starting codon.

In order to answer some of these questions I investigated the *ACR4* promoter both to help validate the ChIP results and to identify other potential ACR4 regulators. I studied six promoter versions of different sizes (*P1-P6*) to identify active regulatory regions (Figure 5.16). I designed these promoters in relation to the positions of three WUS binding sites, four AREs and the L1 box present in the *ACR4* promoter. However, other potential binding sites are present in the promoter which could contribute to the expression pattern of *ACR4* promoter variants that I produced (Figure 5.16). Firstly, it is important to note that the WUS2 binding site in addition to being a potential binding site of WUS (Busch et al. 2010) is known as the core G-box, the binding site of bZIP and GBF transcription factors associated with light responses (Hudson and Quail 2003; Menkens et al. 1995; Terzaghi and Cashmore 1995). Three other binding sites associated with bZIP transcription factors are present in the *ACR4* promoter, together with potential binding sites for MYB, MYC, and WRKY transcription factors (Figure 5.16). Most of these binding sites are associated with stress responses. The possibility that *ACR4* is regulated by stress response signalling pathways reinforces the potential link between epidermal cell fate specification and stress response which I have pinpointed in previous chapters. In particular, my analysis of potential targets of ATML1/PDF2 and new epidermal regulators also suggest a link to stress responses (Chapter 4).

The results of my *ACR4* promoter analysis point to a complex regulation of this gene involving multiple promoter regions in different tissues. To expand my findings it would be interesting to follow the expression pattern of various promoter variants from early stages of embryo development. In order to do this, however, it will be necessary to use a better fluorescent marker than that present in the vector used in my study. Increasing the number of fluorescent proteins driven by the *ACR4* promoter would directly increase the output signal.

Regulation of ACR4 by L1 box

The similarity of *ATML1* and *ACR4* promoter expression patterns, and the results from my analysis, support the theory that the L1 box is required for L1 specific patterns in the embryo, as previously suggested from *ATML1* promoter analysis (Takada and Jurgens 2007).

In order to analyse the role of the L1 box in *ACR4* regulation I produced two versions of this promoter that disrupted the L1 box in different places: m1 and m2. In particular, the m2 mutation also disrupts the WUS1*, which can have a significant impact in the expression of this promoter. Although both mutations of the L1 box are not sufficient to totally abolish *ACR4* expression in embryo, root or inflorescences, the expression in the two constructs in the embryo L1 layer is significantly weaker than that of the full length promoter P1.

The expression of all promoter versions in the embryo is L1 specific. However the intensity of L1 expression and its localization change between the promoter versions. This suggests that the smallest fragment P6 is sufficient to restrict the expression to the epidermis, although it is not sufficient to ensure the full expression pattern of the native *ACR4* promoter, P1, since the expression of P6 in the apical and basal regions of the embryo is considerably weaker than for P1. In inflorescences the expression of all promoter versions is again L1 specific, suggesting once more that P6 is sufficient to drive L1 *ACR4* expression. Assuming that the L1box/WUS1* is the central regulatory region of the P6 promoter, and taking into account the results obtained using full promoters in which this element is mutated I can conclude that the L1 box/WUS1* is important for the L1 specific expression in the embryo and

inflorescence, but is not sufficient to drive expression in the apical embryo. A role for WUS1* in the epidermal pattern of P6 cannot be excluded. It should also be noted, that one ARE remains within the P6 construct, as well as at least two other recognized TF binding sites, so that the activity of P6 cannot be attributed solely to the L1box/WUS1*.

My results underline the importance of the L1 box in the regulation of *ACR4* expression and support the idea that even if PDF2 and ATML1 negatively regulate *ACR4*, other HD-ZIP IV proteins could activate its expression and potentially compete for the L1 box binding site. As previously mentioned it is also possible that although PDF2 and ATML1 negatively regulate *ACR4* in seedling tissues, their regulatory role in embryos is different.

Regulation of ACR4 by AREs

Auxin is integrally involved in embryo patterning. Briefly, after the first division of the zygote, auxin accumulates in the embryo proper until the dermatogen stage (Friml et al. 2003). Auxin accumulation then switches to a maximum in the basal pole in early globular stage and at the late globular stage is also accumulated in the regions were cotyledon will be initiated, in the apical part of the embryo. PIN transporters have an active role in both establishing and remodelling of these dynamic auxin maxima (Friml et al. 2003; Friml et al. 2004; Weijers et al. 2005b). PIN distribution, and in particular the distribution of PIN1, overlaps partially with epidermal cells (Friml et al. 2003; Tanaka et al. 2006).

My *ACR4* promoter analysis suggests that ARE regulatory regions may be involved in *ACR4* regulation in the embryo. The loss of L1 expression in P3 compared to P2, and the formation of the band of L1 expression with loss of basal expression in P4 could be correlated with the loss of the ARE1 and ARE2 binding sites respectively. In addition, the expression in the basal pole of the embryo is partially restored in the P6 construct in which ARE3 has been deleted. One interpretation of these observations is that ARE1 promotes L1 expression, ARE2 drives expression in apical and basal poles of the embryo and ARE3 suppresses the expression in the basal pole. The loss of expression in the basal pole could thus be explained by a loss of response

to auxin accumulation since in embryogenesis a peak of auxin is localized in the basal pole (Benkova et al. 2003; Friml et al. 2003). A lack of expression in the apical part of the embryo might also correspond to a lack of auxin perception, since auxin accumulates preferentially in the epidermis during embryogenesis (Friml et al. 2003; Tanaka et al. 2006). However, the loss of L1 expression of P3 could be an cumulative effect of the deletion of WUS2 and ARE1 since the expression levels of P2 are lower than those of the full promoter. As mentioned before other known binding sites in the *ACR4* promoter (Figure 5.16) could also participate in the changes in expression patterns in these constructs, and only the generation of more detailed mutated promoter versions will be able to clarify the role of the AREs.

The expression pattern of P3 is comparable to that of *m1* and *m2*, consistent with epidermal expression being co-ordinately regulated by the L1box and ARE1 containing regulatory regions. However, the basal pole expression in m1 and m2 is comparable to that of full length promoter whereas P3 expression in the same region is slightly lower, suggesting that both ARE1 and ARE2 are necessary for the expression in the basal pole.

AREs are bound by ARFs. In the embryo, the *ARF* gene promoters which drive expression partially or exclusively in the L1 include those of *ARF4, 9, 10* and *18* and the ARF encoding genes expressed exclusively in the apical or basal region of the embryo are *ARF3* and *ARF6* respectively (Rademacher et al. 2011). Interestingly at the globular stage of embryogenesis, the *ARF5* promoter is expressed in the same lateral regions as the *ACR4* P4-P6 promoters.

In general the expression pattern of the *ACR4* promoter in inflorescences appears less affected by loss of ARE containing promoter domains than was observed in embryos. However, the decrease of expression of P3 constructs compared to the full length P1 or P2 promoters supports the theory that the ARE1 containing fragment promotes L1 expression. Based on their expression domains ARF1, 2, 5, 6, 7 and 8 are the best candidates for potentially regulating *ACR4* expression in inflorescence meristems (Vernoux et al. 2011).

The expression pattern of *ACR4* promoter constructs in roots indicates that the ARE1 and ARE2 containing fragments are required for expression in the primary root tip and lateral root initials respectively. The study of *acr4-2* mutant roots has enabled the identification of a role for *ACR4* in the coordination of pericycle cell divisions in lateral root initiation and the suppression of cell division in cells adjoining the site of lateral root initiation. In addition *ACR4* has been shown to limit cell division in columella cell lineage (De Smet et al. 2008). In particular, the auxin responsiveness of root initiation is impaired in *acr4* mutants (Benkova et al. 2003; De Smet et al. 2008). The relationship between *ACR4* signalling and auxin responses in lateral root initiation, and its role in root tip cell patterning could involve the activation of ACR4 expression by auxin response factors via the ARE1 and ARE2 regulatory elements of the *ACR4* promoter. The GUS expression in early lateral root initials is clearly absent in the P4 version of the *ACR4* promoter, and the same trend is seen in some P3 lines, whilst others have a GUS expression pattern weaker than the full length promoter and more restricted to the regions corresponding to the auxin maximum(Benkova et al. 2003; Tanaka et al. 2006), suggesting that ARE1 could also be active in lateral root initiation. In terms of potential regulators mediating this expression, in the primary root tip *ARF1, 2, 6, 11, 13* are all expressed in regions overlapping with *ACR4* expression (Rademacher et al. 2011).

Organ initiation is widely accepted to occur at auxin maxima established by polar transport mediated by PIN transporters (Benkova et al. 2003). This mechanism is valid for organ primordia in the lateral region of the SAM, lateral root initiation and the inception of cotyledon initials in the embryo (Benkova et al. 2003; Tanaka et al. 2006). Several phenotypes of epidermal mutants are related to organ initiation. These include the ovule development phenotypes of *acr4-2*, *acr4-2pdf2-3* and *acr4-2 atml1-3* (defects in integument initiation), and the formation of monocotyledonous seedlings, or seedlings with non-opposite cotyledons in *acr4-2pdf2-3*and *acr4-2 atml1-3*. One possible explanation for these phenotypes is that auxin transport is physically perturbed in mutants when an intact epidermal cell layer is compromised. To explore this possibility one would analyse PIN1 localization in embryos of mutant backgrounds of interest. Another possibility is that the perception of auxin maxima, and their translation into nascent organs, is intricately linked to elements of

the epidermal maintenance pathway. In order to dissociate these possibilities it would be of interest to study the behaviour of auxin markers such as DII-VENUS (negatively correlated with auxin accumulation) (Vernoux et al. 2011) and DR5-VENUS (positively correlated with auxin signalling) (Ulmasov et al. 1997), in backgrounds with compromised epidermal specification. The fact that epidermal defects are not significantly enhanced in, for example *acr4-2atml1-3* mutants, compared to *acr4-2* single mutants, whereas cotyledon positioning and ovule formation defects are greatly enhanced, might indicate that the later, and potentially more interesting possibility is correct.

Regulation of ACR4 by WOX genes

My results suggest that factors that bind to WUS2 containing fragment of the *ACR4* promoter are partially responsible for the expression in the L1 layer of the embryo and are required for the epidermal expression pattern in the main root tip. This might reflect a role of the WUS2/G-box complementary to that of the L1box in the root and embryo epidermis. The m1 and m2 constructs in which WUS1*/L1 box is disrupted, also shows significant loss of L1 expression in the embryo and main root tip.

The expression patterns of *WOX9*, *PRS* and *WOX2* in the embryo overlap partially with that of *ACR4* (Haecker et al. 2004). In the root *WOX13, WOX14* and to some extent *WOX5* expression overlaps with that of *ACR4*. *WOX13* and *WOX14* expression overlaps with that of *ACR4* in lateral roots and the primary root tip (Deveaux et al. 2008). *WOX5* expression is restricted to the root quiescent centre, where it is required for stem cell maintenance and proper root meristem development (Sarkar et al. 2007). *ACR4* is also expressed in the root meristem, although weakly in the quiescent centre, and recent data from the laboratory of Ruediger Simon suggest that ACR4 interacts with CLV1 to restrict *WOX5* expression and consequently stem cell fate. In addition CLE40 acts upstream CLV1/ACR4, activating this signalling pathway (De Smet et al. 2008; Stahl et al. 2013; Stahl et al. 2009).

Interestingly, for some constructs the expression at the centre of the meristem appeared to be weaker than at the organ boundaries or in organ primordia. This pattern indicates that the *ACR4* promoter is controlled by different regulatory regions

in the diverse regions of the meristem. In particular the binding sites WUS2 and L1box/WUS1* seem to be important for the expression in the centre of the meristem. Since m2 inflorescences display this pattern but not m1, one can tentatively conclude that WUS1* disruption rather than L1 box disruption causes this expression pattern defect. It is formally possible that WUS activates *ACR4* expression. Recently it has been shown that WUS moves to the L1 and L2 layer of the central zone of the meristem (Yadav et al. 2011). Another possibility is that other than WOXs, and ARFs are necessary for this pattern. In the shoot meristem, *PRS* and *WOX1* (organ primordia), *WUS* (central zone) and *WOX13* are the best candidates for *ACR4* regulation based on expression pattern analysis (Deveaux et al. 2008; Matsumoto and Okada 2001; Nakata et al. 2012; Yadav et al. 2011).

Again in order to be sure of the role of WUS2 in *ACR4* regulation, mutated versions of the promoter would need to be studied. It would then be possible to identify potential regulators, for example by using a one hybrid approach, or targeted ChIP studies.

I chose a genetic approach to study the relationship of *ACR4* with selected *WOX* genes. I showed that *acr4-2wox2-1wox8-1* embryos show epidermal defects similar to *acr4-2-pdf2-3* embryos, which may provide a preliminary indication to support the theory that *WOX* genes are involved in the epidermal pathway. The mechanism, however, is unclear. The WOXs could act upstream or downstream of *ATML1/PDF2* or act together with HD-ZIP IVs in the regulation of epidermal targets, including *ACR4*. An interesting possibility was recently uncovered by work from the laboratory of Jennifer Fletcher showing that the CLE8 peptide, expressed in early embryos and endosperm, positively regulates *WOX8* expression in early embryogenesis (Fiume and Fletcher 2012). The loss of *CLE8* results in several seed and aberrant embryo phenotypes that include defects in the boundary of embryo and suspensor that can lead to embryo lethality (Fiume and Fletcher 2012). These embryo phenotypes are similar to the *acr4-2wox2-1wox8-1* triple mutant and basal embryo phenotypes of *pdf2-3*/*atml1-3*/*acr4-2* single or double mutants. It is possible that CLE8 is a potential ligand of ACR4 and this signalling pathway is involved in WOX regulation responsible for apical/basal cell fate. The fact that these phenotypes occur

sporadically in epidermal mutants suggest gene redundancy at the level of the receptor kinases and regulators, which could be a barrier to uncovering ligands and targets. To investigate the theory that CLE8 peptides are the ligands of ACR4, a first experiment would be to cross *cle8* mutant with the epidermal mutants, since the lack of *CLE8* is sufficient to perturb basal cell fates. Another possibility would be the construction of RNAi lines that interfere with the expression of *CLE* genes expressed in and around the early embryo.

In addition to the phenotypes observed in *acr4-2wox2-1wox8-1* embryos, I also showed that *acr4-2* significantly exacerbates the leaf, petal and sepal phenotype of *wox1/prs* double mutants. It has been proposed that one of the functions of epidermal specification is to promote endoreduplication and differentiation. This is one explanation for the lack of giant cells in the sepals of *atml1* and *acr4* mutants. In turn, giant cells play an important role in organ shape both in leaves and sepals (Roeder et al. 2010; Roeder et al. 2012). The link between giant cells and organ shape could explain the sepal phenotypes, but it is not sufficient to explain the strong leaf phenotypes observed in *acr4-2wox1prs* triple mutants.

The flower phenotypes of *acr4-3pdf2-3* and *acr4-2atml1-3* include asymmetric growth, organ curling and organ adhesions of petals, sepals or stamens. The lines overexpressing *PDF2* and *ATML1* show narrow flower organs and leaf epinasty and spiral growth. Some of these phenotypes are superficially similar to those of *wox1prs* double mutants. However, *acr4-2* exacerbates the leaf epinasty and leaf margin indentation of *prs* or *wox1prs* mutants. The interaction between ACR4 signalling and transcriptional regulation by PRS/WOX1 is therefore not clear cut and needs to be explored in further detail. Firstly, the leaf and flower organs phenotypes of epidermal mutants and transgenic lines need to be further analysed in detail in order to understand the causes of the growth defects and compare them with wox1prs phenotypes. This could be an important first step in understanding whether ACR4 signalling works independently from WOX1/PRS with the exacerbation of the triple mutant resulting from the disruption of two functionally similar processes, or whether ACR4 signalling and WOX1/PRS functions are interconnected.

Interestingly, both PRS and WOX1 have been shown to be necessary for the maintenance of abaxial/adaxial pattering in the leaf margins and *prs-2wox1-101* mutants lacks the giant cells that form the limit between the adaxial and abaxial faces of the leaf (Nakata et al. 2012). The description of the adaxial/abaxial identity crisis of *prs-2wox1-101* double mutant has been correlated with a complex network of other transcription factors regulating leaf polarity including HD-ZIP IIIs, ASYMETRIC LEAVES2 and FILAMENTOUS FLOWER (Nakata et al. 2012). Could the epidermal pathway also play a role in abaxial/adaxial cell identity? In order to understand the role of HD-ZIP IV transcription factors in this leaf polarity network it would be necessary to study the influence of ATML1/PDF2 under- and overexpression in multiple mutant backgrounds. The interaction between *acr4-2* and *wox1prs* is encouraging but insufficient to understand if ACR4 has a role in leaf polarity.

Chapter 6 – Discussion

6.1 The role of HD transcription factors in epidermal cell fate

6.1.1 Supremacy for ATML1 and PDF2 in epidermal cell fate formation?

The main traits of basal epidermal cell fate, specified at early stages of embryogenesis will endure throughout plant development. These traits include restricted cell division planes, and the expression of genes necessary for the production of a functional cuticle. However how fate acquisition and the differentiation of epidermal traits are linked remains poorly understood. My results show that the loss of function of *ATML1* and *PDF2*, involved in epidermis cell fate, leads to embryo lethality at early stages of the embryogenesis due to defects in epidermal specification that lead to impairment of embryo patterning.

The atml1-3pdf2-3 phenotype is more severe than the lethal seedling phenotype of *atml1-1pdf2-1* previously reported (Abe et al. 2003). I have shown that *atml1-3* and *pdf2-3* are likely to be null alleles; since in addition to open reading frame disruption, levels of stable transcript are reduced in each case. Therefore the embryo lethal phenotype of the double mutant is likely to result from the total loss of function of *ATML1* and *PDF2* which may not be the case in *atml1-1pdf2-1*. The *pdf2-1* mutant transcript, is expressed at equivalent levels to *PDF2* in wild-type, and can be easily detected in the epidermis (Abe et al. 2003; Kamata et al. 2013). This allele could therefore have residual PDF2 activity.

The early lethal phenotype of the *atml1-3pdf2-3* double mutant shows that *ATML1* and *PDF2* play a central role in epidermal cell-fate specification. At the start of my thesis, I hypothesized based on the weaker *atml1-1pdf2-1* phenotype that other HD-ZIP IV proteins could act redundantly with ATML1 and PDF2 in specifying epidermal cell fate. However, I found very little genetic evidence supporting this hypothesis, with most double mutants between *atml1-3* or *pdf2-2* and other *HD-ZIP IV* mutants showing no additional phenotypes compared to single mutants. Taken together, these results suggest that within the HDZIP IV family, ATML1 and PDF2 are the only members actually required for the establishment of basal epidermal cell fate during embryogenesis.

In this context, the study of the HD-ZIP IV family during embryogenesis reveals a complex and somewhat contradictory picture. On one hand, HD-ZIP IV proteins are likely to compete for similar binding sites (Nakamura et al. 2006; Peterson et al. 2013; Wu et al. 2011). In addition, work in this thesis supports the idea that these proteins can form multiple different complexes, some of which include ATML1 and PDF2. However, genetic analysis, as just discussed, suggests that only homo and heterodimers of ATML1 and PDF2 are directly involved in specifying basal epidermal cell fate during embryogenesis, since other HD-ZIP IVs are not sufficient to promote epidermal cell fate in the absence of *ATML1* and *PDF2*. A possible scenario explaining most of these observations is that a key, but unknown, group of epidermis-specifying genes is activated specifically by heterodimers and homodimers of ATML1 and PDF2. This group of targets is then absolutely vital for the subsequent elaboration of pathways required for epidermal maintenance and differentiation. This group of core targets could well include genes encoding other members of the HD-ZIP IV family. These proteins would then be free to interact with ATML1 and PDF2 to allow the elaboration of epidermal traits (Figure 6.1).

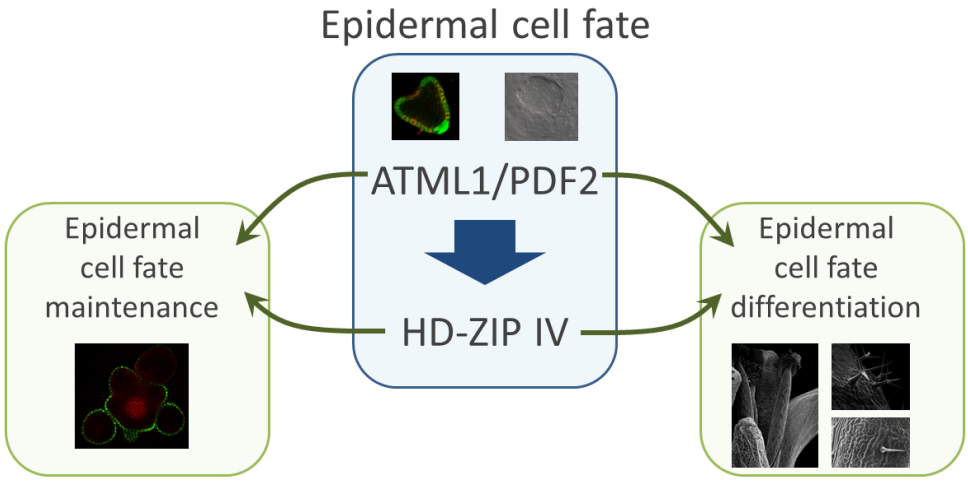


Figure 6.1 – Model of HD-ZIP IV regulatory network in epidermal cell fate. In embryogenesis, ATML1 and PDF2 regulate *HD-ZIP IV* genes required for epidermal establishment, and post-germination ATML1, PDF2 and other HD-ZIP IV are required for the epidermal cell fate maintenance and differentiation.

Consistent with this model (Figure 6.1) I found that many (although not all) members of the HDZIP IV family which are expressed during early embryogenesis contained L1 boxes in their promoters. Thus the multiple HD-ZIP IV proteins expressed during

early embryogenesis could stabilize the regulation of epidermal targets in the presence of different regulators, effectively both downstream of, and in parallel to, the activity of ATML1 and PDF2.

In order to test the veracity of this hypothesis, two key sets of experiments would be required. The first would be to ascertain the temporal sequence of expression of each of the HD-ZIP IV proteins during early embryogenesis. The second would be to test how the accumulation of each of these proteins is affected by the absence of ATML1 and PDF2. These experiments would demand the construction of lines expressing fully functional tagged protein fusions under native promoters for each HD-ZIP IV encoding gene, and their subsequent introduction into lines segregating the *atml1-3 pdf2-3* double mutant. Alternatively specific antibodies against each HD-ZIP IV protein could be generated and used in immunofluorescence experiments in wild-type and mutant backgrounds. In either case, both the generation of materials, and the difficulty of unambiguously identifying double mutant embryos at very early developmental stages in a segregating population, would make these experiments extremely involved and time consuming.

In this thesis I carried out one experiment which seems to go against the idea that all embryonically expressed HD-ZIP IVs act downstream of ATML1 and PDF2. I showed that it was possible to complement the *atml1-3 pdf2-3* double mutant by expressing *GFP-ATML1* or *FLAG-PDF2* under the *PDF1* promoter. This indicates that the *PDF1* promoter can be activated at early stages of embryogenesis independently of *ATML1* or *PDF2*. This is intriguing since analysis of this promoter had previously shown that its activity depended upon an L1 box, the cognate binding site of HD-ZIP IV proteins, leading to the suggestion that it was a target for positive activation by ATML1 and PDF2 (Abe et al. 2003; Abe et al. 2001). My results could suggest that the *PDF1* promoter can be regulated by HD-ZIP IV proteins during early embryogenesis even in the absence of ATML1 and PDF2. Thus either some HD-ZIP IV activity remains in early *atml1-3 pdf2-3* double mutant embryos, or a basal *PDF1* promoter activity can be driven by other transcriptional regulators during early embryogenesis. Disentangling these two possibilities will require considerable further work.

In support of this model (Figure 6.1), several studies suggest that both ATML1/PDF2 and other embryonically expressed HD-ZIP IV transcription factors regulate epidermal differentiation (Nakamura et al. 2006; Peterson et al. 2013; Rerie et al. 1994; Roeder et al. 2012; Takada et al. 2013). The ectopic expression of ATML1 or HDG2 indicates that each gene is sufficient to induce epidermal differentiation and epidermal identity (Peterson et al. 2013; Takada et al. 2013). My results show that high concentrations of ATML1 and PDF2 disturb trichome branching, leaf epinasty and flower organs shape. Additionally, the level of these two *HD-ZIP*s has a direct impact in the frequency of giant cells in sepals. The higher the quantity of *ATML1* and *PDF2*, the higher is the frequency of giant cells. Interestingly loss of function mutations in *HDG11* and *HDG12* also show defects in both trichome branching and in giant cell production (Nakamura et al. 2006; Roeder et al. 2012), again suggesting overlapping functions for these proteins in epidermal cell differentiation.

6.1.2 What is the meaning of the strong dose dependency for ATML1 and PDF2 during embryo development?

The specification of epidermis identity in embryogenesis is dependent on the number of active copies of *ATML1* or *PDF2* suggesting a strong dosage dependency for ATML1 and PDF2. In addition, my transgenic studies suggested that changes in *PDF2* and *ATML1* expression levels are also closely linked to phenotype severity in overexpression lines. Interestingly, in overexpression lines, I noted that early embryonic phenotypes were similar to those observed in *atml1-3* or *pdf2-3* suggesting that there is an optimal range of the expression level of *PDF2* and *ATML1*. Recent published data suggests that *PDF2* and *ATML1* are not fully redundant (Kamata et al. 2013; Peterson et al. 2013; Roeder et al. 2012; Takada et al. 2013). These results taken together with my thesis results, show that ATML1 and PDF2 likely form heterodimers both between themselves and with other members of the HD-ZIP IV family provoking important questions. Firstly, is the fact that single mutants in *ATML1* and *PDF2* have a phenotype only due to dosage (with the possibility of slightly divergent expression patterns exacerbating phenotypes in certain tissues), or is there also a real functional diversification between the two proteins. Phrased another way, are phenotypes simply a question of the total amount

of ATML1/PDF2 protein produced, or are they also due to disturbances in the relative amounts of functionally different protein complexes within the developing plant. Understanding the answer to this question has profound implications for the interpretation of data produced during transcriptional, phenotypic and genetic analysis.

To take an example, in a *pdf2* mutant, only ATML1/ATML1, ATML1/HDG, and HDG/HDG dimers are formed during embryogenesis. The question is then whether the defects I observed during embryogenesis or the transcriptional phenotypes observed in the inflorescences of *pdf2* mutant plants, are due to a lack of PDF2/PDF2, PDF2/ATML1, or PDF2/HDG dimers, or the presence of a disproportionate number of ATML1/ATML1 and ATML1/HDG dimers. My results suggest that homodimerization versus heterodimerization of ATML1 and PDF2 should have only a relatively small impact in embryo patterning, since most of the embryos of *atml1-3* or *pdf2-3* lines develop relatively normally. In theory, it would be possible to test this impact by comparing embryonic phenotypes in *pdf2* or *atml1* single mutant embryos (*ppAA* or *PPaa)* to those of embryos heterozygous for *pdf2* and *atml1* (*PpAa*), where the number of active gene copies is identical to the situation in single mutants, but the relative amount of each dimer is changed. Unfortunately this approach is complicated by the impossibility of genotyping embryos in segregating populations. However transcriptional comparison of the inflorescence apices of *PpAa, ppAA* and *PPaa* individuals, and/or comparison of their post germination phenotypes, might help to clarify this question.

The fact that *ppAa* and *aaPp* mutant embryos show much stronger defects than either *AApp* or *ppAA,* supports the idea that during embryogenesis, there is probably a threshold of ATML1/PDF2 functionality based on absolute protein amounts. Given this fact, *ppAa* and *aaPp* backgrounds would be an excellent subject for future analysis of the transcriptomic effects of reducing the activity of these proteins. Such experiments would again need to be carried out post germination in order to allow genotyping of plants prior to RNA collection.

6.1.3 New horizons in homeodomain overlapping functions?

HD-ZIP transcription factors from other families also have key roles in embryo patterning. *HD-ZIP III* genes are involved in the determination of precursors of vascular tissues with implications in apical/basal cell fate and determination of the shoot apical meristem (Emery et al. 2003; Izhaki and Bowman 2007; Prigge et al. 2005). HD-ZIP II proteins together with HD-ZIP III proteins are involved in SAM specification, adaxial identity in lateral organs and vascular patterning (Turchi et al. 2013). WOX homeodomain transcription factors are also associated with apical/basal cell fate (Breuninger et al. 2008) and possibly regulate ATML1/PDF2 and other epidermal cell fate regulators. My analysis of *ACR4* promoter expression in embryo suggests that *ACR4* may be regulated by WOXs.

Post germination, WOX1, PRS, HD-ZIP III and HD-ZIP II transcription factors have been implicated in the regulation of leaf polarity (Emery et al. 2003; McConnell et al. 2001; Nakata et al. 2012; Prigge et al. 2005; Turchi et al. 2013; Vandenbussche et al. 2009). My analysis of genes coexpressed with *ATML1*, *PDF2* or *ACR4* together with the transcriptomic data of *pdf2* and *pPDF1-FLAGF:PDF2* expressing lines suggests that some transcription factors of the HD-ZIP I (HB7, HB12, and HB21) and HD-ZIP II (HAT3, HAT14) families are potential targets of ATML1 and PDF2. Interestingly the overexpression of transgenic *FLAG-ATML1* or *FLAG-PDF2* under the *PDF1* promoter results in hyponastic and curling leaf phenotypes which could be correlated with defects in leaf polarity caused by the disruption of HD-ZIP III proteins or WOX1/PRS. In addition, I showed that the *acr4* mutation exacerbates the phenotypes of *wox1prs* mutant. Together these results suggest that epidermal cell fate proteins are implicated in leaf polarity.

There are several non-mutually exclusive mechanisms that could underlie interactions between these classes of homeodomain transcription factors. The most obvious is the direct transcriptional regulation by HD-ZIP IV proteins of the genes encoding for HD-ZIP or WOX proteins, or vice versa, as has been recently demonstrated in the case of REV-mediated activation of *ATHB2* in embryo development (Turchi et al. 2013). Alternatively, common target genes could be differentially regulated by different classes of HD transcription factors. In this

scenario HD transcription factors could act cooperatively, as has been shown for HD-ZIP II and HD-ZIP III proteins in leaf polarity and embryo patterning. The analysis of *ACR4* and *ATML1* promoters ((Takada and Jurgens 2007) and this study) is consistent with the hypothesis that both HD-ZIP IV and WOX families of transcription factors are necessary for epidermal cell fate specification. HD transcription factors could also act antagonistically in target gene regulation. The binding sites of HD-ZIP transcription factors are pseudopalindromic and relatively similar. Thus HD-ZIP I and HD-ZIP II proteins bind the sequence *CAAT*N*ATTG* whereas HD-ZIP III bind to *GTAAT*(G/C)*ATTAC* and HD-ZIP IV bind to the L1 box TAAATGYA, *CATTAAATG* or *GCTTAAATG*C (Abe et al. 2001; Johannesson et al. 2001; Palena et al. 1999; Sessa et al. 1993; Sessa et al. 1998; Tron et al. 2001). To date, these binding sites have been assumed to be specific to each class of transcription factors and there has been no evidence to suggest the existence of heterodimers between different classes. Nevertheless, the similarity of binding site structure could in principle allow competition between classes of HD-ZIP transcription factors. In this scenario the lack of function or overexpression of HD-ZIP proteins will ultimately affect the function of different classes of HD-ZIP transcription factors. This could explain the similar leaf phenotypes of *HD-ZIP III* mutants, and transgenic lines overexpressing *FLAG-ATML1* or *FLAG-PDF2*.

In summary, the results presented in my thesis have provided some tantalizing clues regarding potential overlap between epidermal cell fate specification pathways and the activity of a range of HD proteins during embryogenesis. In order to clarify the situation it will be necessary to carry out an in depth study of the genetic interactions between HD-ZIP and WOX encoding genes during embryogenesis, and to analyse the expression levels of different *HD-ZIP* genes in multiple mutant backgrounds of each family (*atml1pdf2*, *phb phv rev*, *hat3 athb4 athb2* or *wox2 wox8*).

6.2 The impact of epidermis cell fate in auxin response

Auxin signalling is indissociable from embryo patterning. As mentioned in the introductory chapter, defects in auxin distribution, transport, synthesis or response pathways lead to defects in embryo patterning including root pole specification and cotyledon initiation (discussed in Section 1.2.2). From the late globular stage

onwards auxin distribution in the embryo is cyclic; on one side auxin flux is conducted from the tip of cotyledon precursors to the root basal pole via the internal tissues and then auxin is redirected to the cotyledon tips via the epidermal cell layer (Benkova et al. 2003; Friml et al. 2003). The main determinant of auxin distribution is the localization of the PIN transporters. During this study, my results regularly suggested that auxin-signalling and epidermal cell fate specification are linked. For instance, several AREs were found to be present in the *ACR4* promoter, and our deletion strategy indicates that ARE containing promoter regions might control the expression of *ACR4* in different zones of the epidermis during embryogenesis (Chapter 4). Several ARFs and IAAs are either coexpressed with *ATML1* and *PDF2* or are differentially expressed in transcriptomic data from the inflorescences of *pdf2-3* or *PDF2* overexpression plants. In particular, *PIN1* and *ARF18* are upregulated in inflorescences of *PDF2* overexpression lines, and their expression pattern at heart stage embryo is relatively specific to the epidermal layer (Friml et al. 2003; Rademacher et al. 2011) suggesting that these genes could be regulated by ATML1 or PDF2 via the L1 boxes present in their respective promoters. Such a regulation could reinforce the auxin distribution pattern in the embryo, especially at the late globular and heart stage of embryogenesis when auxin flux is directed via the L1 layer. Importantly, defects in auxin distribution at this stage in development have been shown to cause defects in cotyledon initiation and positioning; phenotypes which became a recurring theme during my analysis of mutants with defects in epidermal cell fate specification.

The double mutants between *atml1-3*, *pdf2-3* or *acr4-2* show embryo phenotypes that resemble *rpk1toad2* embryo phenotypes, including abnormal periclinal divisions in the epidermis and the patterning defects in the root pole of the embryo characteristic of auxin mutants. Interestingly, in addition to lacking epidermal markers, *rpk1toad2* embryos show misexpression of the DR5 reporter, PIN1 and other embryo patterning markers (Nodine and Tax 2008; Nodine et al. 2007), suggesting that an intact epidermal layer is important for embryo patterning and in particular for auxin distribution. To address this hypothesis, it will be important to analyse PIN1 distribution and the DR5 expression pattern during embryogenesis in *atml1-3 pdf2-3* double mutants.

The intricate relationship between auxin response and embryo patterning regulators probably involves overlapping and antagonistic relationships. For instance, the KAN and HD-ZIP III transcription factors are able to control auxin flux during embryo development (Izhaki and Bowman 2007). It is, for example possible that the disruption of epidermal pathways influences auxin response indirectly by affecting other regulatory programs. There may also be considerable feedback within the system. For example, in the same way that auxin flux through the developing vasculature is thought to be reinforced by the auxin-mediated differentiation of vascular tissues (Floyd and Bowman 2010; Sauer et al. 2006; Scarpella et al. 2006; Wabnik et al. 2011), it is possible that the differentiation of epidermal cells involves their acquiring specific physical characteristics which directly influence auxin flux. For example the organisation of the epidermis as a continuous monolayer of tightly adhering cells may facilitate auxin movement through the tissue, which could both increase flux and also feedback on epidermal specification.

6.3 The robustness of epidermal cell fate?

A major aim of my project was to model the epidermal cell-fate specification and maintenance. Due to the lack of information of signalling pathways or their impact on the transcriptional regulation of *ATML1* and *PDF2*, I focused on the suspected relationship between ACR4, ATML1 and PDF2 (Chapter 5). Before I started this thesis the accepted model in the literature suggested that ATML1 and PDF2 maintained their expression via a positive auto-regulatory feedback loop (Abe et al. 2003; Javelle et al. 2011b; Tanaka et al. 2007). It had also been suggested that the expression of *ACR4* was positively regulated by these factors (Abe et al. 2003). Unfortunately both of these positive regulatory interactions were based on transcriptional analysis in mutants largely lacking epidermal cells, making them difficult to interpret (Abe et al. 2003). Furthermore, unpublished results from the Ingram group suggested that ACR4 signalling presumably mediated by the binding of an unknown ligand activates the transcription of *ATML1* and *PDF2*. Based on the results obtained during this study I proposed an alternative model. I confirmed that the maintenance of expression levels of *ATML1* and *PDF2* requires ACR4, but I

proposed a negative feedback loop model in which the ATML1 and PDF2 proteins repress their own production and the transcription of *ACR4*.

The properties of both models were interesting in the context of the establishment of L1 identity during embryogenesis, but the negative feedback loop model is clearly a closer fit to reality as regards the maintenance of epidermal cell fate post germination due to its superior system robustness and stability. This negative feedback loop model is also consistent with the results of transcriptional analysis in seedlings and inflorescences of different mutant backgrounds and *PDF2* overexpression plants. At least in these tissues, ATML1 and PDF2 participate in a double negative feedback loop, since these transcription factors repress their own expression and that of *ACR4*. The simulations of perturbations in this model display a remarkable robustness and system steadiness.

Interestingly the introduction of perturbations into the positive feedback loop model induced bistable responses, that reflect a sensitivity to fluctuations in signal input which would be predicted to be problematic in the biological context of epidermal cell fate maintenance. The epidermal layer is of critical importance for plant survival, making its maintenance inconsistent with a model showing such vulnerabilities. On the other hand, bistable systems are a recurring theme in the biological mechanisms involved in making cell fate decisions during development (Stelling et al. 2004). During early stages of *Arabidopsis* embryogenesis the definition of the epidermis is accomplished by a specific round of cell divisions that separates "outside cell layer" cell fate and "inner cell layer" cell fate. Critically, inner cells rapidly lose the expression of epidermal markers after this cell division whereas they are maintained in outer embryonic cells. The properties of my positive feedback loop model would be useful in explaining how these two cell-fates are adopted. The generally accepted model of epidermal specification describes the existence of signals that maintain the transcription of the *ATML1* and *PDF2* epidermal regulators. The theory suggests that inner cells of the embryo are somehow cut off from this signal and therefore adopt the inner cell fate. Based on my results I would suggest that ACR4, and possibly other receptors are necessary for the perception of this signal. I would also propose that during early embryogenesis the relationship between epidermal cell fate

regulators resembles the positive feedback loop model, possibly due to a shift in the balance of the activities of combinations of HD-ZIP IV proteins acting together on the promoters of *ATML1*, *PDF2* and *ACR4*.

In order to develop the epidermal feedback loop models it will be of critical importance to identify the initial input signal, and to understand the dynamics of the activation of the ACR4 receptor kinase and downstream steps in the signalling cascade. Additionally the identification of ATML1 and PDF2 targets could help to unveil the pathways required for epidermal specification and maintenance to complete the model. Unfortunately it is not easy to quantify the parameters of the model, especially during early embryogenesis when my work suggests they may be particularly interesting. The embryo is sheltered by several tissues and it is difficult to access with accuracy without disturbing the system. One long-term solution could be the study of somatic embryogenesis, to simplify the system.

6.4 Signalling in epidermal cell fate.

6.4.1 So many RLKs, so few answers!

At least six receptor kinases have been implicated in epidermal cell fate specification. The analysis of the mutant phenotypes and expression patterns of ACR4, ALE2, GSO1, and GSO2 indicates that these receptors may be involved in the determination of epidermal identity (Gifford et al. 2003; Tanaka et al. 2007; Tsuwamoto et al. 2008; Xing et al. 2013). Additionally RPK1 and TOAD2 are involved in radial patterning and the double mutant displays severe defects in embryo patterning including epidermal cell fate (Nodine and Tax 2008; Nodine et al. 2007). Frustratingly, the ligands and targets of these proteins remain largely unknown. In particular the origin of signals involved in specifying and maintaining epidermal cell fate is intriguing. They could be derived from the ESR (endosperm) or could be produced by the epidermal cells themselves and secreted to the cell wall or cuticle layer of the epidermal surface to be detected by neighbouring cells. During early embryogenesis both possibilities might be correct, although obviously, in the context of epidermal maintenance post germination the second possibility is more plausible, and is supported by the fact that ACR4 receptors are preferentially

localized between epidermal cells (Gifford et al. 2003; Gifford et al. 2005; Watanabe et al. 2004).

Recent results from our laboratory together with available data in the literature suggest that the situation is rather complex. A seed specific pathway involving transcriptional activation of the subtilisin-like protease encoding *ALE1* gene in the ESR, is involved in signalling via GSO1 and GSO2, which are expressed in the developing embryo (Tsuwamoto et al. 2008; Xing et al. 2013; Yang et al. 2008). This signalling pathway is necessary for the formation of the embryonic surface (deposition of a functional cuticle), but is not actually required for the specification of epidermal cell fate *per se*. Consistent with this, it acts independently from ATML1 and PDF2. Strikingly *gso1gso2pdf2-3* and *gso1gso2acr4-2* triple mutants show early embryo lethal phenotypes and severe developmental problems, and a significant proportion of *ale1pdf2-3* double mutants also arrest at early stages of embryogenesis again supporting the scenario that the ALE1/GSO1/GSO2 pathway acts in parallel to the ACR4/ATML1/PDF2 feedback loop. Thus the formation of an intact cuticle has a profound influence on epidermal cell fate in backgrounds which would otherwise only show extremely mild developmental defects. A possible explanation for these observations is based on the secretion of the epidermis-maintaining signals by epidermal cells, into the apoplast. In such a scenario the embryonic cuticle could have an important role permitting to build-up these signals, by preventing their leakage into the endosperm. Thus defects in cuticle integrity will have an impact on the stabilization of the epidermal cell fate feedback loop involving ACR4 signalling and ATML1/PDF2 regulation. According to our positive feedback loop model, this might lead to a catastrophic and permanent loss of epidermal cell fate during early embryogenesis leading to embryo lethal phenotypes.

Although seductive this model requires further testing. It would, for example, be interesting to test the effects of mutations in cuticle biosynthetic enzymes in the *pdf2* or *acr4* mutant backgrounds. Furthermore, identification of the ligand of ACR4 would be a critical next step in understanding this pathway further. An intriguing possibility is that this ligand is amongst the secreted peptides encoded by genes

positively regulated by PDF2. This possibility is being explored within the laboratory.

This hypothesis also highlights the question of whether ATML1 and PDF2 themselves directly regulate cuticle biosynthesis. Based on the results of my work, the proposed link between HD-ZIP IV proteins in *Arabidopsis* and cuticle integrity is supported much more weakly than expected. Although the seedlings of several *HD-ZIP IV* double mutants show a significant increase in epidermal permeability to toluidine blue, the levels are to a great extent lower than the levels observed in knockout lines of mutants in the ZOU, ALE1, GSO1, GSO2 pathway. These phenotypes may be due to epidermal discontinuity rather than defective cuticle production. In addition, although some genes involved in cuticle biosynthesis are coexpressed with *ATML1*, *PDF2* or *ACR4,* and the "lipid metabolism" GO term is overrepresented in those lists of genes, the functional analysis our transcriptomic data shows no correlation between potential targets of PDF2 and cuticle biosynthesis. These results are inconsistent with the results of *OCL1* overexpression in maize that identified genes involved in lipid metabolism and cuticle biosynthesis as potential targets of OCL1 (Javelle et al. 2011a).

6.4.2 Overlap between stress signalling and epidermal cell fate

A surprising result of my study of the potential targets of ATML1 and PDF2 is that epidermal cell fate is strongly associated with stress responses. Interestingly several HD-ZIP I and II transcription factors, together with *HDG11,* are associated with different abiotic stress responses (Cao et al. 2009; Harris et al. 2011). Harris and colleagues suggest that HD-ZIP I and II transcription factors are important players in the regulatory network that merges developmental and environmental stimuli to activate the appropriate cell-specific responses for adaptation (Harris et al. 2011). The results of this project suggest that HD-ZIP IV transcription factors may also belong to this regulatory network. The fact that these genes are globally epidermis specific is important since the epidermal layer is the first to perceive many abiotic and biotic stresses. To date the role of HD-ZIP IV genes in stress responses has not been widely analysed. A first step might be to subject the available single and

multiple mutants to a “gauntlet” to assess potential differences in response at the phenotypic level.

When considering the epidermis with respect to stress, it is important to remember that at the biophysical level the epidermis of embryos, meristems and young organ primordial is thought to be a layer of continuously physically stressed cells (Hamant et al. 2008; Javelle et al. 2011b; Kutschera and Niklas 2007; Landrein and Hamant 2013). The epidermis has a profound impact on the organization of inner cell layers, the growth of which in turn provides a constant pressure against the epidermal layer during plant development (Javelle et al. 2011b; Kutschera and Niklas 2007). The epidermis appears to constrain the expansion of underlying tissues, thus maintaining this tension, partially by the formation of a specialized and thickened outer cell wall. The division of epidermal cells and their oriented and regulated expansion is arguably the single most important element in conferring form to developing plant tissues. Because it is under continuous tension, it is possible that the epidermis constantly expresses genes usually associated with stress responses. This could explain some of the data in my analysis. It would be very interesting to look at the tissue specific expression of some of these “stress associated” genes by *in situ* hybridization.

An intriguing possibility, currently being investigated in the RDP (Lyon), is that the epidermis is, in fact, actually defined by its unique phenotypic response to physical stress (in the form of tension generated from the growth of underlying cells). Taking this hypothesis to its logical end would imply that this unique response is mediated by the activities of ATML1, PDF2 and other HD-ZIP IV transcription factors. If this hypothesis is correct it is entirely possible that the HD-ZIP IV transcription factors were a key element is conferring organized forms to multicellular plants.

Chapter 7 – Appendix

ACR4 promoter expression pattern during embryogenesis

In this section I include the results of the *ACR4* promoter expression analysis. I analysed the different *pACR4-GFP:GUS* constructs (P1-P6, m1 and m2 described in Figure A1). In this section, I show the expression profiles of several constructs analysing the GUS expression profiles at heart stage (Figure A1) and GFP signal at globular stage (Figure A2). The expression patterns are described in Section 5.2.3.

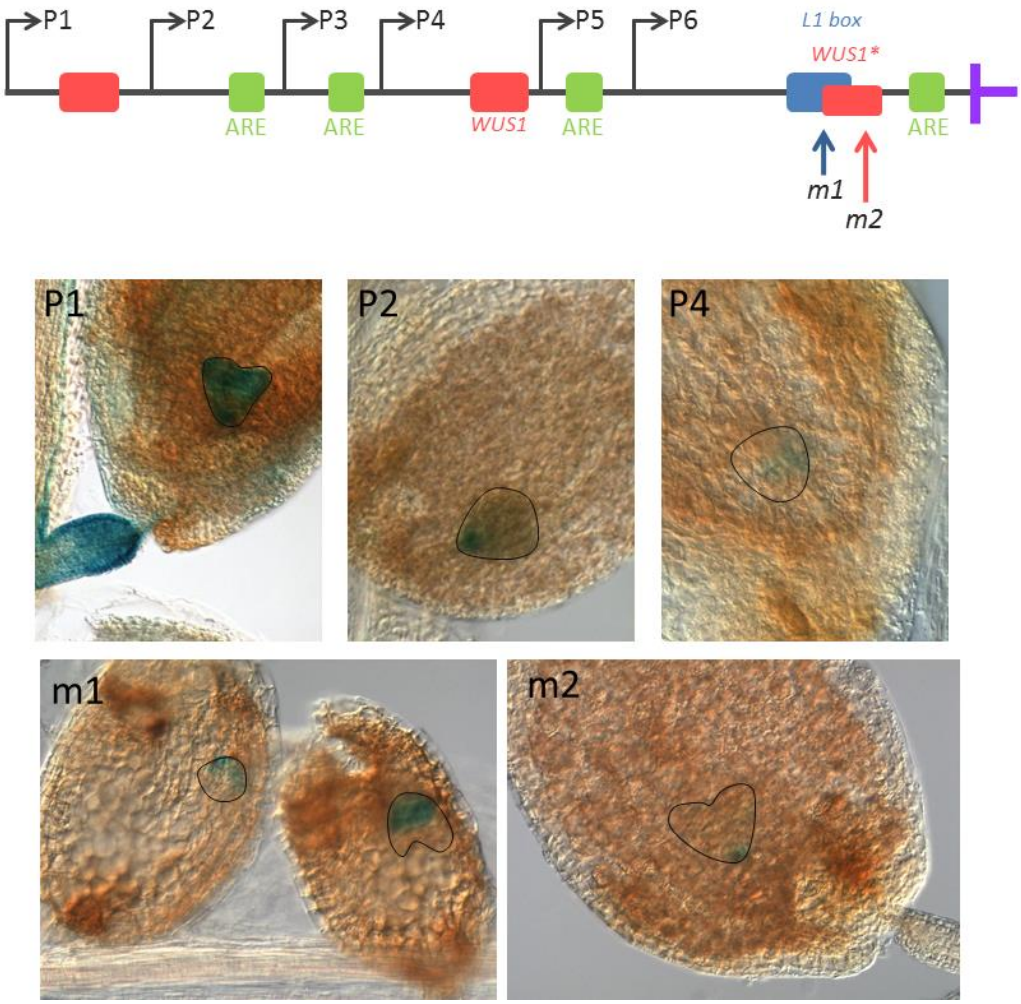


Figure A1 – Expression pattern of *ACR4* promoter during seed development. GUS staining experiments in seeds expressing different versions of *ACR4* promoter. The full length promoter (P1) is expressed in the whole surface of the embryo while P2, P4, m1 and m2 promoters have different expression patterns.

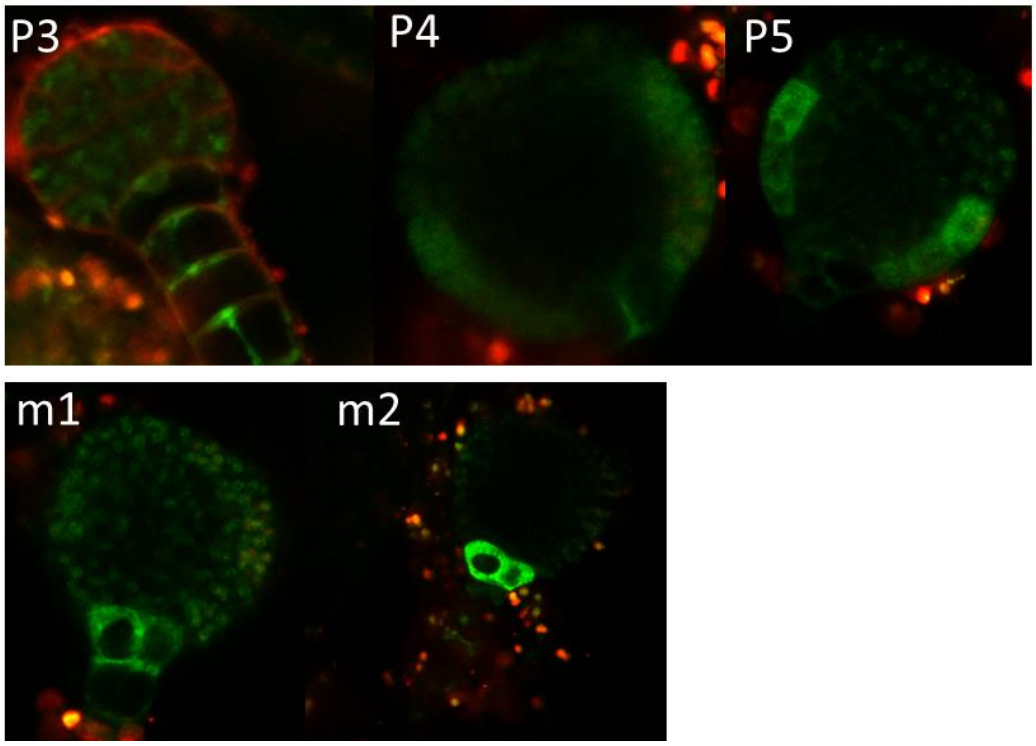


Figure A2 –*ACR4* promoter expression at the globular stage of embryogenesis. Embryos expressing different versions of the *ACR4* promoter display different GFP expression patterns.

Chapter 8 – Bibliography

Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997) Role of arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. Plant Cell 9: 1859-1868

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

Abe M, Takahashi T, Komeda Y (1999) Cloning and characterization of an L1 layer-specific gene in Arabidopsis thaliana. Plant Cell Physiol 40: 571-580

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

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

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

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

Alpy F, Tomasetto C (2005) Give lipids a START: the StAR-related lipid transfer (START) domain in mammals. J Cell Sci 118: 2791-2801

Alves MS, Reis PA, Dadalto SP, Faria JA, Fontes EP, Fietto LG (2011) A novel transcription factor, ERD15 (Early Responsive to Dehydration 15), connects endoplasmic reticulum stress with an osmotic stress-induced cell death signal. J Biol Chem 286: 20020-20030

Ariel FD, Manavella PA, Dezar CA, Chan RL (2007) The true story of the HD-Zip family. Trends Plant Sci 12: 419-426

Bach L, Michaelson LV, Haslam R, Bellec Y, Gissot L, Marion J, Da Costa M, Boutin JP, Miquel M, Tellier F, Domergue F, Markham JE, Beaudoin F, Napier JA, Faure JD (2008) The very-long-chain hydroxy fatty acyl-CoA dehydratase PASTICCINO2 is essential and limiting for plant development. Proc Natl Acad Sci U S A 105: 14727-14731

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

Beaudoin F, Wu X, Li F, Haslam RP, Markham JE, Zheng H, Napier JA, Kunst L (2009) Functional characterization of the Arabidopsis beta-ketoacyl-coenzyme A reductase candidates of the fatty acid elongase. Plant Physiol 150: 1174-1191

Becraft PW, Kang SH, Suh SG (2001) The maize CRINKLY4 receptor kinase controls a cell-autonomous differentiation response. Plant Physiol 127: 486-496

Becraft PW, Li K, Dey N, Asuncion-Crabb Y (2002) The maize dek1 gene functions in embryonic pattern formation and cell fate specification. Development 129: 5217-5225

Becraft PW, Stinard PS, McCarty DR (1996) CRINKLY4: A TNFR-like receptor kinase involved in maize epidermal differentiation. Science 273: 1406-1409

Beeckman T, De Rycke R, Viane R, Inze D (2000) Histological study of seed coat development in Arabidopsis thaliana. J Plant Res 113: 139-148

Beisson F, Li-Beisson Y, Pollard M (2012) Solving the puzzles of cutin and suberin polymer biosynthesis. Curr Opin Plant Biol 15: 329-337

Benjamins R, Ampudia CS, Hooykaas PJ, Offringa R (2003) PINOID-mediated signaling involves calcium-binding proteins. Plant Physiol 132: 1623-1630

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

Benjamins R, Scheres B (2008) Auxin: the looping star in plant development. Annu Rev Plant Biol 59: 443-465

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

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

Berger F (2003) Endosperm: the crossroad of seed development. Curr Opin Plant Biol 6: 42-50

Berger F, Grini PE, Schnittger A (2006) Endosperm: an integrator of seed growth and development. Curr Opin Plant Biol 9: 664-670

Bergmann DC, Lukowitz W, Somerville CR (2004) Stomatal development and pattern controlled by a MAPKK kinase. Science 304: 1494-1497

Bergmann DC, Sack FD (2007) Stomatal development. Annu Rev Plant Biol 58: 163-181

Berleth T, Jurgens G (1993) The Role of the Monopteros Gene in Organizing the Basal Body Region of the Arabidopsis Embryo. Development 118: 575-587

Bernard A, Joubes J (2013) Arabidopsis cuticular waxes: advances in synthesis, export and regulation. Prog Lipid Res 52: 110-129

Besnard F, Vernoux T, Hamant O (2011) Organogenesis from stem cells in planta: multiple feedback loops integrating molecular and mechanical signals. Cell Mol Life Sci 68: 2885-2906

Bethke PC, Libourel IG, Aoyama N, Chung YY, Still DW, Jones RL (2007) The Arabidopsis aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. Plant Physiol 143: 1173-1188

Betsuyaku S, Sawa S, Yamada M (2011) The Function of the CLE Peptides in Plant Development and Plant-Microbe Interactions. Arabidopsis Book 9: e0149

Bird SM, Gray JE (2003) Signals from the cuticle affect epidermal cell differentiation. New Phytologist 157: 9-23

Block A, Dangl JL, Hahlbrock K, Schulze-Lefert P (1990) Functional borders, genetic fine structure, and distance requirements of cis elements mediating light responsiveness of the parsley chalcone synthase promoter. Proc Natl Acad Sci U S A 87: 5387-5391

Bonello JF, Sevilla-Lecoq S, Berne A, Risueno MC, Dumas C, Rogowsky PM (2002) Esr proteins are secreted by the cells of the embryo surrounding region. J Exp Bot 53: 1559-1568

Borghi L, Gutzat R, Futterer J, Laizet Y, Hennig L, Gruissem W (2010) Arabidopsis RETINOBLASTOMA-RELATED is required for stem cell maintenance, cell differentiation, and lateral organ production. Plant Cell 22: 1792-1811

Boudolf V, Vlieghe K, Beemster GT, Magyar Z, Torres Acosta JA, Maes S, Van Der Schueren E, Inze D, De Veylder L (2004) The plant-specific cyclin-dependent kinase CDKB1;1 and transcription factor E2Fa-DPa control the balance of mitotically dividing and endoreduplicating cells in Arabidopsis. Plant Cell 16: 2683-2692

Bouyer D, Geier F, Kragler F, Schnittger A, Pesch M, Wester K, Balkunde R, Timmer J, Fleck C, Hulskamp M (2008) Two-dimensional patterning by a trapping/depletion mechanism: the role of TTG1 and GL3 in Arabidopsis trichome formation. PLoS Biol 6: e141

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

Bulow L, Engelmann S, Schindler M, Hehl R (2009) AthaMap, integrating transcriptional and post-transcriptional data. Nucleic Acids Res 37: D983-986

Busch W, Miotk A, Ariel FD, Zhao Z, Forner J, Daum G, Suzaki T, Schuster C, Schultheiss SJ, Leibfried A, Haubeiss S, Ha N, Chan RL, Lohmann JU (2010) Transcriptional control of a plant stem cell niche. Dev Cell 18: 849-861

Byrne ME, Barley R, Curtis M, Arroyo JM, Dunham M, Hudson A, Martienssen RA (2000) Asymmetric leaves1 mediates leaf patterning and stem cell function in Arabidopsis. Nature 408: 967-971

Byrne ME, Simorowski J, Martienssen RA (2002) ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis. Development 129: 1957-1965

Cao YJ, Wei Q, Liao Y, Song HL, Li X, Xiang CB, Kuai BK (2009) Ectopic overexpression of AtHDG11 in tall fescue resulted in enhanced tolerance to drought and salt stress. Plant Cell Rep 28: 579-588

Casson S, Gray JE (2008) Influence of environmental factors on stomatal development. New Phytol 178: 9-23

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

Chen W, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou G, Whitham SA, Budworth PR, Tao Y, Xie Z, Chen X, Lam S, Kreps JA, Harper JF, Si-Ammour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn T, Dangl JL, Wang X, Zhu T (2002) Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell 14: 559-574

Chen X, Goodwin SM, Boroff VL, Liu X, Jenks MA (2003) Cloning and characterization of the WAX2 gene of Arabidopsis involved in cuticle membrane and wax production. Plant Cell 15: 1170-1185

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

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

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

Chitwood DH, Nogueira FT, Howell MD, Montgomery TA, Carrington JC, Timmermans MC (2009) Pattern formation via small RNA mobility. Genes Dev 23: 549-554

Choi H, Hong J, Ha J, Kang J, Kim SY (2000) ABFs, a family of ABA-responsive element binding factors. J Biol Chem 275: 1723-1730

Ciarbelli AR, Ciolfi A, Salvucci S, Ruzza V, Possenti M, Carabelli M, Fruscalzo A, Sessa G, Morelli G, Ruberti I (2008) The Arabidopsis homeodomain-leucine zipper II gene family: diversity and redundancy. Plant Mol Biol 68: 465-478

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16: 735-743

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

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26: 1367-1372

de Reuille PB, Bohn-Courseau I, Ljung K, Morin H, Carraro N, Godin C, Traas J (2006) Computer simulations reveal properties of the cell-cell signaling network at the shoot apex in Arabidopsis. Proc Natl Acad Sci U S A 103: 1627-1632

De Rybel B, Moller B, Yoshida S, Grabowicz I, Barbier de Reuille P, Boeren S, Smith RS, Borst JW, Weijers D (2013) A bHLH complex controls embryonic vascular tissue establishment and indeterminate growth in Arabidopsis. Dev Cell 24: 426-437

De Smet I, Vassileva V, De Rybel B, Levesque MP, Grunewald W, Van Damme D, Van Noorden G, Naudts M, Van Isterdael G, De Clercq R, Wang JY, Meuli N, Vanneste S, Friml J, Hilson P, Jurgens G, Ingram GC, Inze D, Benfey PN, Beeckman T (2008) Receptor-like kinase ACR4 restricts formative cell divisions in the Arabidopsis root. Science 322: 594-597

Debono A, Yeats TH, Rose JK, Bird D, Jetter R, Kunst L, Samuels L (2009) Arabidopsis LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. Plant Cell 21: 1230-1238

Desvoyes B, Ramirez-Parra E, Xie Q, Chua NH, Gutierrez C (2006) Cell type-specific role of the retinoblastoma/E2F pathway during Arabidopsis leaf development. Plant Physiol 140: 67-80

Dettmer J, Friml J (2011) Cell polarity in plants: when two do the same, it is not the same. Curr Opin Cell Biol 23: 686-696

Deveaux Y, Toffano-Nioche C, Claisse G, Thareau V, Morin H, Laufs P, Moreau H, Kreis M, Lecharny A (2008) Genes of the most conserved WOX clade in plants affect root and flower development in Arabidopsis. BMC Evol Biol 8: 291

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

Dharmasiri N, Dharmasiri S, Weijers D, Karunarathna N, Jurgens G, Estelle M (2007) AXL and AXR1 have redundant functions in RUB conjugation and growth and development in Arabidopsis. Plant J 52: 114-123

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

Dharmasiri S, Dharmasiri N, Hellmann H, Estelle M (2003) The RUB/Nedd8 conjugation pathway is required for early development in Arabidopsis. Embo J 22: 1762-1770

Dhonukshe P, Aniento F, Hwang I, Robinson DG, Mravec J, Stierhof YD, Friml J (2007) Clathrin-mediated constitutive endocytosis of PIN auxin efflux carriers in Arabidopsis. Curr Biol 17: 520-527

Dhonukshe P, Tanaka H, Goh T, Ebine K, Mahonen AP, Prasad K, Blilou I, Geldner N, Xu J, Uemura T, Chory J, Ueda T, Nakano A, Scheres B, Friml J (2008) Generation of cell polarity in plants links endocytosis, auxin distribution and cell fate decisions. Nature 456: 962-966

Di Cristina M, Sessa G, Dolan L, Linstead P, Baima S, Ruberti I, Morelli G (1996) The Arabidopsis Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. Plant J 10: 393-402

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

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. Trends Plant Sci 15: 573-581

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res 19: 1349

Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL (2003) Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. Curr Biol 13: 1768-1774

Eshed Y, Baum SF, Perea JV, Bowman JL (2001) Establishment of polarity in lateral organs of plants. Curr Biol 11: 1251-1260

Eshed Y, Izhaki A, Baum SF, Floyd SK, Bowman JL (2004) Asymmetric leaf development and blade expansion in Arabidopsis are mediated by KANADI and YABBY activities. Development 131: 2997-3006

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends Plant Sci 5: 199-206

Ezcurra I, Ellerstrom M, Wycliffe P, Stalberg K, Rask L (1999) Interaction between composite elements in the napA promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. Plant Mol Biol 40: 699-709

Fiume E, Fletcher JC (2012) Regulation of Arabidopsis embryo and endosperm development by the polypeptide signaling molecule CLE8. Plant Cell 24: 1000-1012

Floyd SK, Bowman JL (2010) Gene expression patterns in seed plant shoot meristems and leaves: homoplasy or homology? J Plant Res 123: 43-55

Fobis-Loisy I, Chambrier P, Gaude T (2007) Genetic transformation of Arabidopsis lyrata: specific expression of the green fluorescent protein (GFP) in pistil tissues. Plant Cell Rep 26: 745-753

Folkers U, Berger J, Hulskamp M (1997) Cell morphogenesis of trichomes in Arabidopsis: differential control of primary and secondary branching by branch initiation regulators and cell growth. Development 124: 3779-3786

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

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

Friml J, Yang X, Michniewicz M, Weijers D, Quint A, Tietz O, Benjamins R, Ouwerkerk PB, Ljung K, Sandberg G, Hooykaas PJ, Palme K, Offringa R (2004) A

PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. Science 306: 862-865

Fujita Y, Fujita M, Shinozaki K, Yamaguchi-Shinozaki K (2011) ABA-mediated transcriptional regulation in response to osmotic stress in plants. J Plant Res 124: 509-525

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

Geisler M, Nadeau J, Sack FD (2000) Oriented asymmetric divisions that generate the stomatal spacing pattern in arabidopsis are disrupted by the too many mouths mutation. Plant Cell 12: 2075-2086

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

Gendrel AV, Lippman Z, Martienssen R, Colot V (2005) Profiling histone modification patterns in plants using genomic tiling microarrays. Nat Methods 2: 213-218

Gifford ML, Dean S, Ingram GC (2003) The Arabidopsis ACR4 gene plays a role in cell layer organisation during ovule integument and sepal margin development. Development 130: 4249-4258

Gifford ML, Robertson FC, Soares DC, Ingram GC (2005) ARABIDOPSIS CRINKLY4 function, internalization, and turnover are dependent on the extracellular crinkly repeat domain. Plant Cell 17: 1154-1166

Gilbertson L (2003) Cre-lox recombination: Cre-ative tools for plant biotechnology. Trends Biotechnol 21: 550-555

Gleave AP (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. Plant Mol Biol 20: 1203-1207

Glover BJ (2000) Differentiation in plant epidermal cells. J Exp Bot 51: 497-505

Gordon SP, Chickarmane VS, Ohno C, Meyerowitz EM (2009) Multiple feedback loops through cytokinin signaling control stem cell number within the Arabidopsis shoot meristem. Proc Natl Acad Sci U S A 106: 16529-16534

Gorton HL, Vogelmann TC (1996) Effects of Epidermal Cell Shape and Pigmentation on Optical Properties of Antirrhinum Petals at Visible and Ultraviolet Wavelengths. Plant Physiol 112: 879-888

Gray WM, Hellmann H, Dharmasiri S, Estelle M (2002) Role of the Arabidopsis RING-H2 protein RBX1 in RUB modification and SCF function. Plant Cell 14: 2137-2144

Grebe M (2012) The patterning of epidermal hairs in Arabidopsis--updated. Curr Opin Plant Biol 15: 31-37

Green KA, Prigge MJ, Katzman RB, Clark SE (2005) CORONA, a member of the class III homeodomain leucine zipper gene family in Arabidopsis, regulates stem cell specification and organogenesis. Plant Cell 17: 691-704

Greer S, Wen M, Bird D, Wu X, Samuels L, Kunst L, Jetter R (2007) The cytochrome P450 enzyme CYP96A15 is the midchain alkane hydroxylase responsible for formation of secondary alcohols and ketones in stem cuticular wax of Arabidopsis. Plant Physiol 145: 653-667

Grunewald W, Friml J (2010) The march of the PINs: developmental plasticity by dynamic polar targeting in plant cells. Embo J 29: 2700-2714

Guimil S, Dunand C (2007) Cell growth and differentiation in Arabidopsis epidermal cells. J Exp Bot 58: 3829-3840

Gutzat R, Borghi L, Gruissem W (2012) Emerging roles of RETINOBLASTOMA-RELATED proteins in evolution and plant development. Trends Plant Sci 17: 139-148

Ha CM, Jun JH, Fletcher JC (2010) Shoot apical meristem form and function. Curr Top Dev Biol 91: 103-140

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

Hamann T, Benkova E, Baurle I, Kientz M, Jurgens G (2002) The Arabidopsis BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. Genes Dev 16: 1610-1615

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

Hamant O, Heisler MG, Jonsson H, Krupinski P, Uyttewaal M, Bokov P, Corson F, Sahlin P, Boudaoud A, Meyerowitz EM, Couder Y, Traas J (2008) Developmental patterning by mechanical signals in Arabidopsis. Science 322: 1650-1655

Hara K, Kajita R, Torii KU, Bergmann DC, Kakimoto T (2007) The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. Genes Dev 21: 1720-1725

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

Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science 290: 2110-2113

Harris JC, Hrmova M, Lopato S, Langridge P (2011) Modulation of plant growth by HD-Zip class I and II transcription factors in response to environmental stimuli. New Phytol 190: 823-837

Heim R, Cubitt AB, Tsien RY (1995) Improved green fluorescence. Nature 373: 663-664

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

Henriksson E, Olsson AS, Johannesson H, Johansson H, Hanson J, Engstrom P, Soderman E (2005) Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships. Plant Physiol 139: 509-518

Hibara K, Karim MR, Takada S, Taoka K, Furutani M, Aida M, Tasaka M (2006) Arabidopsis CUP-SHAPED COTYLEDON3 regulates postembryonic shoot meristem and organ boundary formation. Plant Cell 18: 2946-2957

Higashiyama T (2010) Peptide Signaling in PollenPistil Interactions. Plant and Cell Physiology 51: 177-189

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27: 297-300

Hill AV (1913) The Combinations of Haemoglobin with Oxygen and with Carbon Monoxide. I. Biochem J 7: 471-480

Himmelbach A, Hoffmann T, Leube M, Hohener B, Grill E (2002) Homeodomain protein ATHB6 is a target of the protein phosphatase ABI1 and regulates hormone responses in Arabidopsis. Embo J 21: 3029-3038

Hobbie L, McGovern M, Hurwitz LR, Pierro A, Liu NY, Bandyopadhyay A, Estelle M (2000) The axr6 mutants of Arabidopsis thaliana define a gene involved in auxin response and early development. Development 127: 23-32

Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. Adv Bioinformatics 2008: 420747

Hudson ME, Quail PH (2003) Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. Plant Physiol 133: 1605-1616

Hulskamp M (2004) Plant trichomes: a model for cell differentiation. Nat Rev Mol Cell Biol 5: 471-480

Hulskamp M, Misra S, Jurgens G (1994) Genetic dissection of trichome cell development in Arabidopsis. Cell 76: 555-566

Hunt L, Bailey KJ, Gray JE (2010) The signalling peptide EPFL9 is a positive regulator of stomatal development. New Phytol 186: 609-614

Hunt L, Gray JE (2009) The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. Curr Biol 19: 864-869

Ingram GC (2004) Between the sheets: inter-cell-layer communication in plant development. Philos Trans R Soc Lond B Biol Sci 359: 891-906

Ingram GC (2007) Signalling during epidermal development. Biochem Soc Trans 35: 156-160

Ingram GC (2010) Family life at close quarters: communication and constraint in angiosperm seed development. Protoplasma 247: 195-214

Ingram GC, Boisnard-Lorig C, Dumas C, Rogowsky PM (2000) Expression patterns of genes encoding HD-ZipIV homeo domain proteins define specific domains in maize embryos and meristems. Plant J 22: 401-414

Ingram GC, Magnard JL, Vergne P, Dumas C, Rogowsky PM (1999) ZmOCL1, an HDGL2 family homeobox gene, is expressed in the outer cell layer throughout maize development. Plant Mol Biol 40: 343-354

Ito J, Sono T, Tasaka M, Furutani M (2011) MACCHI-BOU 2 is required for early embryo patterning and cotyledon organogenesis in Arabidopsis. Plant Cell Physiol 52: 539-552

Iwakawa H, Iwasaki M, Kojima S, Ueno Y, Soma T, Tanaka H, Semiarti E, Machida Y, Machida C (2007) Expression of the ASYMMETRIC LEAVES2 gene in the adaxial domain of Arabidopsis leaves represses cell proliferation in this domain and is critical for the development of properly expanded leaves. Plant J 51: 173-184

Izhaki A, Bowman JL (2007) KANADI and class III HD-Zip gene families regulate embryo patterning and modulate auxin flow during embryogenesis in Arabidopsis. Plant Cell 19: 495-508

Jakoby M, Schnittger A (2004) Cell cycle and differentiation. Curr Opin Plant Biol 7: 661-669

Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F (2002) bZIP transcription factors in Arabidopsis. Trends Plant Sci 7: 106-111

Javelle M, Klein-Cosson C, Vernoud V, Boltz V, Maher C, Timmermans M, Depege-Fargeix N, Rogowsky PM (2011a) Genome-wide characterization of the HD-ZIP IV transcription factor family in maize: preferential expression in the epidermis. Plant Physiol 157: 790-803

Javelle M, Vernoud V, Depege-Fargeix N, Arnould C, Oursel D, Domergue F, Sarda X, Rogowsky PM (2010) Overexpression of the epidermis-specific homeodomain-leucine zipper IV transcription factor Outer Cell Layer1 in maize identifies target genes involved in lipid metabolism and cuticle biosynthesis. Plant Physiol 154: 273-286

Javelle M, Vernoud V, Rogowsky PM, Ingram GC (2011b) Epidermis: the formation and functions of a fundamental plant tissue. New Phytol 189: 17-39

Jenik PD, Gillmor CS, Lukowitz W (2007) Embryonic patterning in Arabidopsis thaliana. Annu Rev Cell Dev Biol 23: 207-236

Jeong S, Palmer TM, Lukowitz W (2011) The RWP-RK factor GROUNDED promotes embryonic polarity by facilitating YODA MAP kinase signaling. Curr Biol 21: 1268-1276

Jin P, Guo T, Becraft PW (2000) The maize CR4 receptor-like kinase mediates a growth factor-like differentiation response. Genesis 27: 104-116

Johannesson H, Wang Y, Engstrom P (2001) DNA-binding and dimerization preferences of Arabidopsis homeodomain-leucine zipper transcription factors in vitro. Plant Mol Biol 45: 63-73

Johnson KL, Degnan KA, Ross Walker J, Ingram GC (2005) AtDEK1 is essential for specification of embryonic epidermal cell fate. Plant J 44: 114-127

Johnson KL, Faulkner C, Jeffree CE, Ingram GC (2008) The phytocalpain defective kernel 1 is a novel Arabidopsis growth regulator whose activity is regulated by proteolytic processing. Plant Cell 20: 2619-2630

Kamata N, Okada H, Komeda Y, Takahashi T (2013) Mutations in epidermis-specific HD-ZIP IV genes affect floral organ identity in Arabidopsis thaliana. Plant J

Kanaoka MM, Pillitteri LJ, Fujii H, Yoshida Y, Bogenschutz NL, Takabayashi J, Zhu JK, Torii KU (2008) SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to arabidopsis stomatal differentiation. Plant Cell 20: 1775-1785

Karimi M, Inze D, Depicker A (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci 7: 193-195

Katari MS, Nowicki SD, Aceituno FF, Nero D, Kelfer J, Thompson LP, Cabello JM, Davidson RS, Goldberg AP, Shasha DE, Coruzzi GM, Gutierrez RA (2010) VirtualPlant: a software platform to support systems biology research. Plant Physiol 152: 500-515

Kawashima T, Goldberg RB (2010) The suspensor: not just suspending the embryo. Trends Plant Sci 15: 23-30

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

Kidner CA, Timmermans MC (2010) Signaling sides adaxial-abaxial patterning in leaves. Curr Top Dev Biol 91: 141-168

Kim SY, Chung HJ, Thomas TL (1997) Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. Plant J 11: 1237-1251

Kinoshita A, Betsuyaku S, Osakabe Y, Mizuno S, Nagawa S, Stahl Y, Simon R, Yamaguchi-Shinozaki K, Fukuda H, Sawa S (2010) RPK2 is an essential receptor-like kinase that transmits the CLV3 signal in Arabidopsis. Development 137: 3911-3920

Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, Fischer RL, Kakutani T (2004) One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. Science 303: 521-523

Kirik V, Simon M, Huelskamp M, Schiefelbein J (2004) The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis. Dev Biol 268: 506-513

Kiyohara S, Sawa S (2012) CLE signaling systems during plant development and nematode infection. Plant Cell Physiol 53: 1989-1999

Kleine-Vehn J, Dhonukshe P, Sauer M, Brewer PB, Wisniewska J, Paciorek T, Benkova E, Friml J (2008) ARF GEF-dependent transcytosis and polar delivery of PIN auxin carriers in Arabidopsis. Curr Biol 18: 526-531

Kubo H, Hayashi K (2011) Characterization of root cells of anl2 mutant in Arabidopsis thaliana. Plant Sci 180: 679-685

Kubo H, Peeters AJ, Aarts MG, Pereira A, Koornneef M (1999) ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. Plant Cell 11: 1217-1226

Kunst L, Samuels L (2009) Plant cuticles shine: advances in wax biosynthesis and export. Curr Opin Plant Biol 12: 721-727

Kuo MH, Allis CD (1999) In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. Methods 19: 425-433

Kutschera U, Niklas KJ (2007) The epidermal-growth-control theory of stem elongation: an old and a new perspective. J Plant Physiol 164: 1395-1409

Kwiatkowska D, Dumais J (2003) Growth and morphogenesis at the vegetative shoot apex of Anagallis arvensis L. J Exp Bot 54: 1585-1595

Lai LB, Nadeau JA, Lucas J, Lee EK, Nakagawa T, Zhao L, Geisler M, Sack FD (2005) The Arabidopsis R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. Plant Cell 17: 2754-2767

Lake JA, Woodward FI (2008) Response of stomatal numbers to CO2 and humidity: control by transpiration rate and abscisic acid. New Phytol 179: 397-404

Landrein B, Hamant O (2013) How mechanical stress controls microtubule behavior and morphogenesis in plants: history, experiments and revisited theories. Plant J

Lau S, Ehrismann JS, Schlereth A, Takada S, Mayer U, Jurgens G (2010) Cell-cell communication in Arabidopsis early embryogenesis. Eur J Cell Biol 89: 225-230

Lau S, Slane D, Herud O, Kong J, Jurgens G (2012) Early embryogenesis in flowering plants: setting up the basic body pattern. Annu Rev Plant Biol 63: 483-506

Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Horvath S, Drews GN, Fischer RL, Okamuro JK, Harada JJ, Goldberg RB (2010) Global analysis of gene activity during Arabidopsis seed

development and identification of seed-specific transcription factors. Proc Natl Acad Sci U S A 107: 8063-8070

Lenhard M, Jurgens G, Laux T (2002) The WUSCHEL and SHOOTMERISTEMLESS genes fulfil complementary roles in Arabidopsis shoot meristem regulation. Development 129: 3195-3206

Lenhard M, Laux T (2003) Stem cell homeostasis in the Arabidopsis shoot meristem is regulated by intercellular movement of CLAVATA3 and its sequestration by CLAVATA1. Development 130: 3163-3173

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol 148: 97-107

Li QJ, Xu B, Chen XY, Wang LJ (2007) The effects of increased expression of an Arabidopsis HD-ZIP gene on leaf morphogenesis and anther dehiscence. Plant Science 173: 567-576

Lid SE, Gruis D, Jung R, Lorentzen JA, Ananiev E, Chamberlin M, Niu X, Meeley R, Nichols S, Olsen OA (2002) The defective kernel 1 (dek1) gene required for aleurone cell development in the endosperm of maize grains encodes a membrane protein of the calpain gene superfamily. Proc Natl Acad Sci U S A 99: 5460-5465

Lid SE, Olsen L, Nestestog R, Aukerman M, Brown RC, Lemmon B, Mucha M, Opsahl-Sorteberg HG, Olsen OA (2005) Mutation in the Arabidopisis thaliana DEK1 calpain gene perturbs endosperm and embryo development while over-expression affects organ development globally. Planta 221: 339-351

Lindemose S, O'Shea C, Jensen MK, Skriver K (2013) Structure, function and networks of transcription factors involved in abiotic stress responses. Int J Mol Sci 14: 5842-5878

Lofke C, Luschnig C, Kleine-Vehn J (2013) Posttranslational modification and trafficking of PIN auxin efflux carriers. Mech Dev 130: 82-94

Lohmann JU, Hong RL, Hobe M, Busch MA, Parcy F, Simon R, Weigel D (2001) A molecular link between stem cell regulation and floral patterning in Arabidopsis. Cell 105: 793-803

Long JA, Woody S, Poethig S, Meyerowitz EM, Barton MK (2002) Transformation of shoots into roots in Arabidopsis embryos mutant at the TOPLESS locus. Development 129: 2797-2806

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

Lu S, Zhao H, Des Marais DL, Parsons EP, Wen X, Xu X, Bangarusamy DK, Wang G, Rowland O, Juenger T, Bressan RA, Jenks MA (2012) Arabidopsis ECERIFERUM9 involvement in cuticle formation and maintenance of plant water status. Plant Physiol 159: 930-944

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

Lynn K, Fernandez A, Aida M, Sedbrook J, Tasaka M, Masson P, Barton MK (1999) The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. Development 126: 469-481

MacAlister CA, Ohashi-Ito K, Bergmann DC (2007) Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. Nature 445: 537-540

Mano Y, Nemoto K (2012) The pathway of auxin biosynthesis in plants. J Exp Bot 63: 2853-2872

Marks MD, Wenger JP, Gilding E, Jilk R, Dixon RA (2009) Transcriptome analysis of Arabidopsis wild-type and gl3-sst sim trichomes identifies four additional genes required for trichome development. Mol Plant 2: 803-822

Martin C, Bhatt K, Baumann K, Jin H, Zachgo S, Roberts K, Schwarz-Sommer Z, Glover B, Perez-Rodrigues M (2002) The mechanics of cell fate determination in petals. Philos Trans R Soc Lond B Biol Sci 357: 809-813

Masucci JD, Rerie WG, Foreman DR, Zhang M, Galway ME, Marks MD, Schiefelbein JW (1996) The homeobox gene GLABRA2 is required for position-dependent cell differentiation in the root epidermis of Arabidopsis thaliana. Development 122: 1253-1260

Matsubayashi Y (2011) Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol 52: 5-13

Matsumoto N, Okada K (2001) A homeobox gene, PRESSED FLOWER, regulates lateral axis-dependent development of Arabidopsis flowers. Genes Dev 15: 3355-3364

Mayer KFX, Schoof H, Haecker A, Lenhard M, Jurgens G, Laux T (1998) Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. Cell 95: 805-815

McConnell JR, Barton MK (1998) Leaf polarity and meristem formation in Arabidopsis. Development 125: 2935-2942

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

Menkens AE, Schindler U, Cashmore AR (1995) The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. Trends Biochem Sci 20: 506-510

Meyer MR, Lichti CF, Townsend RR, Rao AG (2011) Identification of in vitro autophosphorylation sites and effects of phosphorylation on the Arabidopsis CRINKLY4 (ACR4) receptor-like kinase intracellular domain: insights into conformation, oligomerization, and activity. Biochemistry 50: 2170-2186

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

Mockaitis K, Estelle M (2008) Auxin receptors and plant development: a new signaling paradigm. Annu Rev Cell Dev Biol 24: 55-80

Moller B, Weijers D (2009) Auxin control of embryo patterning. Cold Spring Harb Perspect Biol 1: a001545

Morelli G, Ruberti I (2002) Light and shade in the photocontrol of Arabidopsis growth. Trends Plant Sci 7: 399-404

Mukherjee K, Brocchieri L, Burglin TR (2009) A comprehensive classification and evolutionary analysis of plant homeobox genes. Mol Biol Evol 26: 2775-2794

Muller B, Sheen J (2008) Cytokinin and auxin interaction in root stem-cell specification during early embryogenesis. Nature 453: 1094-1097

Nadakuduti SS, Pollard M, Kosma DK, Allen C, Jr., Ohlrogge JB, Barry CS (2012) Pleiotropic phenotypes of the sticky peel mutant provide new insight into the role of CUTIN DEFICIENT2 in epidermal cell function in tomato. Plant Physiol 159: 945-960

Nadeau JA (2009) Stomatal development: new signals and fate determinants. Curr Opin Plant Biol 12: 29-35

Nadeau JA, Sack FD (2002) Control of stomatal distribution on the Arabidopsis leaf surface. Science 296: 1697-1700

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

Nakamura M, Katsumata H, Abe M, Yabe N, Komeda Y, Yamamoto KT, Takahashi T (2006) Characterization of the class IV homeodomain-Leucine Zipper gene family in Arabidopsis. Plant Physiol 141: 1363-1375

Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. Plant Mol Biol 60: 51-68

Nakashima K, Takasaki H, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2012) NAC transcription factors in plant abiotic stress responses. Biochim Biophys Acta 1819: 97-103

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

Nawy T, Bayer M, Mravec J, Friml J, Birnbaum KD, 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

Nishiuchi T, Shinshi H, Suzuki K (2004) Rapid and transient activation of transcription of the ERF3 gene by wounding in tobacco leaves: possible involvement of NtWRKYs and autorepression. J Biol Chem 279: 55355-55361

Nodine MD, Bryan AC, Racolta A, Jerosky KV, Tax FE (2011) A few standing for many: embryo receptor-like kinases. Trends Plant Sci 16: 211-217

Nodine MD, Tax FE (2008) Two receptor-like kinases required together for the establishment of Arabidopsis cotyledon primordia. Dev Biol 314: 161-170

Nodine MD, Yadegari R, Tax FE (2007) RPK1 and TOAD2 are two receptor-like kinases redundantly required for arabidopsis embryonic pattern formation. Dev Cell 12: 943-956

Nover L, Bharti K, Doring P, Mishra SK, Ganguli A, Scharf KD (2001) Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? Cell Stress Chaperones 6: 177-189

Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic Acids Res 37: D987-991

Ogo Y, Itai RN, Nakanishi H, Inoue H, Kobayashi T, Suzuki M, Takahashi M, Mori S, Nishizawa NK (2006) Isolation and characterization of IRO2, a novel iron-regulated bHLH transcription factor in graminaceous plants. J Exp Bot 57: 2867-2878

Ohashi-Ito K, Bergmann DC (2006) Arabidopsis FAMA controls the final proliferation/differentiation switch during stomatal development. Plant Cell 18: 2493-2505

Ohyama K, Shinohara H, Ogawa-Ohnishi M, Matsubayashi Y (2009) A glycopeptide regulating stem cell fate in Arabidopsis thaliana. Nat Chem Biol 5: 578-580

Oliva M, Farcot E, Vernoux T (2013) Plant hormone signaling during development: insights from computational models. Curr Opin Plant Biol 16: 19-24

Olsen OA (2004) Nuclear endosperm development in cereals and Arabidopsis thaliana. Plant Cell 16 Suppl: S214-227

Olsson AS, Engstrom P, Soderman E (2004) The homeobox genes ATHB12 and ATHB7 encode potential regulators of growth in response to water deficit in Arabidopsis. Plant Mol Biol 55: 663-677

Otsuga D, DeGuzman B, Prigge MJ, Drews GN, Clark SE (2001) REVOLUTA regulates meristem initiation at lateral positions. Plant J 25: 223-236

Palena CM, Gonzalez DH, Chan RL (1999) A monomer-dimer equilibrium modulates the interaction of the sunflower homeodomain leucine-zipper protein Hahb-4 with DNA. Biochem J 341 (Pt 1): 81-87

Panikashvili D, Savaldi-Goldstein S, Mandel T, Yifhar T, Franke RB, Hofer R, Schreiber L, Chory J, Aharoni A (2007) The Arabidopsis DESPERADO/AtWBC11 transporter is required for cutin and wax secretion. Plant Physiol 145: 1345-1360

Payne CT, Zhang F, Lloyd AM (2000) GL3 encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with GL1 and TTG1. Genetics 156: 1349-1362

Pekker I, Alvarez JP, Eshed Y (2005) Auxin response factors mediate Arabidopsis organ asymmetry via modulation of KANADI activity. Plant Cell 17: 2899-2910

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF (2000) B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature 405: 200-203

Pesch M, Hulskamp M (2009) One, two, three...models for trichome patterning in Arabidopsis? Curr Opin Plant Biol 12: 587-592

Peterson KM, Shyu C, Burr CA, Horst RJ, Kanaoka MM, Omae M, Sato Y, Torii KU (2013) Arabidopsis homeodomain-leucine zipper IV proteins promote stomatal development and ectopically induce stomata beyond the epidermis. Development 140: 1924-1935

Pighin JA, Zheng H, Balakshin LJ, Goodman IP, Western TL, Jetter R, Kunst L, Samuels AL (2004) Plant cuticular lipid export requires an ABC transporter. Science 306: 702-704

Pillitteri LJ, Sloan DB, Bogenschutz NL, Torii KU (2007) Termination of asymmetric cell division and differentiation of stomata. Nature 445: 501-505

Pillitteri LJ, Torii KU (2012) Mechanisms of stomatal development. Annu Rev Plant Biol 63: 591-614

Ploense SE, Wu MF, Nagpal P, Reed JW (2009) A gain-of-function mutation in IAA18 alters Arabidopsis embryonic apical patterning. Development 136: 1509-1517

Pollard M, Beisson F, Li Y, Ohlrogge JB (2008) Building lipid barriers: biosynthesis of cutin and suberin. Trends Plant Sci 13: 236-246

Prigge MJ, Otsuga D, Alonso JM, Ecker JR, Drews GN, Clark SE (2005) Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. Plant Cell 17: 61-76

Qian P, Hou S, Guo G (2009) Molecular mechanisms controlling pavement cell shape in Arabidopsis leaves. Plant Cell Rep 28: 1147-1157

Qing L, Aoyama T (2012) Pathways for epidermal cell differentiation via the homeobox gene GLABRA2: update on the roles of the classic regulator. J Integr Plant Biol 54: 729-737

Rademacher EH, Moller B, Lokerse AS, Llavata-Peris CI, van den Berg W, Weijers D (2011) A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. Plant J 68: 597-606

Rajjou L, Debeaujon I (2008) Seed longevity: survival and maintenance of high germination ability of dry seeds. C R Biol 331: 796-805

Reina-Pinto JJ, Yephremov A (2009) Surface lipids and plant defenses. Plant Physiol Biochem 47: 540-549

Reinhardt D, Frenz M, Mandel T, Kuhlemeier C (2005) Microsurgical and laser ablation analysis of leaf positioning and dorsoventral patterning in tomato. Development 132: 15-26

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

Rerie WG, Feldmann KA, Marks MD (1994) The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. Genes Dev 8: 1388-1399

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. Cell 110: 513-520

Roeder AH, Chickarmane V, Cunha A, Obara B, Manjunath BS, Meyerowitz EM (2010) Variability in the control of cell division underlies sepal epidermal patterning in Arabidopsis thaliana. PLoS Biol 8: e1000367

Roeder AH, Cunha A, Ohno CK, Meyerowitz EM (2012) Cell cycle regulates cell type in the Arabidopsis sepal. Development 139: 4416-4427

Rowland O, Zheng H, Hepworth SR, Lam P, Jetter R, Kunst L (2006) CER4 encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. Plant Physiol 142: 866-877

Sabatini S, Beis D, Wolkenfelt H, Murfett J, Guilfoyle T, Malamy J, Benfey P, Leyser O, Bechtold N, Weisbeek P, 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, 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

Saiga S, Moller B, Watanabe-Taneda A, Abe M, Weijers D, Komeda Y (2012) Control of embryonic meristem initiation in Arabidopsis by PHD-finger protein complexes. Development 139: 1391-1398

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

Sauer M, Balla J, Luschnig C, Wisniewska J, Reinohl V, Friml J, Benkova E (2006) Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. Genes Dev 20: 2902-2911

Sawa S, Ohgishi M, Goda H, Higuchi K, Shimada Y, Yoshida S, Koshiba T (2002) The HAT2 gene, a member of the HD-Zip gene family, isolated as an auxin inducible gene by DNA microarray screening, affects auxin response in Arabidopsis. Plant J 32: 1011-1022

Scarpella E, Marcos D, Friml J, Berleth T (2006) Control of leaf vascular patterning by polar auxin transport. Genes Dev 20: 1015-1027

Schellmann S, Hulskamp M (2005) Epidermal differentiation: trichomes in Arabidopsis as a model system. Int J Dev Biol 49: 579-584

Schellmann S, Schnittger A, Kirik V, Wada T, Okada K, Beermann A, Thumfahrt J, Jurgens G, Hulskamp M (2002) TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. Embo J 21: 5036-5046

Schena M, Davis RW (1992) HD-Zip proteins: members of an Arabidopsis homeodomain protein superfamily. Proc Natl Acad Sci U S A 89: 3894-3898

Schlereth A, Moller B, Liu W, Kientz M, Flipse J, Rademacher EH, Schmid M, Jurgens G, Weijers D (2010) MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. Nature 464: 913-916

Schnittger A, Hulskamp M (2002) Trichome morphogenesis: a cell-cycle perspective. Philos Trans R Soc Lond B Biol Sci 357: 823-826

Schreiber L (2005) Polar paths of diffusion across plant cuticles: new evidence for an old hypothesis. Ann Bot 95: 1069-1073

Schrick K, Nguyen D, Karlowski WM, Mayer KF (2004) START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. Genome Biol 5: R41

Sessa G, Carabelli M, Sassi M, Ciolfi A, Possenti M, Mittempergher F, Becker J, Morelli G, Ruberti I (2005) A dynamic balance between gene activation and

repression regulates the shade avoidance response in Arabidopsis. Genes Dev 19: 2811-2815

Sessa G, Morelli G, Ruberti I (1993) The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. Embo J 12: 3507-3517

Sessa G, Steindler C, Morelli G, Ruberti I (1998) The Arabidopsis Athb-8, -9 and -14 genes are members of a small gene family coding for highly related HD-ZIP proteins. Plant Mol Biol 38: 609-622

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

Shinozaki K, Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Curr Opin Plant Biol 3: 217-223

Shpak ED, Berthiaume CT, Hill EJ, Torii KU (2004) Synergistic interaction of three ERECTA-family receptor-like kinases controls Arabidopsis organ growth and flower development by promoting cell proliferation. Development 131: 1491-1501

Shpak ED, McAbee JM, Pillitteri LJ, Torii KU (2005) Stomatal patterning and differentiation by synergistic interactions of receptor kinases. Science 309: 290-293

Siegfried KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL (1999) Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. Development 126: 4117-4128

Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Two different novel cis-acting elements of erd1, a clpA homologous Arabidopsis gene function in induction by dehydration stress and dark-induced senescence. Plant J 33: 259-270

Smith LG, Oppenheimer DG (2005) Spatial control of cell expansion by the plant cytoskeleton. Annu Rev Cell Dev Biol 21: 271-295

Smith ZR, Long JA (2010) Control of Arabidopsis apical-basal embryo polarity by antagonistic transcription factors. Nature 464: 423-426

Soccio RE, Breslow JL (2003) StAR-related lipid transfer (START) proteins: mediators of intracellular lipid metabolism. J Biol Chem 278: 22183-22186

Stadler R, Lauterbach C, Sauer N (2005) Cell-to-cell movement of green fluorescent protein reveals post-phloem transport in the outer integument and identifies symplastic domains in Arabidopsis seeds and embryos. Plant Physiol 139: 701-712

Stahl Y, Grabowski S, Bleckmann A, Kuhnemuth R, Weidtkamp-Peters S, Pinto KG, Kirschner GK, Schmid JB, Wink RH, Hulsewede A, Felekyan S, Seidel CA, Simon R (2013) Moderation of Arabidopsis Root Stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 Receptor Kinase Complexes. Curr Biol

Stahl Y, Wink RH, Ingram GC, Simon R (2009) A signaling module controlling the stem cell niche in Arabidopsis root meristems. Curr Biol 19: 909-914

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

Stelling J, Sauer U, Szallasi Z, Doyle FJ, 3rd, Doyle J (2004) Robustness of cellular functions. Cell 118: 675-685

Stepanova AN, Robertson-Hoyt J, Yun J, Benavente LM, Xie DY, Dolezal K, Schlereth A, Jurgens G, Alonso JM (2008) TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. Cell 133: 177-191

Sugano SS, Shimada T, Imai Y, Okawa K, Tamai A, Mori M, Hara-Nishimura I (2010) Stomagen positively regulates stomatal density in Arabidopsis. Nature 463: 241-244

Suh MC, Samuels AL, Jetter R, Kunst L, Pollard M, Ohlrogge J, Beisson F (2005) Cuticular lipid composition, surface structure, and gene expression in Arabidopsis stem epidermis. Plant Physiol 139: 1649-1665

Swanson R, Clark T, Preuss D (2005) Expression profiling of Arabidopsis stigma tissue identifies stigma-specific genes. Sex Plant Reprod 18: 163-171

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

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

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

Tanaka H, Dhonukshe P, Brewer PB, Friml J (2006) Spatiotemporal asymmetric auxin distribution: a means to coordinate plant development. Cell Mol Life Sci 63: 2738-2754

Tanaka H, Onouchi H, Kondo M, Hara-Nishimura I, Nishimura M, Machida C, Machida Y (2001) A subtilisin-like serine protease is required for epidermal surface formation in Arabidopsis embryos and juvenile plants. Development 128: 4681-4689

Tanaka H, Watanabe M, Sasabe M, Hiroe T, Tanaka T, Tsukaya H, Ikezaki M, Machida C, Machida Y (2007) Novel receptor-like kinase ALE2 controls shoot development by specifying epidermis in Arabidopsis. Development 134: 1643-1652

Tanaka H, Watanabe M, Watanabe D, Tanaka T, Machida C, Machida Y (2002) ACR4, a putative receptor kinase gene of Arabidopsis thaliana, that is expressed in the outer cell layers of embryos and plants, is involved in proper embryogenesis. Plant Cell Physiol 43: 419-428

Tanaka T, Tanaka H, Machida C, Watanabe M, Machida Y (2004) A new method for rapid visualization of defects in leaf cuticle reveals five intrinsic patterns of surface defects in Arabidopsis. Plant J 37: 139-146

Tao Y, Ferrer JL, Ljung K, Pojer F, Hong F, Long JA, Li L, Moreno JE, Bowman ME, Ivans LJ, Cheng Y, Lim J, Zhao Y, Ballare CL, Sandberg G, Noel JP, Chory J (2008) Rapid synthesis of auxin via a new tryptophan-dependent pathway is required for shade avoidance in plants. Cell 133: 164-176

Terzaghi WB, Cashmore AR (1995) Light-Regulated Transcription. Annu Rev Plant Phys 46: 445-474

Thomas-Chollier M, Defrance M, Medina-Rivera A, Sand O, Herrmann C, Thieffry D, van Helden J (2011) RSAT 2011: regulatory sequence analysis tools. Nucleic Acids Res 39: W86-91

Tian Q, Olsen L, Sun B, Lid SE, Brown RC, Lemmon BE, Fosnes K, Gruis DF, Opsahl-Sorteberg HG, Otegui MS, Olsen OA (2007) Subcellular localization and functional domain studies of DEFECTIVE KERNEL1 in maize and Arabidopsis suggest a model for aleurone cell fate specification involving CRINKLY4 and SUPERNUMERARY ALEURONE LAYER1. Plant Cell 19: 3127-3145

Torii KU (2012) Two-dimensional spatial patterning in developmental systems. Trends Cell Biol 22: 438-446

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7: 562-578

Tripathi P, Rabara RC, Langum TJ, Boken AK, Rushton DL, Boomsma DD, Rinerson CI, Rabara J, Reese RN, Chen X, Rohila JS, Rushton PJ (2012) The WRKY transcription factor family in brachypodium distachyon. BMC Genomics 13: 270

Tron AE, Bertoncini CW, Palena CM, Chan RL, Gonzalez DH (2001) Combinatorial interactions of two amino acids with a single base pair define target site specificity in plant dimeric homeodomain proteins. Nucleic Acids Res 29: 4866-4872

Tron AE, Welchen E, Gonzalez DH (2004) Engineering the loop region of a homeodomain-leucine zipper protein promotes efficient binding to a monomeric DNA binding site. Biochemistry 43: 15845-15851

Tsuwamoto R, Fukuoka H, Takahata Y (2008) GASSHO1 and GASSHO2 encoding a putative leucine-rich repeat transmembrane-type receptor kinase are essential for the normal development of the epidermal surface in Arabidopsis embryos. Plant J 54: 30-42

Turchi L, Carabelli M, Ruzza V, Possenti M, Sassi M, Penalosa A, Sessa G, Salvi S, Forte V, Morelli G, Ruberti I (2013) Arabidopsis HD-Zip II transcription factors control apical embryo development and meristem function. Development 140: 2118-2129

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

Ulmasov T, Hagen G, Guilfoyle TJ (1999) Dimerization and DNA binding of auxin response factors. Plant J 19: 309-319

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

van der Graaff E, Laux T, Rensing SA (2009) The WUS homeobox-containing (WOX) protein family. Genome Biol 10

van Zanten M, Snoek LB, Proveniers MC, Peeters AJ (2009) The many functions of ERECTA. Trends Plant Sci 14: 214-218

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: 2269-2283

Vandepoele K, Quimbaya M, Casneuf T, De Veylder L, Van de Peer Y (2009) Unraveling transcriptional control in Arabidopsis using cis-regulatory elements and coexpression networks. Plant Physiol 150: 535-546

Vanneste S, Friml J (2009) Auxin: a trigger for change in plant development. Cell 136: 1005-1016

Verkest A, Manes CL, Vercruysse S, Maes S, Van Der Schueren E, Beeckman T, Genschik P, Kuiper M, Inze D, De Veylder L (2005) The cyclin-dependent kinase inhibitor KRP2 controls the onset of the endoreduplication cycle during Arabidopsis leaf development through inhibition of mitotic CDKA;1 kinase complexes. Plant Cell 17: 1723-1736

Vernoux T, Brunoud G, Farcot E, Morin V, Van den Daele H, Legrand J, Oliva M, Das P, Larrieu A, Wells D, Guedon Y, Armitage L, Picard F, Guyomarc'h S, Cellier C, Parry G, Koumproglou R, Doonan JH, Estelle M, Godin C, Kepinski S, Bennett M, De Veylder L, Traas J (2011) The auxin signalling network translates dynamic input into robust patterning at the shoot apex. Mol Syst Biol 7: 508

Vielle-Calzada JP, Baskar R, Grossniklaus U (2000) Delayed activation of the paternal genome during seed development. Nature 404: 91-94

Vieten A, Vanneste S, Wisniewska J, Benkova E, Benjamins R, Beeckman T, Luschnig C, Friml J (2005) Functional redundancy of PIN proteins is accompanied by auxin-dependent cross-regulation of PIN expression. Development 132: 4521-4531

Wabnik K, Govaerts W, Friml J, Kleine-Vehn J (2011) Feedback models for polarized auxin transport: an emerging trend. Mol Biosyst 7: 2352-2359

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC (1999) The TRANSPARENT TESTA

GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. Plant Cell 11: 1337-1350

Wang C, Barry JK, Min Z, Tordsen G, Rao AG, Olsen OA (2003a) The calpain domain of the maize DEK1 protein contains the conserved catalytic triad and functions as a cysteine proteinase. J Biol Chem 278: 34467-34474

Wang G, Fiers M (2010) CLE peptide signaling during plant development. Protoplasma 240: 33-43

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

Wang Y, Henriksson E, Soderman E, Henriksson KN, Sundberg E, Engstrom P (2003b) The Arabidopsis homeobox gene, ATHB16, regulates leaf development and the sensitivity to photoperiod in Arabidopsis. Dev Biol 264: 228-239

Wang ZY, Xiong L, Li W, Zhu JK, Zhu J (2011) The plant cuticle is required for osmotic stress regulation of abscisic acid biosynthesis and osmotic stress tolerance in Arabidopsis. Plant Cell 23: 1971-1984

Watanabe M, Suwabe K, Suzuki G (2012) Molecular genetics, physiology and biology of self-incompatibility in Brassicaceae. Proc Jpn Acad Ser B Phys Biol Sci 88: 519-535

Watanabe M, Tanaka H, Watanabe D, Machida C, Machida Y (2004) The ACR4 receptor-like kinase is required for surface formation of epidermis-related tissues in Arabidopsis thaliana. Plant J 39: 298-308

Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000) Activation tagging in Arabidopsis. Plant Physiol 122: 1003-1013

Weijers D, Benkova E, Jager KE, Schlereth A, Hamann T, Kientz M, Wilmoth JC, Reed JW, Jurgens G (2005a) Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators. Embo J 24: 1874-1885

Weijers D, Jurgens G (2004) Funneling auxin action: specificity in signal transduction. Curr Opin Plant Biol 7: 687-693

Weijers D, Sauer M, Meurette O, Friml J, Ljung K, Sandberg G, Hooykaas P, Offringa R (2005b) Maintenance of embryonic auxin distribution for apical-basal patterning by PIN-FORMED-dependent auxin transport in Arabidopsis. Plant Cell 17: 2517-2526

Weijers D, Schlereth A, Ehrismann JS, Schwank G, Kientz M, Jurgens G (2006) Auxin triggers transient local signaling for cell specification in Arabidopsis embryogenesis. Dev Cell 10: 265-270

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS One 2: e718

Wu G, Lin WC, Huang T, Poethig RS, Springer PS, Kerstetter RA (2008) KANADI1 regulates adaxial-abaxial polarity in Arabidopsis by directly repressing the transcription of ASYMMETRIC LEAVES2. Proc Natl Acad Sci U S A 105: 16392-16397

Wu R, Li S, He S, Wassmann F, Yu C, Qin G, Schreiber L, Qu LJ, Gu H (2011) CFL1, a WW domain protein, regulates cuticle development by modulating the function of HDG1, a class IV homeodomain transcription factor, in rice and Arabidopsis. Plant Cell 23: 3392-3411

Wu X, Chory J, Weigel D (2007) Combinations of WOX activities regulate tissue proliferation during Arabidopsis embryonic development. Dev Biol 309: 306-316

Wu X, Dabi T, Weigel D (2005) Requirement of homeobox gene STIMPY/WOX9 for Arabidopsis meristem growth and maintenance. Curr Biol 15: 436-440

Wysocka-Diller JW, Helariutta Y, Fukaki H, Malamy JE, Benfey PN (2000) Molecular analysis of SCARECROW function reveals a radial patterning mechanism common to root and shoot. Development 127: 595-603

Xing Q, Creff A, Waters A, Tanaka H, Goodrich J, Ingram GC (2013) ZHOUPI controls embryonic cuticle formation via a signalling pathway involving the subtilisin protease ABNORMAL LEAF-SHAPE1 and the receptor kinases GASSHO1 and GASSHO2. Development 140: 770-779

Yadav RK, Girke T, Pasala S, Xie M, Reddy GV (2009) Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. Proc Natl Acad Sci U S A 106: 4941-4946

Yadav RK, Perales M, Gruel J, Girke T, Jonsson H, Reddy GV (2011) WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. Genes Dev 25: 2025-2030

Yadegari R, Drews GN (2004) Female gametophyte development. Plant Cell 16 Suppl: S133-141

Yanai O, Shani E, Dolezal K, Tarkowski P, Sablowski R, Sandberg G, Samach A, Ori N (2005) Arabidopsis KNOXI proteins activate cytokinin biosynthesis. Curr Biol 15: 1566-1571

Yang S, Johnston N, Talideh E, Mitchell S, Jeffree C, Goodrich J, Ingram G (2008) The endosperm-specific ZHOUPI gene of Arabidopsis thaliana regulates endosperm breakdown and embryonic epidermal development. Development 135: 3501-3509

Yephremov A, Wisman E, Huijser P, Huijser C, Wellesen K, Saedler H (1999) Characterization of the FIDDLEHEAD gene of Arabidopsis reveals a link between adhesion response and cell differentiation in the epidermis. Plant Cell 11: 2187-2201

Yeung EC, Meinke DW (1993) Embryogenesis in Angiosperms: Development of the Suspensor. Plant Cell 5: 1371-1381

Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E (2011) AGRIS: the Arabidopsis Gene Regulatory Information Server, an update. Nucleic Acids Res 39: D1118-1122

Yoshida S, Saiga S, Weijers D (2013) Auxin regulation of embryonic root formation. Plant Cell Physiol 54: 325-332

Yu D, Chen C, Chen Z (2001) Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. Plant Cell 13: 1527-1540

Yu H, Chen X, Hong YY, Wang Y, Xu P, Ke SD, Liu HY, Zhu JK, Oliver DJ, Xiang CB (2008) Activated expression of an Arabidopsis HD-START protein confers drought tolerance with improved root system and reduced stomatal density. Plant Cell 20: 1134-1151

Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. Development 130: 4859-4869

Zhang Z, Laux T (2011) The asymmetric division of the Arabidopsis zygote: from cell polarity to an embryo axis. Sex Plant Reprod 24: 161-169

Zhao H, Li X, Ma L (2012) Basic helix-loop-helix transcription factors and epidermal cell fate determination in Arabidopsis. Plant Signal Behav 7

Zhao Y (2010) Auxin biosynthesis and its role in plant development. Annu Rev Plant Biol 61: 49-64

Zhu H, Hu F, Wang R, Zhou X, Sze SH, Liou LW, Barefoot A, Dickman M, Zhang X (2011) Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell 145: 242-256

Zwiewka M, Feraru E, Moller B, Hwang I, Feraru MI, Kleine-Vehn J, Weijers D, Friml J (2011) The AP-3 adaptor complex is required for vacuolar function in Arabidopsis. Cell Res 21: 1711-1722