A thesis submitted to the University of East Anglia for the Degree of Doctor of Philosophy

Submitted August 2010

Comparative Investigations of the Regulatory Roles of MYB Transcription Factors across the Plant Kingdom

Silvia Mª Jiménez Sanz

Department of Metabolic Biology
John Innes Centre
Norwich, UK

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to Cathie Martin for the opportunities she has given me: to be part of her group and providing me with the environment in which to develop a valuable understanding of plant science. It has been an unforgettable experience that I will always value. Thanks also to the EU for the Marie Curie Plants and Microbes MEST-CT-2005-019727 project for funding my PhD studies.

Thanks to my family, friends, colleagues, JIC staff and those special persons for their love, help and support in both the good times and the bad. There are too many of you to mention but I have enjoyed working alongside you, your friendship and all the shared experiences over the years. I will treasure the memories we share and I am grateful that you are part of my life and have made me feel so loved and happy. I would like to express thanks to Goyi for a lot of things but above all, for being always there and being my best friend. I do not want to forget Norwich, my home for the past years and a very special place in my thoughts forever.

Finally, I especially want to dedicate this thesis to the most important people to me, my parents, Jose Manuel and María del Carmen for their unconditional love, dedication and support. Thanks so much for making me the person I am today. I owe you a great debt and I love you.

En principio, me gustaría expresar mi gratitud a Cathie Martin por las oportunidades que me ha dado: ser parte de su grupo y proveerme del entorno en el que desarrollar un valioso conocimiento científico en plantas. Gracias también a la Unión Europea por financiar mis estudios de doctorado (proyecto MEST-CT-2005-019727).

Agradecer a mi familia, amigos, compañeros, personal del JIC y a esas otras personas especiales su amor, ayuda y apoyo en los buenos y malos momentos. Sois muchos para mencionar pero no os olvido; he disfrutado el trabajar con vosotros, vuestra amistad y todas las experiencias compartidas durante estos años. Guardaré en mi memoria todo aquello que compartimos y agradecida quedo de que seáis parte de mi vida y me hayáis hecho sentir tan querida y feliz. Me gustaría dar las gracias también a Goyi por muchas cosas pero sobre todo, por estar siempre ahí y ser mi mejor amigo. Tampoco olvidar Norwich, mi hogar en estos años y un lugar muy especial para mis recuerdos siempre.

Finalmente, quiero dedicar especialmente esta tesis a las personas más importantes para mí, mis padres, Jose Manuel y María del Carmen, por su amor, dedicación y apoyo. Gracias por hacer de mí la persona que soy hoy, os debo todo, os quiero.

ABSTRACT

MYB proteins belong to a family of transcription factors that regulate an important number of plant-specific processes. The family is highly restricted in slime moulds, fungi and animals whereas in higher plants it is large; there are 126 *R2R3 MYB* genes in Arabidopsis, more than 80 in maize and rice and about 140 in Lotus. In the moss, *Physcomitrella patens*, however, 48 *R2R3 MYB* genes have been identified from the genome sequence. The *P. patens* genes that share common functions with those in Angiosperms may well represent genes that had similar functions in the last common ancestor between Angiosperms and Bryophytes, some 400 Mya.

I wished to examine the function of two members of this family of transcription factors from *P. patens* as a basis for understanding the functional diversification of the MYB family during the evolution of plants. These members were believed to participate in a signalling pathway that is conserved between mosses and higher plants. If the function of these regulatory proteins could be established in *P. patens,* then comparisons could be made to determine which biological functions have been conserved in the evolution of the R2R3 MYB family in Angiosperms and where neo-functionalisation has occurred.

Careful phylogenetic analysis of two *MYB*-like genes from *P. patens*, *PpMYB4* and *PpMYB5*, revealed that two members of the Arabidopsis *MYB* family, *AtMYB8* and *AtMYB6,* were closely related structurally, but functionally poorly characterised. However, a mutant of Arabidopsis with a deficiency in tolerance to cold and salt stress (dehydration stress in general), *hos10*, had just been reported at the start of my project as a loss of function mutant of *AtMYB8*. This work suggested that the principal biochemical role of the AtMYB8/HOS10 transcription factor was to regulate ABA biosynthesis. In moss, the ability to survive dehydration is of primary physiological importance. Molecular responses to ABA and dehydration stress have also been shown to be conserved between moss and angiosperms. The aim of my research was to investigate whether the moss *MYB*-like genes regulate the same processes to

their Arabidopsis counterpart(s) and whether this function of *MYB*-like genes in plants had been conserved over 400 million years since these lineages diverged.

During the progress of this project I discovered that *hos10* was not a mutation of the *AtMYB8* gene, as previously published. Consequently, the initial hypothesis about the roles of the MYB proteins in abiotic stress had to be dismissed and I generated new data suggesting the involment of these MYB proteins in phenylpropanoid metabolism, a pathway responsible for the production of a wide range of secondary metabolites that seems to have been conserved for more than 400 Mya. In this work, I propose *AtMYB6* and *AtMYB8* and closely related genes, *PpMYB4* and *PpMYB5,* in *P. patens*, as putative regulators of the phenylpropanoid biosynthetic pathway, suggesting that these genes have been functionally conserved during the evolution of plants.

LIST OF CONTENTS

ABBREVIATIONS

aa	amino acid
amu	atomic mass unit
BAP	benzylaminopurine
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bp	base pair
BSA	bovine serum albumin
ºC	degrees centigrade
CaMV	cauliflower mosaic virus
cDNA	complementary DNA
cfu	colony forming unit
cm	centimetre
Col-0	Columbia-0 ecotype
dH_2O	distilled water
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTPs	deoxyribonucleotide triphosphates
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
g	gramme
g	gravitational acceleration
GFP	green fluorescent protein
GST	glutathione S-transferase
GUS	β-glucuronidase
h	hour
HPLC	high performance liquid chromatography
IP	immunoprecipitation
IPTG	isopropyl beta-D-thiogalactoside

kb	kilobase pair
kDa	kilodalton
KO	knockout
kV	kilovolt
l	litre
LB	Luria Bertani medium
M	molar
MBP	maltose binding protein
MES	2-(N-morpholino)ethanesulphonic acid
MetOH	methanol
min	minute
mg	miligramme
mm	millimetre
mM	millimolar
MS	Murashige and Skoog nutrient medium
MOPS	3-(N-morpholino)propanesulfonic acid
NAA	naphthaleneacetic acid
NBT	4-nitroblue tetrazolium chloride
ng	nanogramme
nm	nanometre
nptII	neomycin phosphotransferase II
OD	optical density
o/n	overnight
ORF	open reading frame
OX	overexpression
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pJAM	plasmid John Innes Centre Antirrhinum majus
PMSF	phenylmethylsulfonyl fluoride
prom	promoter

PVPP	polyvinylpyrrolidone
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse transcription-polymerase chain reaction
s	second
SDS	sodium dodecylsulphate
SSC	standard saline citrate
TBE	tris-borate / EDTA
TE	tris / EDTA
Tm	melting temperature
Tris	2 amino-2-(hydroxymethyl)-l ,3-propanediol
UV	ultra violet light
V	volts
v/v	Volume/volume
w/v	Weight/volume
Wt	Wild-type
X-Glu	5-bromo-4-chloro-3-indolyl B-D-glucuronide
µg	microgramme
µM	micromolar
µl	microlitre
µm	micrometre

GENERAL MATERIALS AND METHODS

1. Gateway cloning

For all the constructs, the GATEWAY™ conversion technology (Invitrogen, Gaithersburg, MD, USA) was used.

Sequences to be cloned were amplified by PCR using primers containing the *attB* sites to generate an entry clone by recombination between *attB*-flanked DNA fragments and *attP*-containing vector (BP reaction). The PCR products were gel-purified with Qiaquick gel extraction kit (Qiagen) and cloned into the Donor vectors pDONR201 or pDONR207 to create an entry clone.

Reactions were performed as described in the Gateway manual (Invitrogen/Life Technologies). For the BP reaction, 1 µl BP Clonase II mix, 2 µl *attB*-PCR product (~ 3.5 ng/ µl), 1 µl Donor vector (~150 ng/ µl), and TE 1X buffer up to 6.37 µl were used. Reactions were incubated usually at RT for 1-2 hours and stopped with 1 µl Proteinase K at 37ºC for 10 min. An aliquot of the reaction (3 µl) was used for transformation of *E.coli* DH5α cells. Transformed cells were grown on L-agar plates supplemented with the appropriate antibiotic depending on the Donor vector used. For pDONR207, 20 µg/ml gentamycin was added and 50 µg/ml kanamycin for pDONR201.

Entry clones (pDONR201 or pDONR207 containing the *attB* PCR product) were used in a second recombination reaction (LR reaction) to transfer the insert into different expression vectors using 1.33 µl LR Clonase II mix, 1 µl Destination vector at 300 ng/ µl, 1 µl Entry clone at 20–200ng/ µl, and TE 1X buffer up to 6.37 µl. After 1-2 hours incubation at RT, reactions were stopped using 1 µl Proteinase K at 37ºC for 10 min. The resulting expression clones were transformed into *DH5α* competent cells and plated on L-agar plates supplemented with the appropriate antibiotic. Colonies were confirmed by sequencing with specific primers. Plasmids were prepared using QIAprep Spin Miniprep Kit (Qiagen).

2. Bacterial transformation

For cloning experiments and plasmid propagation *E.coli DH5α* was used. To transform competent cells (John Innes SOP, access via Intranet)1-50 ng of plasmid was added to an aliquot of 100 µl of thawed-on-ice cells and left on ice for 30 min. The cells were then given a heat shock at 42ºC for 45 s and returned to ice for 5 min. L-broth (800 µl) was then added and the cells incubated at 37ºC for 1 h to allow the expression of the antibiotic resistance. Cells were collected by centrifugation in a benchtop centrifuge and resuspended in 100 µl of L-broth. Aliquots of 50 µl were then spread onto L-agar plates containing the appropriate antibiotic and left to grow at 37ºC overnight.

Protein expression was carried out in the strain BL-21 (DE3) Codon Plus. Cells were transformed following the protocol above and colonies selected on L-agar plates supplemented with ampicillin and cloramphenicol (100 and 34 µg/ml final concentration respectively).

For *Agrobacterium tumefaciens* transformation, 1 (1-50 ng) DNA was added to a 100 µl aliquot of competent cells which were then flash-frozen in liquid nitrogen. Cells were then incubated at 37ºC for 5 min in a waterbath. 1 ml of L-broth was then added and the cells incubated at 28ºC for 2 h with gentle shaking. Cells were then centrifuged briefly in a benchtop microcentrifuge and then resuspended in 100 µl of L-broth before being spread onto L-agar plates containing suitable antibiotics and incubated at 28ºC for 48 h. All strains were grown with rifampicin at 50 µg/ml. The pSOUP plasmid vector was selected using tetracycline 2 µg/ml. Other antibiotics were used depending on the vectors introduced. For long term storage, strains were suspended in sterile 40% glycerol (v/v) and frozen at - 80ºC.

3. Bacterial DNA isolation

Small and medium scale preparations of plasmid DNA from bacteria were carried out with a Qiagen miniprep kit or a Qiagen plasmid midi kit respectively. *E.coli* cells were grown from a single colony in either 10 ml (small

scale) or subcultured in 100 ml (medium scale) of L-broth at 37ºC over a period of 16 hours with agitation and appropiate antibiotic selection. Bacterial cells were then processed using the kit following the manufacturer's instructions.

4. Isolation of total RNA

Small scale RNA extractions were carried out from tissue frozen in liquid nitrogen and ground to a fine powder with a chilled pestle and mortar and then processed using a Qiagen RNeasy plant mini kit as per the manufacturer's instructions. Samples were treated with an on-column DNase (Qiagen) to remove any possible DNA contamination. After washing to remove contamination, RNA was eluted from the membrane in 30 µl RNase-free water.

5. First strand cDNA synthesis

Total RNA extracted was treated with RNase-free DNaseI (Amersham). Treated samples were used directly in the cDNA synthesis reaction using SuperScript™ III Reverse Transcriptase (Invitrogen). First strand cDNA was synthesized from 2 µg total RNA using an oligo(Briggs et al.)$_{20}$ primer and accompanying reagents according to the manufacturer's instructions.

6. Polymerase chain reaction (PCR)

In a typical reaction, 25 µl final volume was used. Reactions contained approximately 2 ng/µl DNA template (either genomic, cDNA or plasmid DNA), 0.3 µl dNTPs (from a 10 mM stock), 0.3 µl each primer (10 µM stock) (see list of primers used in this work in Appendix 1), 1X Qiagen Taq Polymerase buffer (10X stock) and 0.2 µl Qiagen Taq Polymerase (5 U/µl stock). The reaction volume was made up to the total with sterile dH_2O. Usually, cycles of amplification were performed under the following conditions: initial denaturation at 96ºC for 4 min, followed by 35 cycles of 96 ºC for 30 s, primer annealing in between 55-57 ºC (depending on the primer Tm) for 30 s, 72 ºC for 1min for every 1000 bp being amplified and an extension at 72 ºC for 5 min. For cloning,

a high fidelity proofreading Taq (Qiagen) was used instead. Reactions were carried out in a HYBAID thermal cycler.

The above protocol was modified slightly to screen bacterial colonies for the presence of certain DNA fragments during cloning. In these colony PCRs a bacterial colony was used as a template rather than a DNA solution. Bacterial cells were broken down during the initial denaturation stage of the PCR due to the combination of heat and detergents present in the AmpliTaq buffer, releasing DNA into the reaction mixture.

7. Separation of DNA fragments on agarose gels

Electrophoresis through horizontal 0.8% to 2.5% (w/v) agarose gels was used to separate DNA fragments. Gels were prepared by melting the agarose in 1x TBE (0.089 M Tris pH 8.0, 0.089 M boric acid, 0.002 M EDTA) in a microwave oven, then 0.5 g/ml ethidium bromide was added and the gel poured into a horizontal gel tray. One-tenth of the volume of loading buffer (20% ficoll, 1 mM EDTA pH 8.0, 0.25% bromophenol blue, 0.25% xylene cyanol) was added to the DNA samples before loading onto the gel. An aliquot of 1 kb ladder (Invitrogen) was also loaded as a DNA size marker. Gels were run in 1x TBE at 100-200 V and the DNA fragments subsequently visualised and photographed on an ultraviolet transilluminator using the GelDoc software (BioRad).

8. Isolation of DNA from agarose gels

After electrophoresis, the appropriate DNA fragments for either probe or construct production were excised using a razor blade. DNA was extracted from the gel and purified using a Qiagen Gel Extraction Kit according to the manufacturer's instructions.

9. Restriction digests of DNA

Restriction digests were used for the cloning of DNA fragments into plasmids and for checking the presence and orientation of DNA inserted into

plasmids. Restriction digests were carried out in 1.5 ml tubes usually at 37°C for a minimum of one hour. Restriction digestion reactions typically consisted of 10% (v/v) 10x restriction buffer supplied by the enzyme manufacturers, 10% (v/v) restriction enzyme, with the remainder of the reaction volume being the template DNA to be digested. In situations where two restriction enzymes were used together in the same digestion reaction, the total volume of both enzymes together did not exceed 10% of the reaction volume. This limitation was to prevent cutting at non-designated sites (star activity) as a result of the high concentrations of glycerol in which the enzymes are stored.

10. Ligation of DNA fragments

Ligation reactions were set up containing restriction digested insert and plasmid in a roughly 3:1 insert to plasmid ratio to reduce the probability of plasmid self ligation. Reactions contained 10% (v/v) 10x T4 DNA ligase buffer and 10% (v/v) T4 DNA ligase (Roche). Reactions were incubated o/n at 12°C before being transformed into *E. coli* strain *DH5α*.

11. Arabidopsis transformation

Arabidopsis thaliana Col-0 ecotype was selected for the transformation with *Agrobacterium tumefaciens* by the floral dip method (Clough and Bent, 1998). *Agrobacterium* GV3101 strain harbouring the relevant construct was grown overnight at 28°C in 10 ml of LB medium with the appropriate antibiotics. The overnight culture was added to 300 mL of fresh medium with the same antibiotics and grown to stationary phase (OD_{600} 0.8-1.0). Cells were harvested by centrifuging at 6000 rpm for 20 min. The pellet was resuspended in infiltration medium (2.8 g/l MS+MES (Duchefa), 50 g/l sucrose [adjusted to pH 5.5 with KOH]) and just before dipping 10 µl/l BAP and 50 µl/l Silwet L-77 were added.

Plants with numerous immature floral buds were inoculated by submerging inflorescences in the bacterial suspension. 4-6 pots containing 9 plants each were used. Autoclave bags were then used to cover the plants

which were left on the bench overnight. Seeds were collected when all siliques were dry. The selection of transformed plants was done in five Petri dishes containing 300-400 sterilised seeds per plate on medium containing containing 50 µg/ml kanamycin. Seeds were incubated in a growth chamber at 23ºC±1ºC under long days (16 hr light/8 hr dark cycles) for approximately 10 d, until plants reached the 4-leaf stage, to ensure kanamycin resistance. T2 (obtained after growing the resistant plants up and allowing them to self) lines were transferred to soil into a controlled environment room under short days (8 hr light/16 hr dark cycles for 4 few weeks and after placed under long days. Plants transformed with the plasmid containing the basta resistance gene were selected by spraying with Basta 0.1%. The T3 plants selected from kanamycin or basta resistant T2 plants were used to perform subsequent experiments.

12. Arabidopsis DNA isolation

For screening a large number of plants (*e.g.* genotyping), a quick NaOH-Tris method was used for the extraction of genomic DNA from the *A.thaliana* leaves (Collard *et al*., 2007). A leaf disc from 3-4 week old plants was collected from individual plants by capping the lid of a sterile microfuge tube over the leaf. A volume of 100 µl of 0.5 M NaOH was added to the tissue in the tube containing two small ball bearings. The samples were ground in a GenoGrinder (BT&C/OPS Diagnostics) 2000 for 1 min at 400 rpm followed by addition of 800 µl of 100 mM Tris-HCl (0.1M pH 8) and centrifugation for 2 min at 10000 rpm. For subsequent PCR analysis, 2 µl of the supernatant were used. For cleaner extractions of DNA, a Qiagen kit was used.

CHAPTER 1

Chapter 1: Introduction

1.1. Transcription factors

Regulation of gene expression is a fundamental process in all living organisms. A number of different factors are required for the process of transcription. Some of these factors are referred to as transcription factors (TFs).

TFs are proteins that bind *cis* regulatory elements, such as promoters and enhancers, to regulate the transcription of different genes (Lee and Young, 2000). Promoters are composed of common sequence elements, like TATA boxes and binding sites for other TFs, which work together to recruit the general (basal) transcriptional machinery to the transcriptional start site of a target gene. Enhancer elements are also bound by TFs but are located some distance from the site of transcription initiation (Sandelin *et al*., 2007).

Generally, TFs usually contain at least two domains: a DNA binding domain for recruitment to DNA and an activation or repression domain (or both) for regulating transcription. Other proteins can be recruited to promoters in the absence of a DNA-binding domain by protein-protein interactions and may modulate the rate of transcription. These are termed cofactors or co-adaptors. Furthermore, in some TFs the DNA binding domains themselves may act as transcriptional regulatory domains. TFs usually do not function alone, but function in concert with other TFs and accessory proteins to regulate transcription (Sharrocks, 2000).

Some TF families are found in both plants and animals and originated from common ancestral genes while others evolved specifically in plants (Yanagisawa, 1998). On the basis of similarities in their DNA-binding domains, most TFs can be classified into structural families on the basis of the presence of specific DNA-recognition motifs or folds. Currently, 64 TF families have been characterized in plants (Fig. 1.1) (Guo *et al*., 2005). The PlnTFDB database contains the protein models for TFs found in 22 species (Guo *et al*., 2008).

ABI3VP1	CAMTA	LFY	SBP
Alfin-like	CCAAT	LIM	Sigma70-like
AP2-EREBP	CPP	LOB	SRS
ARF	CSD	MADS	TAZ
ARR-B	DBP	mTERF	TCP
BBR/BPC	E2F-DP	MYB	Tify
BES1	EIL	MYB-related	TIG
bHLH	FAR1	NAC	Trihelix
BSD	FHA	NOZZLE	TUB
bZIP	G2-like	OFP	ULT
C2C2-CO-like	GeBP	Orphans	VARL
C2C2-Dof	GRAS	PBF-2-like	VOZ
C2C2-GATA	GRF	PLATZ	WRKY
C2C2-YABBY	HB	RWP-RK	zf-HD
C2H2	HRT	S1Fa-like	Zn-clus
C3H	HSF	SAP	

Figure 1.1 Transcription factors families

Transcription factors families listed by the Plant Transcription Factor Database (PlnTFDB). The database was built from analyses of several plant species and currently contains 28193 protein models, arranged in 84 families (transcription factor families plus other transcriptional regulators). The assignment of proteins to each of the families is based on the presence of one or more characteristic domains previously described in the literature.

Over 1600 genes encoding transcription factors have been identified in the Arabidopsis genome sequence, representing approximately 6% of the total number of genes (Riechmann *et al*., 2000).

As regulators of cellular processes, TFs are excellent candidates for modifying complex traits in crop plants and TF-engineering is likely to be important in the next generation of successful “biotech” crops. There are multiple transcriptional regulators involved in controling traits that could be useful for crop improvement. For example, in tomato, expression of the transgenes, *Delila* and *Rosea1*, from the snapdragon *Antirrhinum majus*, resulted in the accumulation of anthocyanins which are potent antioxidants good for health (Butelli *et al*., 2008), while alterations in the expression of *OsTB1* could be used to increase or decrease tiller number and thereby adapt rice morphology to different agronomic situations (Taito *et al*., 2003). TFs can play important roles, and improving exploitation of our current knowledge of molecular pathways influencing crop yields and understanding their regulatory activity should help us to produce crops with improved agronomic traits.

1.2. The plant MYB transcription factor family

MYB proteins constitute a diverse class of TFs of particular importance in transcriptional regulation in plants. Members of this family are characterised by having a structurally conserved MYB DNA-binding domain. The *MYB* acronym is derived from "*myeloblastosis*", a type of leukaemia (cancer of blood cells). The first *MYB* gene identified was the oncogene *v-Myb* derived from the avian myeloblastosis virus (AMV). This retrovirus causes monoblastic leukemia in chickens and possibly may have originated by the insertion of a retrovirus into a proto-oncogene (c-*Myb*), which mutated once it became part of the virus (Klempnauer *et al*., 1982).

The MYB TF family is an unusually large protein family in plants, and in fact, it is considered one of the largest (Riechmann *et al*., 2000) and different members regulate an important number of processes that include regulation of secondary metabolism, development, defence and stress reponses, light and

hormone signalling and the cell cycle (Jin and Martin, 1999; Yanhui *et al*., 2006). In plants, the first MYB-like protein identified was encoded by the *C1* gene of *Zea mays*, which is required for the synthesis of anthocyanins (Paz-Ares *et al*., 1987). The functions of many more MYB TFs have been described since then. For instance, 126 R2R3 MYB genes are present in the Arabidopsis genome (Riechmann *et al*., 2000; Stracke *et al*., 2001), more than 80 MYB genes in maize (Rabinowicz *et al*., 1999), at least 30 in *Petunia hybrida* (Avila *et al*., 1993) about 100 in Lotus, around 200 in cotton (Cedroni *et al*., 2003) and 6 were originally identified in the moss, *Physcomitrella patens* (Leech *et al*., 1993). Since the release of the *P.patens* genome sequence in 2008 they are now estimated to number 48.

In contrast to plants, the MYB-related genes of vertebrates comprise a small family with a central role in controlling cellular proliferation (Martin and Paz-Ares, 1997). Many vertebrates contain three genes related to *v-Myb*: *c-Myb*, *A-Myb* and *B-Myb* (Fig. 1.2) (Weston, 1998) and other similar genes have been identified in insects, plants, fungi and slime moulds (Lipsick, 1996). The most completely characterized protein of the group is c-Myb, which is involved in the regulation of proliferation and differentiation of immature hematopoietic cells (Duprey and Boettiger, 1985). The animal MYB proteins all have three repeats comprising their MYB DNA binding domain. There is a small group of three repeat MYB proteins in plants, including 9 members in Arabidopsis (Stracke *et al*., 2001), which also function in controlling the cell cycle and cellular proliferation. However the largest group of MYB proteins in plants have only two repeats in their DNA binding domains (R2R3) and these are unique to plants.

The remarkable number of MYB proteins in plants reflects the diverse number of roles these transcription factors have assumed in plants. The functions of many of the members of this plant gene family are still unknown although in 1995 a European Union Project was initiated to carry out a search for the function of MYB transcription factors in plants.

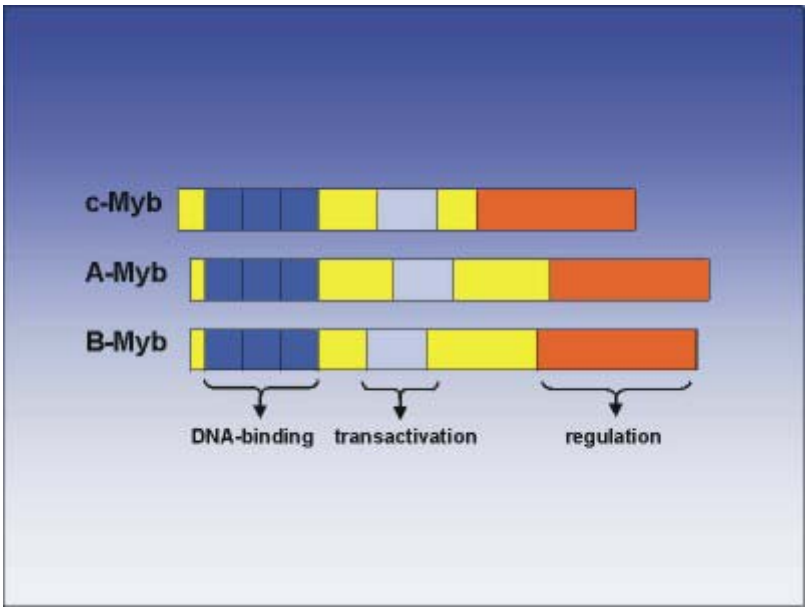


Figure 1.2. **The Myb family in vertebrates**

Schematic structure of vertebrate c-Myb, A-Myb and B-Myb proteins. Blue, grey and orange boxes show the different domains in the protein. v-Myb, from Avian Myoblastosis Virus (AMV) is a truncated version of c-Myb, lacking aminoacid sequences from the amino- and carboxyl terminus of c-Myb (http://www.uni-muenster.de/Chemie.bc/fields/ak-klempnauer/khk topics.html).

This project aimed to identify the function of most members the *MYB* gene family in *Arabidopsis* through a collaboration of several European research groups. Initially, 80 MYB genes in Arabidopsis were described (Romero *et al*., 1998). Although this project identified the function of only a few *MYB* genes, it generated tools and resources for searching further for the functions of members of the MYB family. Recently, 198 genes encoding proteins which contain regions with homology to the MYB repeats have been identified from an analysis of the complete Arabidopsis genome sequence (Yanhui *et al*., 2006). Because of their structural and functional diversity, MYB proteins offer an excellent opportunity to study the evolution of biochemical pathways and their integration into complex developmental cascades.

1.2.1. MYB transcription factor structure

It has been mentioned previously that the *R2R3 MYB* gene family encodes proteins which share a structurally conserved MYB DNA-binding domain located at the N-terminus of each protein. MYB domains are composed of one, two or three repeat motifs of approximately 50 amino acids each. Each motif is capable of forming 3 α-helices with the second and third helices forming a helix-turn-helix structure when bound to DNA (Martin and Paz-Ares, 1997). Tandem repeats of these motifs are referred as the R1, R2, and R3 repeats (Fig. 1.3) after the repeats identified in the prototypic MYB protein c-Myb.

MYB proteins in animals generally contain repeats R1, R2, and R3 (encoded by R1R2R3 class genes) whereas most of MYB proteins in plants have just two repeats (R2, R3) very similar to R2 and R3 of the animal c-MYB proteins. These differences may have resulted from duplication or triplication of the basic repeat unit during evolution (Martin and Paz-Ares, 1997). The large R2R3 MYB family of proteins has been found only in plants to date. Although this type is the predominant MYB group in plants, proteins with four (MYB4R), three (MYB3R) and one (MYB1R) MYB repeats have also been identified in the plant kingdom (Baranowskij *et al*., 1994; Kranz *et al*., 2000; Stracke *et al*., 2001).

Figure 1.3. Structure of plant MYB proteins

Representation of different MYB proteins in the plant kingdom. Dark shaded boxes represent the conserved DNA MYB binding domain with one (R1), two (R2) or three (R3) repeats. The regulation domain contains interaction sites for other proteins in the regulation of gene expression.

In fact, R1R2R3 MYBs, closely related to c-MYB, have been identified in all major lineages of land plants and in the slime mould *D. discoideum* (Otsuka and Van Haastert, 1998). These plant three repeat MYB proteins have been termed pc-Myb (plant c-Myb-like) and were first identified through the *Arabidopsis* genome project (Braun and Grotewold, 1999).

Each repeat in the MYB DNA binding domain contains three regularly-spaced tryptophan residues (Trp) separated by basic amino acids within a single motif. The role of these tryptophans is to provide the fold for the hydrophobic core of the MYB domain to maintain the helix-turn-helix structure of the DNA binding domain (Saikumar *et al*., 1990). The tryptophans are generally conserved in all MYB proteins although the first tryptophan residue in the R3 repeat is substituted by another hydrophobic or aromatic amino acid in most plant R2R3 MYB proteins. Some MYB proteins in fungi and plants lack the typical regularly spaced Trp residues or have partial MYB domains and therefore are non-typical MYB proteins (Jiang *et al*., 2004).

There is greater conservation between the same repeat from different proteins than between R2 and R3 repeats from the same protein suggesting that each repeat has a specialized function in binding DNA. R2 and R3 domains in c-MYB are required for binding to the DNA although it has been proposed that R3 is involved in more specific interactions with the nucleotides of the binding motif than R2 (Ogata *et al*., 1992). The R2R3 MYB proteins are believed to bind DNA in a similar way to c-Myb in contrast to single MYB domain proteins that seem to bind DNA in a different manner, often requiring dimerisation (Jin and Martin, 1999). Differences in the base-contacting residues of this domain change the DNA-binding specifity. This means that MYB proteins may compete for common target motifs although the final activation of transcription will depend on the quantity of different MYB proteins in the nucleus and their abilities to bind DNA and activate transcription as well (Martin and Paz-Ares, 1997).

The MYB domain shows little variation among all R2R3 MYB proteins but their C-termini differ greatly and are only related through the conservation of

small domains which are thought to be responsible for activation and repression of transcription (Jin and Martin, 1999). Most of the reported *MYB* genes are positive regulators of transcription although a few have been characterized as repressors. That is the case for *AmMYB308* and *AmMYB330* (Tamagnone *et al*., 1998), *ZmMYB31* and *ZmMYB42* (Fornalé *et al*., 2006) and *AtMYB4* (Jin *et al*., 2000) which down-regulate expression of some genes of the phenylpropanoid metabolic pathway. While one activation motif is characterized by the sequence L1srGIDPxT/SHRxI/L, a putative negative regulatory domain (NRD) has the motif pdLHLD/LLxi G/S, which has been found only in the R2R3 MYB factors of subgroup 4 (which includes AmMYB308, AmMYB330, ZmMYB42 and AtMYB4) which also contain a conserved domain of unknown function in their C-termini. Many of the divergences in the C-terminal regions appear not to affect the function of the proteins significantly, as shown by divergence between the C-termini of orthologous MYB proteins in different species (Dias *et al*., 2003), although the motifs in that region that are conserved generally reflect similarities in functions, as shown by some subgroups of Arabidopsis MYBs (Kranz *et al*., 1998). Generally, the C-terminal regions of R2R3 MYB proteins appear to be involved in modulating the rate of transcriptional initiation (either by activation or repression).

1.2.2. The MYB family classification

MYB proteins can be classified into three subfamilies depending on the number of adjacent repeats in the MYB domain (one, two or three) and a heterogeneous group referred to as the MYB-related proteins, which usually but not always contain a single MYB repeat (Rosinski and Atchley, 1998; Jin and Martin, 1999; Stracke *et al*., 2001).

In order to categorize the plant R2R3 MYB family, several classifications have been proposed.

The MYB domain from plant R2R3 MYB proteins recognizes different DNA motifs to those recognised by vertebrate MYB proteins. In MYB proteins

identified in plants, three major subdivisions can be made on the basis of the sequence of the DNA-binding domain giving subgroups A, B and C (Romero *et al*., 1998). Mammalian R1R2R3 MYBs such as c-MYB, A-MYB, B-MYB and closely related proteins from invertebrates and cellular slime moulds all bind to the cognate site T/CAACG/TGA/C/TA/C/T (MBSI). Some plant R2R3 proteins can recognise this binding site while others can not. The group of plant R2R3 MYB proteins that bind preferentially to MBSI (group A) are also more closely related in the primary structure of their DNA-binding domains to the c-MYB family. In group B are those plant MYB proteins that bind to MBSI but will also bind to a second site, TAACTAAC (MBSII). This is a small group with only four members in Arabidopsis. The MBSII sequence is the one recognised by the majority of plant R2R3 MYB proteins which encompasses seventy members in Arabidopsis (subgroup C).

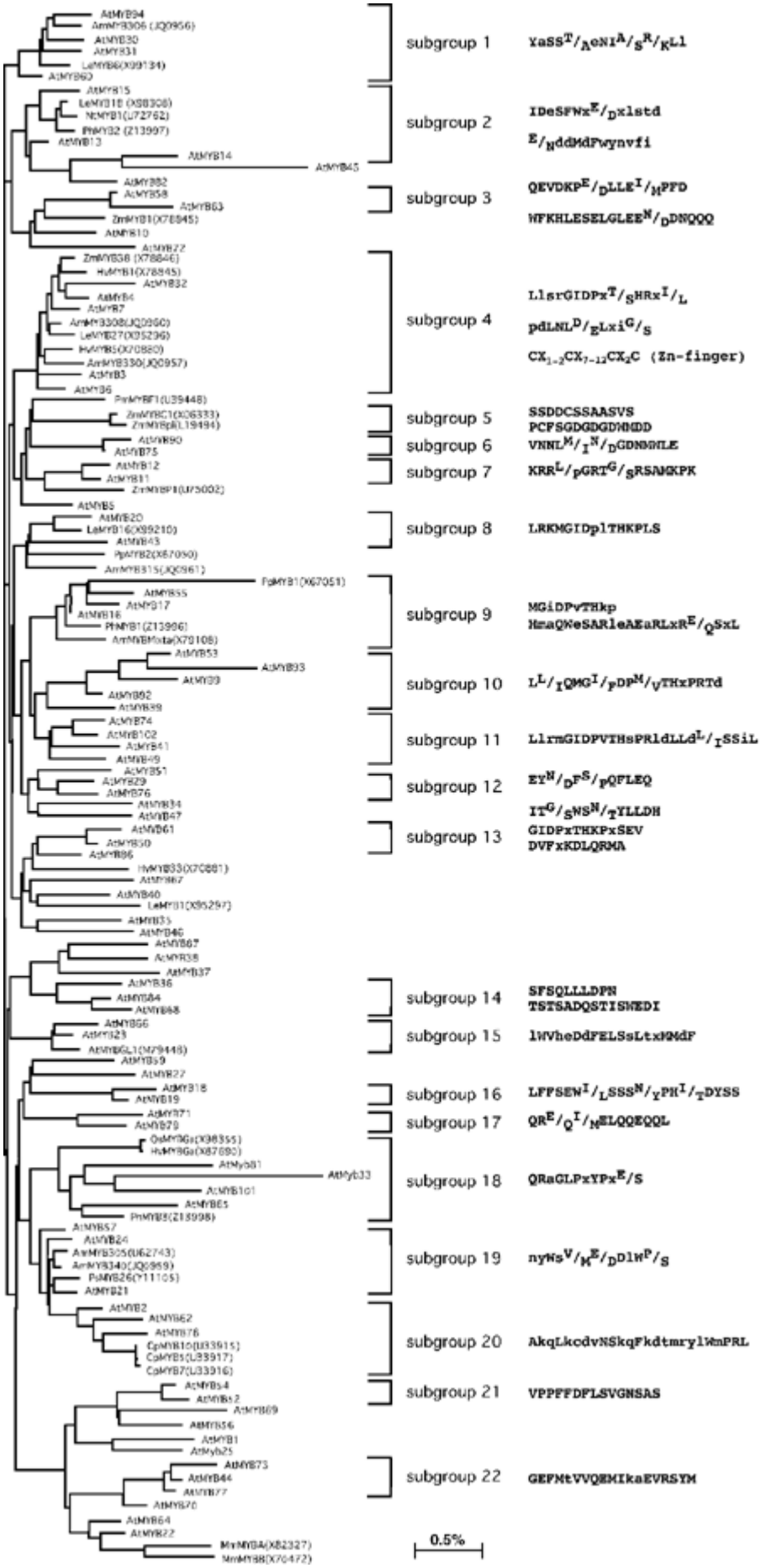
Based on the conserved amino acid sequence motifs in the C-terminal region, the plant R2R3 MYB proteins can be divided into 22 subgroups (Kranz *et al*., 1998). Kranz and colleagues suggested that members from the same subgroup may have similar or redundant functions and the presence of the conserved motifs in the C-terminal domains supports that idea (Fig. 1.4).

1.2.3. MYB evolution

The general evolutionary scenario of the MYB gene family still remains unclear. It is thought that the MYB domain first originated over 1 billion years ago, shortly after the divergence of eubacteria and eukaryotes (Lipsick, 1996).

Some authors believe that R2R3 and R1R2R3 MYB genes co-existed in eukaryotes before the divergence of plants and animals. In fact, the presence of R1R2R3 MYBs in both kingdoms suggests that they had a common ancestor. According to previous studies, most plant MYB genes originated by gene duplications during a period of plant evolution 200 to 500 million years ago. The monophyletic origin of R1, R2 and R3 repeats suggests that these tandem repeats arose as the result of duplication events that occurred prior to the separation of the plant and animal kingdoms (Jiang *et al*., 2004).

Figure 1.4. Phylogenetic relationships of plant *R2R3-MYB* proteins calculated on the basis of 320 amino acid long sequences. The data set (deduced aa sequences) contained 86 *AtMYB* genes and 35 *MYB* genes from other plant species. *Mmmyb-A and B* from *Mus musculus* were included as outgroups. The bar indicates an evolutionary distance of 0.5%. Subgroups of *R2R3-MYB* genes are indicated; members of these subgroups share at least one aa motif in addition to the N-terminal MYB DNA-binding domain. Upper case letters indicate aa's found in all members of a subgroup, lower case letters indicate a conservation in more than 50% of the members. If two amino acids are found at the same position, both are given and these are separated by a slash. From Kranz et al., (1998).


AtMYB94
AmMYB306 (JQ0956)
AtMYB30
AtMYB31
LeMYB6(X99134)
AtMYB60
AtMYB15
LeMYB1B (X88308)
NtMYB1(U72762)
PhMYB2 (Z13997)
AtMYB13
AtMYB14
AtMYB45
AtMYB82
AtMYB58
AtMYB63
ZmMYB1(X78845)
AtMYB10
AtMYB22
ZmMYB38 (X78846)
HvMYB1(X78845)
AtMYB32
AtMYB4
AtMYB7
AmMYB308(JQ0960)
LeMYB27(X95236)
HvMYB5(X70880)
AmMYB330(JQ0957)
AtMYB3
AtMYB6
PhMYBF1(U39448)
ZmMYBC1(X06333)
ZmMYBpl(L19494)
AtMYB90
AtMYB75
AtMYB12
AtMYB11
ZmMYBP1(U75002)
AtMYB5
AtMYB20
LeMYB16(X89210)
AtMYB43
PpMYB2(X87050)
AmMYB315(JQ0961)
FgMYB1(X67051)
AtMYB55
AtMYB17
AtMYB16
PhMYB1(Z13996)
AmMYBMixta(X70108)
AtMYB53
AtMYB93
AtMYB39
AtMYB92
AtMYB89
AtMYB74
AtMYB102
AtMYB41
AtMYB49
AtMYB51
AtMYB29
AtMYB76
AtMYB34
AtMYB47
AtMYB61
AtMYB50
AtMYB96
HvMYB33(X70881)
AtMYB67
AtMYB40
LeMYB1(X95297)
AtMYB35
AtMYB46
AtMYB87
AtMYB38
AtMYB37
AtMYB36
AtMYB84
AtMYB68
AtMYB66
AtMYB23
AtMYBGL1(M79448)
AtMYB39
AtMYB27
AtMYB18
AtMYB19
AtMYB71
AtMYB70
OsMYB3Ga(X98355)
HvMYB3Ga(X87690)
AtMyb61
AtMyb33
AtMYB101
AtMYB65
PhMYB5(Z18998)
AtMYB57
AtMYB24
AmMYB305(U62743)
AmMYB340(JQ0959)
PsMYB26(Y11105)
AtMYB21
AtMYB2
AtMYB62
AtMYB78
CpMYB10(U33915)
CsMYB5(U33917)
CsMYB7(U33916)
AtMYB54
AtMYB52
AtMYB69
AtMYB56
AtMYB1
AtMyb25
AtMYB73
AtMYB44
AtMYB77
AtMYB70
AtMYB64
AtMYB22
MnMYB4(X82327)
MnMYB8(X70472)
subgroup 1
subgroup 2
subgroup 3
subgroup 4
subgroup 5
subgroup 6
subgroup 7
subgroup 8
subgroup 9
subgroup 10
subgroup 11
subgroup 12
subgroup 13
subgroup 14
subgroup 15
subgroup 16
subgroup 17
subgroup 18
subgroup 19
subgroup 20
subgroup 21
subgroup 22
YaSS[T/A]eN[I/A][A/S][R/K]Ll
IDeSFWx[E/D]xlstd
[E/N]ddMdFwynvfi
QEVDKP[E/D]LLE[I/M]PFD
WFKHLESELGLEE[N/D]DNQQQ
LlsrGIDPx[T/S]HRx[I/L]
pdLNL[D/E]Lxi[G/S]
CX1-2CX7-12CX2C (Zn-finger)
SSDDCSSAASVS
PCFSGDGDGDWMDD
VNNL[M/I][N/D]GDNWMLE
KRR[L/p]GRT[G/S]RSNWQPK
LRKMGIDplTHKPLS
MGiDPvTHkp
HmaQWeSARleAEaRLxR[E/Q]SxL
L[L/I]QMG[I/V]pDP[M/T]THxPRTd
LlrmGIDPVTHsPRldLLd[L/I]SSiL
EY[N/D]F[S/P]pQFLEQ
IT[G/S]WS[N/T]YLLDH
GIDPxTHKPxSEV
DVFxKDLQRMA
SFSQLLLDPN
TSTSADQSTISWEDI
lWVheDdFELSsLtxWMdF
LFFSEW[I/L]SSS[N/Y]PH[I/T]DYSS
QR[E/Q][I/N]ELQQEQQL
QRaGLPxYPx[E/S]
nyWs[V/M][E/D]DlW[P/S]
AkqLkcdvNSkqFkdtmrylWnPRL
VPPFFDPLSVGNSAS
GEFMtVVQEMIkaEVRSYM
0.5%

Several models have been proposed to explain how R2R3 and R1R2R3 MYB genes evolved. One model proposes that the R2R3 *MYB* gene family arose by the loss of R1 motif from a *pc-myb*-like gene (R1R2R3) (Braun and Grotewold, 1999). In contrast to this model, Jiang *et al*. 2004 suggested that the ancestral R2R3 MYB was produced by an intragenic domain duplication and subsequently, the ancestral R1R2R3 MYB was formed by a further intragenic domain duplication. This hypothesis is flawed because it suggests that *R1R2R3* genes arose from *R2R3* progenitors, yet *R2R3 MYB* genes are restricted to plants, whereas R1R2R3 MYB genes are present in both animals and plants suggesting an older origin.

The R2R3 MYB family seems to have evolved in plants to regulate specialized plant-specific processes and its expansion occurred before the divergence of monocots and dicots (Martin and Paz-Ares, 1997; Rabinowicz *et al*., 1999; Jiang *et al*., 2004). In fact, with 48 members (*Physcomitrella patens*) it may already have been fairly expanded by the time mosses diverged 400 Mya.

1.2.4. Multifunctionality of R2R3 MYB transcription factors

The precise functions of most of the plant MYB transcription factors are still unknown.They have been reported to be involved in development and cell morphogenesis as, for example the *MIXTA* and *AmMYBML2* genes from *Antirrhinum majus* (Fig. 1.5A), *PhMYB1* from *Petunia hybrida* and *AtMYB16* from *Arabidopsis thaliana* also control petal cell development (Baumann *et al*., 2007). *GL1* (*AtMYB0*) and *GL3* (*AtMYB23*) from Arabidopsis determine trichome fate (Hulskamp, 2004). *WER* regulates root epidermal cell fate (Lee and Schiefelbein, 2002) and *GhMYB109* is necessary for *Gossypium hirsutum* (cotton) fibre development (Pu *et al*., 2008).

An extensive number of R2R3 MYB factors are also activated in defence and stress responses. *AtMYB30* is a positive regulator of programmed cell death (Vailleau *et al*., 2002) and *BOS1* is required to restrict the spread of pathogens (Fig. 1.5B) (Mengiste *et al*., 2003). *OsMYB3R-2* increases tolerance

to cold, drought, and salt stress (Dai *et al*., 2007) as do other MYB genes which are involved in responses to stresses such as *AtMYB15* (cold) (Agarwal *et al*., 2006) or *BcMYB1* (drought) (Chen *et al*., 2005). UV-B light (*AtMYB4, PopMYB134*)(Jin *et al*., 2000; Mellway *et al*., 2009) and nutritional deficiency (*AtPhr2*, *AtNsr1 or CrPsr1*) (Todd *et al*., 2004) are other signal transduction pathways in which plant MYB transcription factors are involved (Fig. 5C).

In hormone signalling pathways, R2R3 MYB transcriptions factors also play important roles. For instance *AtMYB2* and *MYC2* from Arabidopsis are activators of ABA signalling (Abe *et al*., 2003). Gibberellin in conjunction with jasmonate control the growth of stamen filament through the expression of Arabidopsis *AtMYB21*, *AtMYB24* and *AtMYB57* genes (Fig. 5D) (Cheng *et al*., 2009) and these two hormones also regulate, respectively, the expression of *GAMYB* to activate α-amylase gene expression in barley and rice aleurone cells (Gubler *et al*., 1995) and *JAMYB*, involved in inducing host cell death in the hypersensitive response (Lee *et al*., 2007).

Changes in plant metabolism are, in some cases, due to the action of MYB proteins. *AtMYB41* is involved, amongst other functions, in alterations of primary metabolism including the accumulation of amino acids, organic acids, and sugars (Lippold *et al*., 2009). In Arabidopsis, *AtMYB51, AtMYB122, AtMYB28, AtMYB29* and *AtMYB76* are regulators of aliphatic glucosinolate biosynthesis for protection of plants against hervibores (Gigolashvili *et al*., 2009) and atr1D (AtMYB34) and *HAG1/MYB28* are also activators of tryptophan biosynthesis (Bender and Fink, 1998). Without doubt, the most extensively studied biosynthetic pathway in which the MYB transcription factors are involved is phenylpropanoid metabolism, one of the three main types of secondary metabolism in plants involving compounds derived from phenylalanine.

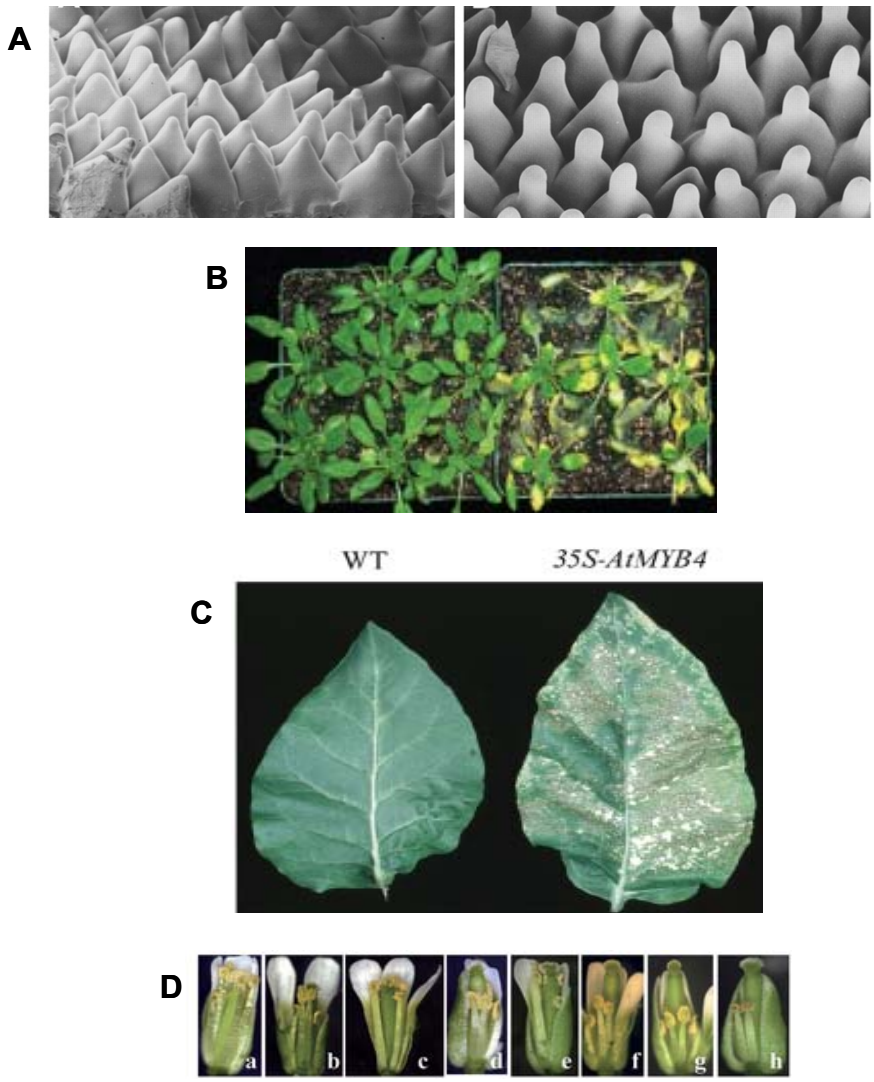


Figure 1.5. Multifunctionality of MYB transcription factors

(A) SEM micrographs of conical inner epidermal cells of wild-type tobacco petal limb (left), and of tobacco petals transformed with *2x35S::AmMYBML2* (right) (Baumann *et al*., 2007). (B) Wild-type (left) and *bos1* (right) plants at 5 days after inoculation with Botrytis (Mengiste *et al*., 2003). (C) Phenotypic effects of *AtMYB4* overexpression in tobacco, overexpression caused the appearance of white lesions on older leaves (Jin *et al*., 2000). (D) Comparison of the stamen phenotype among different *MYB21*, *MYB24*, and *MYB57* Arabidopsis single, double and triple mutant lines (Cheng *et al*., 2009).

Compared to the MYB transcription factor family in vertebrates, (specifically c-MYB that controls the progression of the cell cycle from G1 to S-phase) plant MYB genes are not usually involved in cellular proliferation (Martin and Paz-Ares, 1997). Their role in this function remains still unclear. An R1R2R3 MYB from Arabidopsis was reported to have a conserved function in differentiation and/or cell cycle control (Yanhui *et al*., 2006) but recently, two new R2R3 MYBs, *AtMYB59*, and *DUO1* (*AtMYB125*), have been characterized to control the transition of the cell cycle (review by Cominelli and Tonelli, 2009). *AtMYB59* has a role in controlling root length probably by the activation of *CYCB1;1* a gene coding for a cyclin B, critical in G2/M progression of the cell cycle (Mu *et al*., 2009). Similarly, *DUO1* also regulates *CYCB1;1* to correct male germ line differentiation. These new studies suggest the involvement of other R2R3MYB members in the local control of cell proliferation.

1.2.5. Genetic redundancy in MYBs

Redundancy is a phenomenon widespread in genomes of higher organisms and means that two or more genes encode proteins that perform the same function and that inactivation of one of these genes has little or no effect on the biological phenotype of the organism (Nowak *et al*., 1997). Discovery of high levels of gene duplication suggests that many genes might encode at least partially redundant functions in plants. Redundancy facilitates evolutionary change because of the reduction in the probability of lethal effects by mutations in one of the redundant genes such that new functions may evolve and be subject to natural selection. In the case of transcription factors, redundancy is reflected in the existence of gene families composed of a large number of members encoding proteins with a high degree of homology and similar specificities to bind the DNA.

The R2R3 MYB family in Arabidopsis is one of the largest families of transcription factors described and the extended sequence similarity is due to the highly conserved MYB domain in each member. In this family, the high degree of protein sequence similarity suggests frequent functional redundancy.

For instance, AmMYB305 and AmMYB340, transcription factors from Antirrhinum, are expressed specifically in flowers and compete for the same targets to activate flavonoid metabolism (Moyano *et al*., 1996). Also, in flowers of *A. majus*, the *Rosea1*, *Rosea2*, and *Venosa* genes encode MYB-related transcription factors. Despite their structural similarity, they influence the expression of target genes encoding the enzymes of anthocyanin biosynthesis with different specificities (Schwinn *et al*., 2006). Frequently, different MYB genes have the same/similar functions but have different expression patterns. MYB genes expressed in anthers such as *AtMYB32*, *AtMYB26* and *AtMYB103* in Arabidopsis, *ZmMYBP2* in maize and *NtMYBAS1* and *NtMYBAS2* in tobacco provide an example. Their functions are all related to male fertility but their tissue specific expression patterns are slightly different involving different plant parts (Preston *et al*., 2004).

1.2.6. MYB transcriptional control

Gene-specific regulation of transcription is fundamental for cellular functions. Specificity is provided by the action of transcription factors with activation or repression activities. In eukaryotes, gene expression is frequently mediated by multi-protein complexes. The formation of these complexes involves the combinatorial action of transcription factors that bind conserved promoter elements in precise spatial orientation and on the basis of both specific protein–DNA and protein–protein interactions. This type of transcriptional regulation, termed combinatorial control, is thought to facilitate the complex regulatory networks found in higher eukaryotes (Wolberger, 1999).

For instance, phenylpropanoid metabolism is a complex pathway that leads to thousands of different secondary metabolites. Transcriptional control plays an important role in regulating the overall activity of phenylpropanoid metabolism in response to endogenous (metabolic and hormonal), biotic and abiotic signals. However, this level of control is not, in itself, sufficient to balance the activities of a complex branched metabolic system, particularly where several branch pathways are active. Promoter analysis of genes

involved in phenylpropanoid metabolism reveals the existence of *cis*-regulatory elements such as ACEs (AC-elements) or G-box-like elements, bHLH factor binding sites (R-motifs), and MYB recognition elements whose combinatorial interaction provides specificity in promoter- and stimulus-dependent gene activation process (Weisshaar and Jenkins, 1998). This type of transcriptional regulation facilitates the control of complex regulatory networks (Koes et al., 2005)(Fig. 1.6).

The regulation of pigment production in maize is one of the best-characterized examples in plants of the importance of combinatorial interactions in gene regulation. There is an intimate functional relationship between MYB and bHLH proteins that has been most extensively studied with respect to the phenylpropanoid biosynthetic pathway. MYB and bHLH transcription factors have been studied in petunia, snapdragon and maize as regulators of anthocyanin biosynthesis (Mol *et al*., 1998). The developmental regulation of anthocyanin expression is the result of combinatorial interactions between two distinct families of plant transcription factors. The first is the C1 family, which are MYB-related regulatory proteins and the second is the R family, containing the bHLH (basic helix-loop-helix) motif. Individual family members are not sufficient to induce the anthocyanin biosynthetic genes, but, rather, a member from each family must be coexpressed in a particular tissue for anthocyanin biosynthesis to occur. Similarly, Arabidopsis PAP1 (MYB75) and PAP2 (MYB90), which regulate several enzymes of the anthocyanin biosynthetic pathway, need a bHLH partner (Borevitz *et al*., 2000) which may differ between cell types (GL3, EGL3 and TT8). The combinatorial action of MYB and BHLH proteins has also been demonstrated through the molecular analysis of the *transparent testa* mutants where several steps are controlled in combination by MYB and bHLH transcription factors. This is the case for TT8, a bHLH protein and TT2, a R2R3-MYB which regulate expression of DFR and BAN genes in Arabidopsis and control the production of proanthocyanins in the seed coat. Proanthocyanins are another product of phenylpropanoid metabolism (Nesi *et al*., 2000).

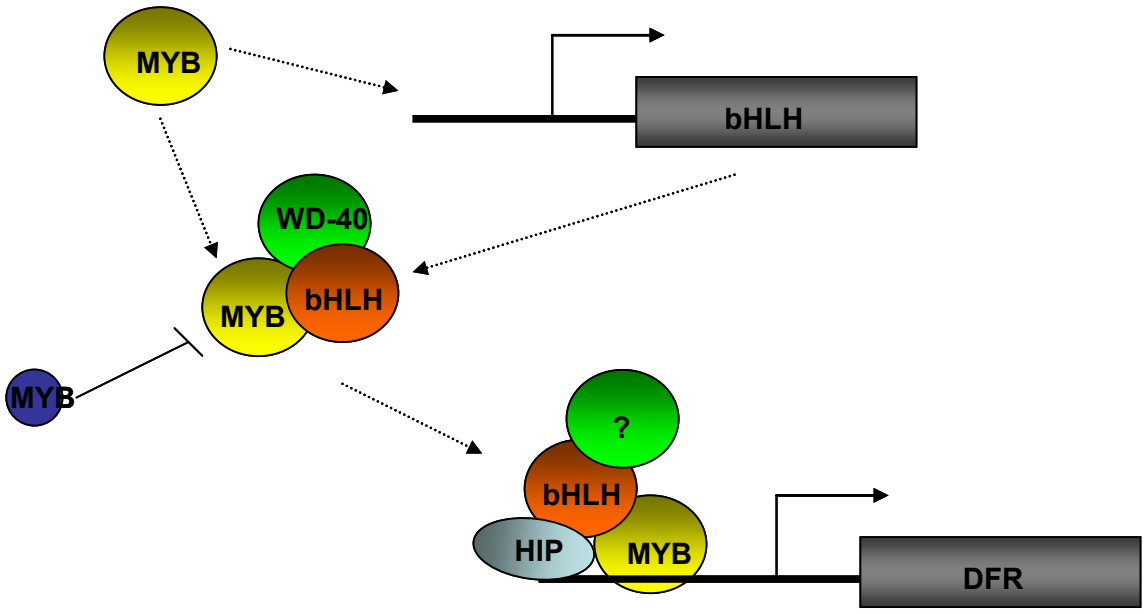


Figure 1.6. MYB protein involvement in the MYB–bHLH–WD40 (**MBW) complex for transcriptional control of anthocyanin, proanthocyanin and cell fate determination in Arabidopsis**

Model for the transcriptional activation of a structural pigmentation gene dihydroflavonol-4-reductase (DFR). Proteins (circles) are coloured. Genes, including upstream sequences are indicated by rectangles. MYBs control transcription of bHLH factors, and subsequently form a complex that also involves a WD40 repeat protein. The MYB factors contact the DNA directly, whereas the bHLH proteins probably bind indirectly via a hypothetical bHLH INTERACTING PROTEIN (HIP). The small R3-MYB acts as an inhibitor, probably by sequestration of the HLH protein into an inactive complex. Figure adapted from Koes, R. 2005.

The WDR (WD-repeat) proteins comprise a diverse superfamily of regulatory proteins characterized by the WD motif (also known as the Trp-Asp or WD40 motif) (van Nocker and Ludwig, 2003). The additional requirement for a WDR protein in the WD40–bHLH–MYB transcription complex that controls anthocyanin biosynthesis has been shown in several plant species (Arabidopsis, Petunia, maize) which indicates that interactions between these kinds of transcription factors have been conserved throughout the Plant Kingdom. The genes *AN1* and *AN2* encode bHLH and R2R3-MYB proteins, respectively. AN2 interacts with AN1 and JAF13 (which encodes another bHLH protein) to activate the expression of anthocyanin biosynthetic genes in petunia (Spelt *et al*., 2000). In addition to the AN1 and AN2 proteins, the cytoplasmic WDR protein, AN11, is also required for anthocyanin accumulation, probably by stabilising the transcriptional complex bound to the promoter of the target genes (de Vetten *et al*., 1997).

Combinatorial interactions between a set of very similar regulatory proteins to those controlling anthocyanin biosynthesis influence epidermal cell identity in Arabidopsis as exemplified in root epidermal cell fate specification, where a cascade of several transcription factors plays an important role. WEREWOLF (Moore et al.) is a MYB transcription factor that induces transcription of *CPC* (*CAPRICE)* to promote root hair cell fate but also controls coordinately GLABRA2 (GL2) which encodes a homeodomain-leucine zipper (HD-Zip) protein which promotes non-hair cell fate (Ryu *et al*., 2005). The transcription of GL2, which is thought to act farthest downstream in the root hair regulatory pathway, is controlled by a protein complex that includes WER, GL3, ENHANCER OF GLABRA3 (EGL3), and TRANSPARENT TESTA GLABRA1 (TTG1) proteins. GL3 and EGL3 encode basic helix-loop-helix (bHLH) proteins and promote non-hair cell differentiation in a redundant manner. The CPC protein (which is a 1R MYB protein) has been proposed to disrupt this protein complex by competitive binding to bHLH proteins with WER, leading to repression of *GL2* expression and therefore determining root hair cell fate (Tominaga *et al*., 2007).

1.3. Phenylpropanoid metabolism

Plants synthesize a wide variety of natural products based on the phenylpropane skeleton from phenylalanine. As a major component of specialized plant metabolism, phenylpropanoid biosynthetic pathways provide anthocyanins for pigmentation, flavonoids such as flavones, for protection against UV photodamage, signalling molecules (various flavonoids and isoflavonoids are inducers of Rhizobium nodulation genes), polymeric lignin for structural support and assorted antimicrobial phytoalexins as plant defences (Liang *et al*., 1989). A variety of phenylpropanoid compounds are present in plants but genes homologous with those encoding enzymes of the phenylpropanoid biosynthetic pathways have also been found in bacteria and fungi (Emiliani *et al*., 2009).

1.3.1. Biochemistry of phenylpropanoid metabolism

Phenylpropanoid metabolism starts initially with phenylalanine, supplied by the shikimate pathway which has a crucial role in the regulation of the phenylpropanoid pathway (Bentley and Haslam, 1990). The shikimate pathway is initiated by condensation of phosphoenol pyruvate and erythrose 4-phosphate followed by several reactions to yield shikimate. The condensation of shikimate results in the formation of chorismate, where the pathway bifurcates to produce tryptophan, phenylalanine or tyrosine (Fig. 1.7). Phenylpropanoid biosynthesis is initiated by the deamination of phenylalanine to give cinnamic acid, catalyzed by phenylalanine ammonia-lyase (PAL). Successive hydroxylations of cinnamic acid to give p-coumaric acid and caffeic (3,4-dihydroxycinnamic) acid are followed by the formation of hydroxycinnamoyl CoA thioesters, which are the substrates for branch pathways for the production of lignin, flavonoids, benzoic acid, salicylic acid, coumarins, and simple esters. Since PAL has a metabolically important position linking the phenylpropanoid pathway to primary metabolism, the regulation of overall flux into phenylpropanoid metabolism has been suggested to be modulated by this protein as a rate-limiting enzyme (Noel *et al*., 2005; Ferrer *et al*., 2008).

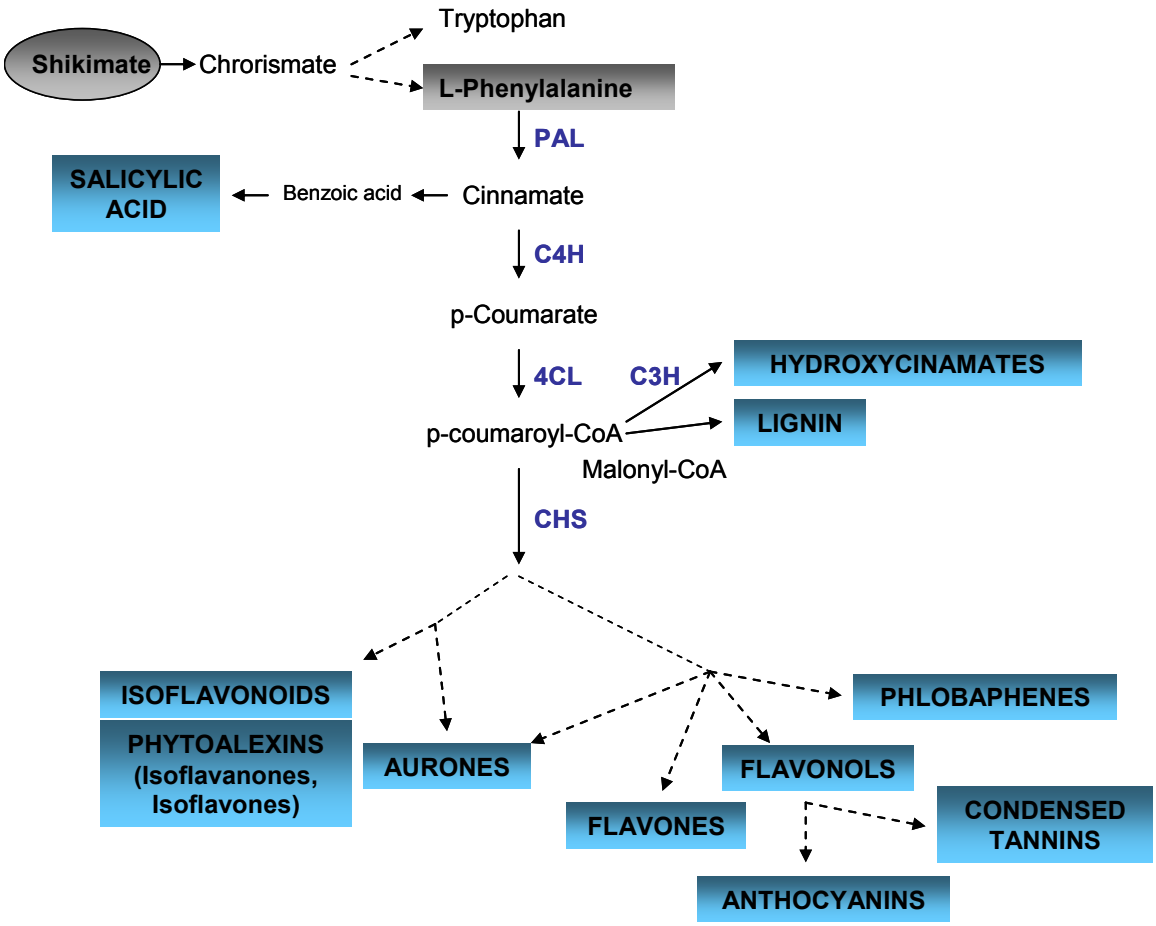


Figure 1.7. General phenylpropanoid metabolic pathway. Enzymes are indicated in blue uppercase letters. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; C3H, p-coumarate 3-hydroxylase; 4CL, 4-coumarate-CoA ligase; CHS, chalcone synthase. Dashed lines arrows indicate multiple enzymes involved in that step. Blue boxes show final products of the different branches of the phenylpropanoid metabolism.

Due to its importance, PAL is the general enzyme from the phenylpropanoid pathway whose structure and function have been most extensively studied in plants and bacteria. The other two enzymes in the general phenylpropanoid pathway are C4H (cinnamate-4-hydroxylase) and 4CL (4-coumarate-CoA ligase). The general phenylpropanoid pathway serves to provide precursors of various secondary branches for the synthesis of compounds including monolignols (lignin monomers), flavonols, anthocyanins, proanthocyanins, salicylic acid and other derived metabolites (coumarins, stilbenes, sinapoyl esters *etc*).

The enzymes of the phenylpropanoid metabolism may be organized as multiprotein enzyme complexes that channel the flux of substrate through specific subpathways. The phenylpropanoid pathway is located in the cytoplasm, however the shikimate pathway, source of the phenylalanine, is believed to be located in plastids (Winkel-Shirley, 1999).

The phenylpropanoid pathway can be activated by multiple abiotic/biotic factors (Dixon and Paiva, 1995). Phenolic compounds are induced in plants by pathogen attack (stilbenes, flavonols, salicylic acid). Many phenylpropanoid compounds are produced in response to wounding or to feeding by herbivores (coumarins, chlorogenic acid, kaempferol). Anthocyanins and flavones increase in response to high visible light levels whereas UV irradiation induces flavonoids (particularly kaempferol derivatives) and sinapate esters in Arabidopsis or isoflavonoids and psoralens in other species. Levels of anthocyanins also increase following cold and nutritional stress (such as phosphate limitation) and this has been suggested as a mechanism for protecting the photosynthetic apparatus under stress conditions. Low nitrogen induces flavonoid and isoflavonoid nod gene inducers and chemoattractants for nitrogen-fixing symbionts whereas low iron levels can cause increased release of phenolic acids.

1.3.2. Flavonoid biosynthetic pathway

Flavonoids constitute a diverse family of aromatic molecules that are derived from the amino acid phenylalanine. There are more than 6000 different flavonoids described that can be grouped into different classes. These compounds include, in higher plants: chalcones, flavones, flavonols, flavandiols, anthocyanins, condensed tannins (or proanthocyanidins) and aurones, which are widespread, but not ubiquitous. Some plant species also synthesize specialised types of flavonoids, such as the isoflavonoids which are made by legumes (for review, see Lepiniec *et al*., 2006).

These plant secondary metabolites have important developmental and physiological functions in the control of flower and fruit pigmentation, auxin transport, signalling to symbiotic microorganisms, UV protection, pathogen defence, pollen tube germination, pollinator attraction and also serve as important micronutrients in human and animal diets.

Chalcone synthase is the first enzyme committed to flavonoid biosynthesis (Fig. 1.8). Curiously, the chalcone synthase gene of *Petunia* is famous for being the first gene for which the phenomenon of RNA interference was observed (Napoli *et al*., 1990). CHS catalyzes the condensation of three acetate residues from malonyl-CoA with *p*-coumaroyl-CoA to form naringenin chalcone. Then the chalcone is isomerized by chalcone isomerase (CHI) to form the flavonone naringenin, which is converted to dihydrokaempferol by flavonone 3-hydroxylase (F3H). Subsequently, the dihydrokaempferol is further hydroxylated by flavonoid 3′-hydroxylase (F3′H) or flavonoid 3′, 5′-hydroxylase (F3′5′H) to form the dihydroflavonols dihydroquercetin (DHQ) and dihydromyricetin (DHM), which are required for the production of cyanidin and delphinidin-based anthocyanins respectively. These molecules are converted into anthocyanins by at least three steps involving the enzymes dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), and UDP-glucose:flavonoid 3-O-glucosyl-transferase (UFGT).

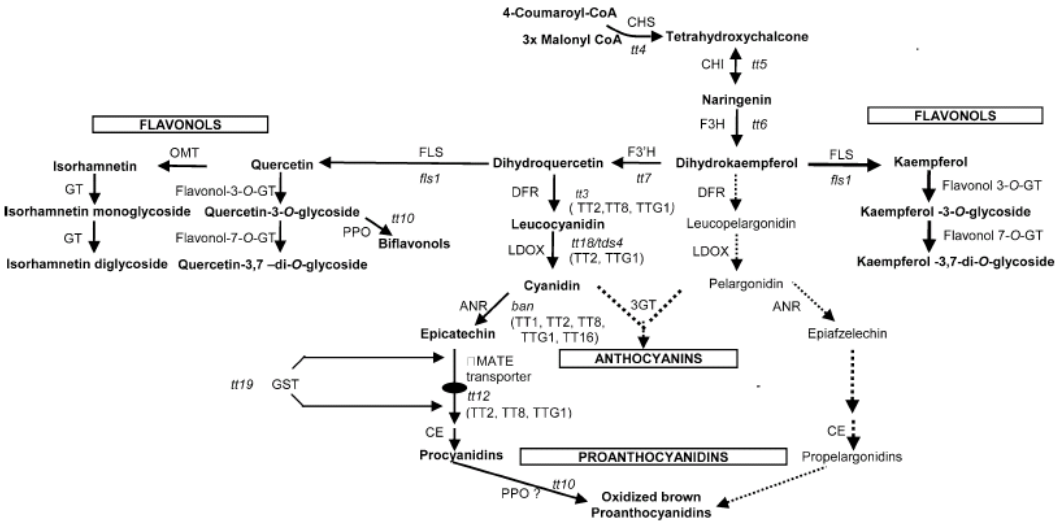


Figure 1.8. Scheme of the flavonoid biosynthetic pathway in Arabidopsis seed

The different steps leading to the formation of flavonoids in Arabidopsis seed are indicated by boldface arrows. The pathways to propelargonidins and anthocyanins that are not present in wildtype seed are indicated with dashed lines. Where mutants are available, for respective enzymatic steps, they are indicated in lowercase italic letters. Regulatory proteins are given in brackets beside their target genes that are shown in capital letters. ANR anthocyanidin reductase, CE condensing enzyme, CHI chalcone isomerase, CHS chalcone synthase, DFR dihydroflavonol-4-reductase F3H flavonol 3-hydroxylase, F3'H flavonol 3'-hydroxylase, FLS flavonol synthase; GST glutathione S-transferase; GT, glycosyltransferase, LDOX leucoanthocyanidin dioxygenase; OMT methyltransferase. From Routabul JM, 2006.

Branches of phenylpropanoid metabolism are regulated by interplay between branch-specific activating and repressing MYB transcription factors, some of which depend on specific bHLH protein partners, although a few additional genes regulating the expression of flavonoid biosynthesis genes have been identified. Two such regulators are the Petunia AN11 protein (de Vetten *et al*., 1997) and Arabidopsis transparent testa glabra1 (TTG1) (Walker *et al*., 1999). AN11 and TTG1 are highly similar, and both contain four WD repeat motifs. The WDR protein interacts with the bHLH proteins which interact with the MYB protein to form a regulatory complex.

Elucidation of the flavonoid pathway has been extensively studied since a collection of mutant lines defective in flavonoid biosynthesis were identified in Arabidopsis on the basis of altered seed coat color. They are known as the *transparent testa* or *tt* mutants (Koornneef, 1990; Shirley *et al*., 1995) because the phenotype involves a reduction or absence of pigments in the testa (seed coat) . Seeds from these mutants are yellow or pale-brown in color (Arabidopsis seeds are brown) because they fail to produce brown proanthocyanidins or condensed tannins in the seed coat. Although the first flavonoid mutant was not discovered by Koorneef and colleagues, they were responsible for the isolation and characterization of numerous additional *tt* mutants. The loci affected in these mutants, encode structural proteins (DFR, CHS, CHI, F3H, LDOX, FLS1), regulatory proteins (TT1, TT2, TT8, TT16, TTG1, and TTG2), and proteins that are probably involved in flavonoid compartmentation (TT12, TT19, and AHA10) (some of them represented in Fig.8 and 9A) (reviewed by Lepiniec *et al*., 2006). Although extensively studied, the *transparent testa* mutants have little information available on specific architectural phenotypes resulting from flavonoid level perturbations. New phenotypes in these mutants related to root development, hypocotyl length, inflorescence production *etc*., have been recently described (Buer and Djordjevic, 2009). One other mutant, named *banyuls* (Bang et al.) (Fig. 9B) because of the similarity with the colour of a local wine, affects the colour of the seed coat but differently to the *tt* mutants; the seeds

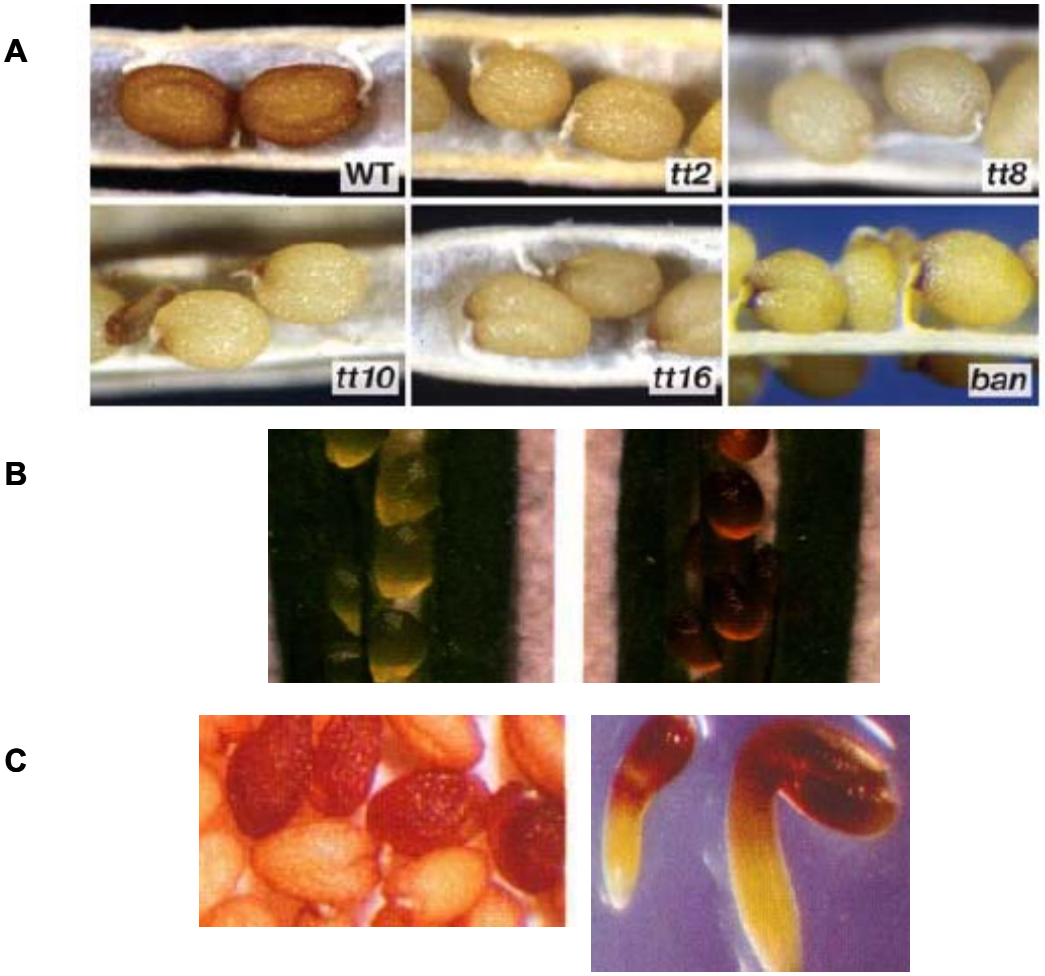


Figure 1.9. Arabidopsis flavonoid mutants

(A) *Arabidopsis* seed phenotypes in wild-type and some *transparent testa* mutants. (B) Immature siliques showing seeds with the wild-type phenotype (left) and the *ban* phenotype. (C) Anthocyanin accumulation in *fusca* mutants; left, dry seeds from a plant heterozygous for *fus* and right, *fus* phenotype in different stages of development.

are darker because mutation of ANR (*banyuls*) causes an accumulation of red anthocyanin in the seed coat (Sylvie *et al*., 1997). *Fusca* genes also have roles in flavonoid biosynthesis because the *fusca* (*fus*) (Fig. 1.9C) mutation induces ectopic accumulation of anthocyanins, not in the seed coat but in immature embryos and developing seedlings (Castle and Meinke, 1994).

Flavonoids are likely synthesized by a cytoplasmic complex resulting in the production of an enormous variety of products that are then transported into the vacuole (Winkel-Shirley, 2001). The tissue-specific localization of flavonoids has been demonstrated in Arabidopsis seedlings. Quercetin, kaempferol and naringenin chalcone were shown to be concentrated in three tissues: the upper hypocotyl, the hypocotyl-root transition zone and the distal elongation region of the root. Kaempferol in the epidermis localized in the nuclear region and plasma membrane whereas quercetin localized subcellularly in the nuclear region, in the cortex (Murphy *et al*., 2000; Peer *et al*., 2001). Long-distance movement of secondary metabolites is largely unexplored. According to Peer *et al*., (2001), flavonoids are synthesized in the cells in which they accumulate and serve local functions, although recently several studies have suggested that flavonoids may move from their site of synthesis (Winkel-Shirley, 2001; Buer *et al*., 2007).

1.3.3. Metabolic engineering and applications of phenylpropanoids

The chemical nature of phenylpropanoids has been studied extensively. Recent advances in genomics, proteomics, and metabolomics provide new approaches to define the roles of phenylpropanoids in plant development and explore their potential application in agriculture and medicine. Genetic engineering can be applied to manipulate levels of secondary metabolites of economic value as well as those of potential importance in diet and health (Fig. 1.10).

Figure 1.10. Phenylpropanoids in nature and engineering

From above left, visualisation of capsidiol, a phytoalexin produced during the tobacco defence response (right hand image shows leaf viewed under UV light); right, wild-type (top) and transgenic purple tomatoes (bottom) rich in anthocyanins. Below, from left to right, engineered Petunia showing a new orange colour, pelargonidin 3-rutinoside not normally produced in flowers of Petunia; products from black grapes such as red wine are natural sources of antioxidants such anthocyanins; resveratrol pills as a dietary supplement (Images from Google).

1.3.3.1. Importance of the phenylpropanoid pathway in Agriculture

Engineering plant growth and development has potential applications in agriculture such as reducing the growth periods required for development of particular organs, such as fruits or favouring other aspects of plant development. Changing auxin distribution using flavonoid signals, either through genetic engineering or by manual application of flavonoids, could allow modulation of growth and development in diverse plant species. For instance, introduction of the UDP-glucuronosyltransferase-encoding gene from pea (*PsUGT1*) into the alfalfa genome resulted in modified auxin levels that caused diverse changes in the regulation of the cell cycle, gene expression, growth and life cycle indicating that such genes are an attractive target for improved crop production (Woo *et al*., 2003).

Phytoalexins are secondary metabolites that act as defence chemicals produced by plants in response to bacterial and fungal pathogens and also in response to abiotic stresses. Engineering novel phytoalexins in heterologous plants provides a unique strategy to enhance plant innate defence systems.

For example, heterologous expression of *stilbene synthase* (*STS*) resulted in accumulation of the phytoalexin resveratrol in tobacco and alfalfa and improved disease resistance toward some fungal pathogens. These results indicate that boosting crop phenylpropanoid phytoalexin production can reduce the use of chemical fungicides/insecticides (Richard *et al*., 2002).

Lignin composition is a problem for forage digestibility in ruminant animals and also in the forestry industry. Chemical treatments for removal of lignin are both costly and potentially polluting (Dixon *et al*., 1996). Downregulation of cinnamyl alcohol dehydrogenase (CAD) in alfalfa altered lignin composition improving its digestibility, while reduction of bi-specific caffeic acid/5-hydroxyferulic acid *O*-methyltransferase (COMT) activity in poplar changed lignification (Van Doorsselaere *et al*., 1995; Baucher *et al*., 1999) rendering the polymer more readily delignified and thereby rendering the plant more readily pulped for paper manufacture or digested. Increasing the extractability of lignin from cell walls is of vital importance in the development of

second generation biofuels, so that polysaccharide cell wall material is more readily available for fermentation.

1.3.3.2. Biotechnology

Other phenylpropanoid compounds have a great influence not only on the yield and quality of many crops, but also for their commercial appeal.

It is possible to change flower colour through the metabolic engineering of flavonoid biosynthesis. Orange petunia flowers have been produced from white petunia by introducing a maize gene encoding dihydroflavonol 4-reductase (DFR), which converts dihydrokaempferol into the substrate for pelagonidin, an orange pigment (Meyer et al., 1987). Violet flowers of Dianthus were generated from a white flowering cultivar by introducing petunia DFR and flavonoid 3´,5´-hydroxylase (F3´5´H) the endogenous enzyme of the plant (Forkmann and Martens, 2001). From these studies it can be concluded that through a combination of biochemistry, breeding and genetic engineering it is possible to generate unique flower colours in a cultivar with commercial potential.

The phenylpropanoid pathway gives rise to metabolites that determine some floral fragances. ODORANT1, a member of the R2R3-type MYB family, is a candidate for the regulation of production of volatile benzenoids in *Petunia hybrida* (Verdonk *et al*., 2005). Pigment1 (*Pap1*), another MYB transcription factor from Arabidopsis has been expressed in petunia flowers and showed an increase of up to tenfold in the production of volatile phenylpropanoid/ benzenoid scent compounds, but also enhanced the production of colour (Ben Zvi *et al*., 2008). The knowledge of factors activating the production of scent creates a new biotechnological strategy for the metabolic engineering of fragrance and also allows the development of novel genetic variability for breeding purposes because fragrances produced by plants attract pollinators.

1.3.3.3. Medicine, health and diet

Dietary flavonoids have biological activities in promoting health. Genistein may perturb the process of phosphorylation/dephosphorylation of tyrosine residues of Cdc2 kinase, leading to cell cycle arrest at G2–M suggesting that dietary genistein could help treat cancer. Genistein may also work as a topoisomerase II inhibitor to arrest the cell cycle at G2. In contrast, other flavonoids such as the flavones luteolin, and daidzen, which are structurally similar to estrogen (phytostrogens) arrest the cell cycle at G1 (Bai *et al*., 1998). Quercetin inhibits the expression of specific oncogenes and genes controlling the cell cycle and also up-regulates the expression of several tumor suppressor genes (Nair *et al*., 2004). Flavopiridol inhibits cell cycle progression, induces apoptosis, and inhibits transcriptional activity of RNA polymerase II (Dai et al., 2003)

Anthocyanins offer protection against certain cancers, cardiovascular disease and also have anti-inflammatory activity, help boost eyesight, and may help stave off obesity and diabetes. Engineered tomatoes expressing two transcription factors from *Antirrhinum*, a basic helix-loop-helix and a MYB-related transcription factor, Del (bHLH) and Ros1 (MYB) respectively, raised the anthocyanin levels in fruit higher than anything previously achieved, giving them an intense purple colour. Tests on mice susceptible to cancer showed that animals whose diets were supplemented with the purple tomatoes had a significantly longer lifespan than those who received diets supplemented with normal red tomatoes (Butelli *et al*., 2008). This kind of metabolic engineering, offers the potential to promote health through diet by reducing the impact of chronic diseases.

Other flavonoids and anthocyanins are strong antioxidants; epigallocatechin gallate and related catechins in green tea have been shown to reduce cardiovascular disease and to have a role in limiting obesity and have effects in slowing the progression of prostate cancer (McKay and Blumberg, 2007) . Isoflavones are predominantly produced by legumes and have been linked with alleviating postmenopausal symptoms and reducing breast and

colon cancers by consumption of soybeans (Burke *et al*., 2000). Furthermore, resveratrol, which is found in red wine, has been shown to function not only as an antioxidant, but also as a regulator of metabolism and longevity (Halls and Yu, 2008)

It is reported that ferulic acid can inhibit thrombosis and blood platelet clotting, chlorogenic acid has antimicrobial and anticoagulant properties, and caffeic acid can be used as an antimicrobial agent and immunization accelerant (Li *et al*., 2007).

As constituents of plant-rich diets and an assortment of herbal medicinal agents, dietary phenylpropanoids have been shown to have considerable health benefits. There is increasing the awareness in the medical community and the public at large as to the potential dietary importance of these plant-derived compounds. Detailed knowledge of these enzymes of specialized metabolism provides a basis for the enzyme and metabolic engineering of production platforms for diverse novel compounds with desirable dietary and medicinal properties. Moreover, genetic manipulation of microorganisms has also been used to engineer the phenylpropanoid pathway. Microbial biotechnology applications for flavonoid or stilbenoid targets cause the bacterial strains or yeast to use 4-coumarate-CoA ligase for the production of resveratrol in *S. cerevisiae* or flavone synthase expression in *E. coli* to allow the biosynthesis of flavone derivatives (Filippos *et al*., 2007).

1.3.4. Phenylpropanoid evolution

The first land plants had to adapt to a wide array of new environmental challenges including desiccation, varying temperatures, and increased UV radiation. The radiation of land plants was marked by a long list of morphological innovations and adaptations, and among them the phenylpropanoid pathway which was crucial. Phenolic compounds played a critical role in evolution of land plants protecting plants from UV radiation, providing lignin for vascularization, mechanical support and water proofing and

colours to enhance reproductive strategies through pollinator attraction (Waters, 2003).

While charophycean algae, the pioneer green algal ancestors of the land plants, possess some phenolic compounds, they lack the diversity of phenolic compounds and pathways found in embryophytes (Raven, 2000). The "early" steps in the pathway are found first in the bryophytes (mosses). In fact, enzymes from the phenylpropanoid pathway have been characterized in mosses including chalcone synthase and 4-coumarate-CoA ligase (4CL) in the moss *Physcomitrella patens* (Jiang *et al*., 2006; Silber *et al*., 2008) and CHS in the liverwort *Marchantia paleacea* (Harashima *et al*., 2004). In addition, in the descendant of the most primitive vascular plants, the pteridophyte *Psilotum nudum*, genes encoding several isoforms of the CHS family have been cloned and a CHS was also cloned from the green sprouts of *Equisetum arvense* (horsetail) which is one of the most ancient living vascular plant taxa, second only to the Psilotaceae (Yamazaki *et al*., 2001). Phenylalanine ammonia-lyase (PAL), the enzyme which catalyzes the first step of the phenylpropanoid pathway, has been extensively characterized in all land plant lineages, including the early emerging bryophytes, fungi, cyanobacteria and algae (Emiliani *et al*., 2009). These discoveries have given rise to considerable speculation that phenylpropanoid biosynthetic systems in plants have been built by successive recruitment of enzymes from primary metabolism and/or as the product of a cooperation of a green alga and a fungus-like organism (Jorgensen, 1993).

Although phenylpropanoid metabolism is being extensively studied, there is, as yet, little information about the pathway during its first 400–700 million years of existence despite its great importance for the radiation of plants in terrestrial environments.

CHAPTER 2

Chapter 2: Characterization of two Arabidopsis MYB transcription factors

2.1. Introduction

Abiotic stresses caused by drought, salinity, temperature extremes *etc*., are among the major constraints to plant growth and crop production worldwide; a better understanding of the genetic and biological mechanisms underpinning stress adaptation and response is needed to improve crop productivity and to underpin food security.

The *hos* (for high expression of osmotically responsive genes) mutants belong to a collection of Arabidopsis mutants that was assembled for the identification of genes involved in responses to low temperature, drought, salinity, and ABA with the aim of creating a model for stress signalling in higher plants (Ishitani *et al*., 1997). The *hos10* mutant was characterized as affecting a gene encoding an R2R3-type MYB transcription factor (*AtMYB8*) involved in dehydration stress tolerance and cold acclimation response. The reported mutant had a clear, strong phenotype showing a reduced size and fertility and was hypersensitive to cold and NaCl treatments (Zhu *et al*., 2005). Several other MYB transcription factors have been described as involved in the response mechanisms to different abiotic stresses (Urao *et al*., 1993; Denekamp and Smeekens, 2003; Agarwal *et al*., 2006). The gene *AtMYB6* (At4g09460) is, phylogenetically, the closest gene to *AtMYB8* (At1g35515) as shown by phylogenetic analysis of their encoded proteins which group in subfamily IV of the MYB transcription factor (Kranz *et al*., 1998; Stracke *et al*., 2001) (Fig.1A). The phylogenetic position of *AtMYB6* with respect to *AtMYB8,* which is based on the high degree of structural similarity of the encoded proteins, led us to hypothesize that they could have redundant functions. Deciphering whether or not the *AtMYB6* and *AtMYB8* proteins share any common functions was one of the main goals of this project.

In this chapter, the approach taken for comparing the functions of *AtMYB6* and *AtMYB8* is described. The first step was to test if alleles of the genes encoding these two proteins shared a common phenotype with *hos10* and responded similarly to different abiotic stress treatments to establish whether there is redundancy in their functionality or whether their roles diverged during evolution.

Preliminary results indicated that even the *AtMYB8* alleles we analyzed were not phenotypically related to *hos10* although mutations in different genetic backgrounds (ecotypes) were being compared.

2.2 Materials and methods

2.2.1. Plant material

Arabidopsis T-DNA insertion lines were obtained from the Nottingham Arabidopsis Stock Center, NASC (arabidopsis.org.uk/) or Max Planck Institute (www.gabi-kat.de/). The *hos10-1* mutant was kindly provided by Dr. Ray A. Bressan from Purdue University, Indiana (Table 1).

Arabidopsis thaliana accessions Columbia (Col-0), C24 and mutants were grown in soil at 22ºC (following vernalization at 4ºC for 1 week) in a growth room under short days (8h light/16h dark) for 3-4 weeks and then transferred to long days (16h light/8h dark).

For seedling experiments, seeds were surface-sterilized using a sterilization buffer (70% ethanol)[w/v] and SDS 0.05% [w/v]) for 3 min and rinsed with 100% ethanol for 1 min. Afterwards, seeds were air-dried in the flow hood and plated on Murashige and Skoog nutrient medium (MS) solidified with 0.8% (w/v) agar and 3% sucrose unless indicated otherwise. Plates containing seeds were placed in a cold room for 2-3 days at 4ºC in darkness for stratification. Treatments on plates were undertaken in a growth chamber 23ºC±1ºC under long days (16hr light/8 hr dark cycles).

Genes	Lab No.	Line	Mutants	Stock Centre	T-DNA	Polymorphism	Marker	Ecotype
	ARA 168	MYB6-KO#4	Salk_075174	NASC	pROK2	Exon	Km	Col-0
AtMYB6	ARA 169	MYB6-KO#1	Salk_126058	NASC	pROK2	Exon	Km	Col-0
At4g09460	ARA 170	MYB6-KO#2	AL936127	Max Planck	pAC106	Exon	Sulf	Col-0
	ARA 171	MYB6-KO#3	Salk_074789	NASC	pROK2	Exon	Km	Col-0
	ARA 172	MYB8-KO#3	Salk_069625	NASC	pROK2	Intron	Km	Col-0
	ARA 173	MYB8-KO#2	BX290224(F5K24)	Max Planck	pAC161	Intron	Sulf	Col-0
AtMYB8	ARA 258	MYB8-KO#1	Salk_088230	NASC	pROK2	Intron	Km	Col-0
At1g35515	ARA286	MYB8-KO#4	GK-036E11	NASC	pAC161	Intron	Sulf	Col-0
	ARA287	MYB8-KO#5	GK-062B07	NASC	pAC161	Intron	Sulf	Col-0
	HOS10	HOS10	HOS10	NASC	pSKI015	Intron	Basta	C24

Table1. T-DNA mutants lines used in this project

List of mutant lines genotyped. Sulf = sulfadiazine. Mutants highlighted in grey were discarded after genotyping.

2.2.2. Genotyping T-DNA lines

The presence of T-DNA insertions in the different alleles of *AtMYB6* and *AtMYB8* was confirmed by PCR using gene-specific primers and a T-DNA left border (LB) primer (Table 2). PCR amplification was initiated with denaturation at 94ºC for 2 min, followed by 35 cycles of 94ºC for 15 s, 55-57ºC for 30 s, 72ºC for 1.30 min, and an extension at 72 ºC for 5 min (for more detail see General Materials and Methods section 6 and 12).

The presence or absence of bands was used to distinguish between hemizygous, homozygous and wild type seedlings.

Genes	Lab No.	Line	Wt primers	T-DNA insertion
	ARA 168	MYB6-KO#4	SJS1-SJS2	SJS1-LBSalk
AtMYB6	ARA 169	MYB6-KO#1	SJS1-SJS2	SJS1-LBSalk
At4g09460	ARA 170	MYB6-KO#2	SJS1-SJS2	SJS1-LBGabi
	ARA 171	MYB6-KO#3	SJS1-SJS2	SJS1-LBSalk
	ARA 172	MYB8-KO#3	SJS9-SJS10	SJS9-LBSalk
AtMYB8	ARA 173	MYB8-KO#2	SJS7-SJS10	SJS7-LBGabi
At1g35515	ARA 258	MYB8-KO#1	SJS7-SJS10	SJS10-LBSalk
	ARA 286	MYB8-KO#4	SJS8-SJS9	SJS8-LBGabi
	ARA 287	MYB8-KO#5	SJS8-SJS9	SJS26-LBGabi
	HOS10	HOS10	SJS8-SJS9	xxxx

Table 2. Oligonucleotide primers used for genotyping the T-DNA mutants

2.2.3. Segregation analysis

Seeds from T2 hemizygous plants (nomenclature used by the Gabi-Kat database/project http://www.gabi-kat.de/) were used for a segregation analysis to find homozygous mutant plants with a single T-DNA insertion for each of these alleles. The analyses were done with populations of 60 sterile seeds from individual T2 plants (T3 generation). Sterile seeds were sown in 90-cm Petri dishes, MS 0.8% agar, with the appropriate antibiotic depending on the line (Table1), 50 µg/ml kanamycin or 7.5 µg/ml sulfadiazine. The imbibing seeds were chilled for 2-3 days at 4ºC in darkness for stratification. Then, Petri dishes were placed at 23ºC in long days (16 h light) for one week.

Seedlings without fully green, expanded cotyledons were counted as antibiotic/herbicide sensitive. Only seedlings with T-DNA insertions can grow, and theoretically, only one insertion should be segregating in progeny of each hemizygous line where 50 % of the population should be hemizygous and 25 % homozygous (for the insertion) according to Mendel's laws. The next step was to sow resistant seedlings in soil for screening by PCR to identify the homozygous mutant individuals (T3 generation). Seeds from the resultant homozygous transgenic lines were collected and used for all further experiments.

2.2.4. Abiotic stress treatments

2.2.4.1. Freezing tolerance test

Wild-type and mutant plants were grown in soil in a growth chamber at 22°C, under a short-day photoperiod (8 h light/16 h dark) for 4 weeks and then incubated at 4°C for 3 days for cold acclimation under a short-day photoperiod. The plants were then placed in darkness in a cooled incubator (Gallenkamp model) with the following freezing temperature regimen: from 4°C to -3°C±1°C then hold at -3°C for 3 h; then in a similar timing sequence until -6°C±1°C was reached and held there for 3 h. After removal from the freezing chamber,

seedlings were incubated at 4°C overnight in darkness then returned to the original growth conditions. Plant damage was assessed 2 weeks later.

2.2.4.2. Hormone treatment

ABA (+-Abcisic acid, 99%, Sigma) was dissolved in 100% ethanol. Sterilised seeds were plated on basal medium without sucrose and solidified with 0.8% agar supplemented with 0, 0.5 or 1 μM of ABA.

PAC (paclobutrazol), a gibberellin biosynthesis inhibitor, was dissolved in acetone and was added to MS medium (1% sucrose) to different final concentrations (0.4 and 4 μM).

After stratification for 2-3 d at 4ºC, plates were transferred to a growth room (23ºC±1ºC under long days, 16hr light/8 hr dark) for 1-2 weeks.

2.2.4.3. Salt tolerance

For salt tolerance assays, mutant and wild-type sterilised seeds were planted vertically on germination medium. After 4-6 days growth, seedlings were transferred carefully to new MS plates supplemented with 50 mM or 125 mM NaCl and placed with roots pointing downward (with plates placed vertically). Pictures were taken 2 weeks later.

2.2.4.4. Dehydration stress

For water loss measurements, 6 week-old plants (at the rosette stage) which had been grown on soil in a growth cabinet under short day photoperiod were used. Rosettes were detached from their roots at the soil surface and weighed immediately in a plastic weighing boat. The same procedure was used for detached leaves. The boat with the tissue then was placed on a laboratory bench (Mengiste *et al*., 2003; Verslues *et al*., 2006) and weighed at designated time intervals. There were four replicates for each line. The percentage loss of fresh weight was calculated based on the initial weight of the leaves/rosette.

2.3. Results

2.3.1. Identification and main features of AtMYB6 and AtMYB8

Phylogenetic analysis of the R2R3-MYB proteins of Arabidopsis, places MYB6 and MYB8 in subgroup 4 under an established classification scheme for the R2R3 MYB transcription factor family (Kranz *et al*., 1998). Subgroup 4 has four other members, MYB3, MYB4, MYB7 and MYB32, all of them already characterized functionally except MYB7 (Fig. 2.1A). This family is characterized by having motifs representing a putative activation domain LlsrGIDPx$^{T}/_{S}$ HRx$^{I}/_{L}$ at the end of the R3 repeat of the DNA binding domain and a putative negative regulatory domain pdLHL$^{D}/_{E}$Lxi $^{G}/_{S}$ in their C-termini (Fig. 2.2) (Jin *et al*., 2000). Deletion and mutational analysis of AtMYB4 showed that the repression motif forms part of the region involved in the repression of transcription. It is also interesting to note that this motif has similarities to a characterized repressor motif called the ERF-associated amphiphilic repression motif (EAR motif), present in class II AP2/ERF transcriptional repressors and also in zinc finger proteins which function as repressors (Ohta *et al*., 2001). The fact that these repressors have this EAR-like domain indicates that they function as active repressors, which is interesting because few transcription factors that act as repressors have been identified in plants.

AtMYB6 and AtMYB8 share both putative regulatory domains but not the putative C-terminal zinc-finger domain ($CX_{1-2}CX_{7-12}$ CX_2C) present in other members of the subfamily. The functional roles of this group are diverse; UV-protection, pollen development or abiotic stress signalling amongst others (Jin *et al*., 2000; Preston *et al*., 2004; Bang *et al*., 2008).

The *AtMYB6* gene encodes an R2R3 MYB-like protein comprising 237 amino acids with an estimated molecular weight of 26.3 kDa. *AtMYB6* contains two exons (316 and 667 bp) and a single intron (308 bp) which interrupts the sequence encoding the R3 repeat. *AtMYB6* is present as a single copy gene on chromosome IV (Fig. 2.1B).

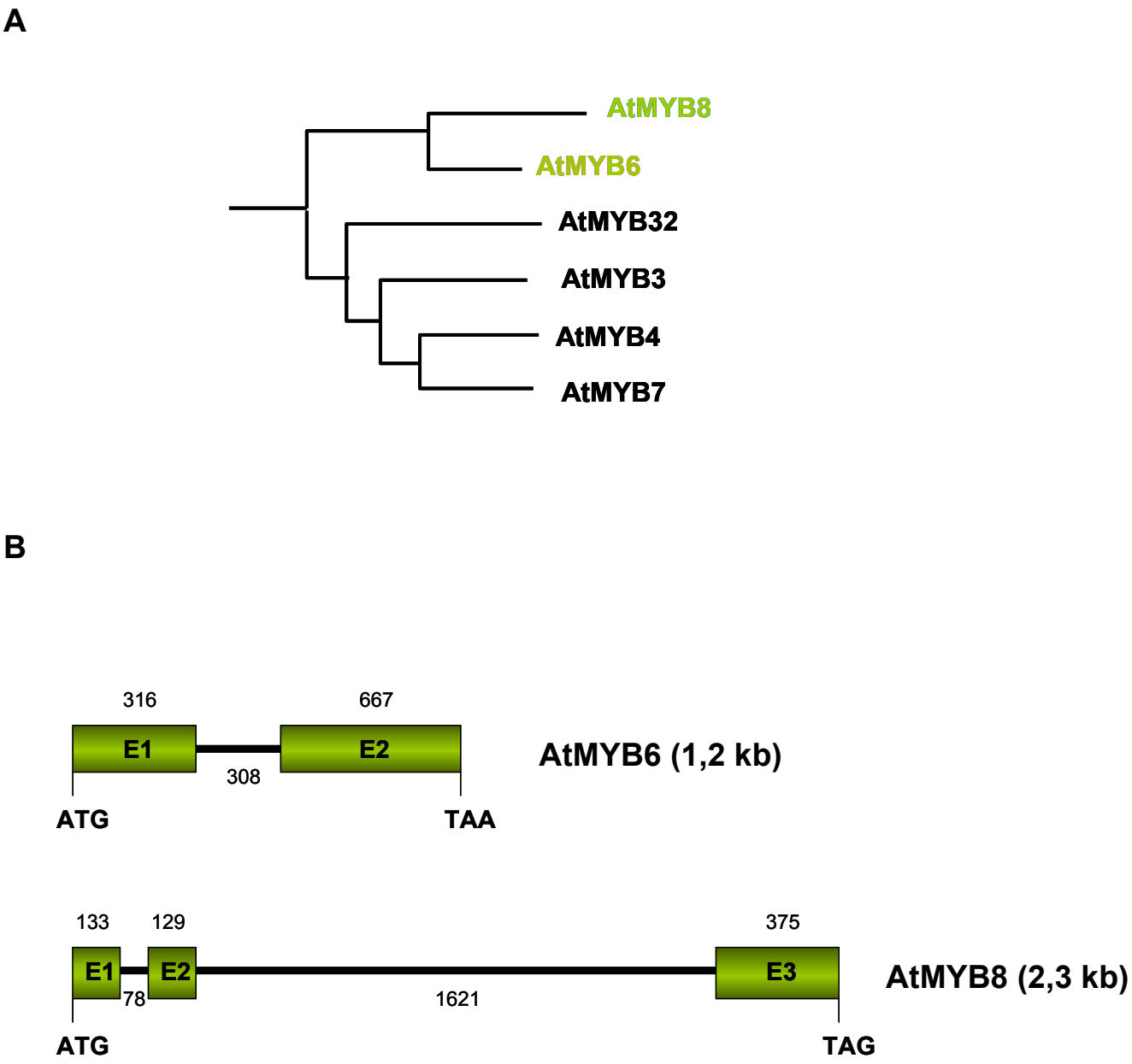


Figure 2.1. Phylogenetic relationship and gene structure of *AtMYB6* and *AtMYB8*

(A) Subgroup IV of the *Arabidopsis thaliana* MYB transcription factors family showing the phylogenetic proximity of *AtMYB6* and *AtMYB8* genes (Courtesy of Paul Bailey). Detailed bootstrap values can be seen in Fig. 5.6 (Chapter 5). (B) Main features of *AtMYB6* and *AtMYB8* genes indicating exon (E, boxed in green) and intron positions (black lines) and their sizes in base pairs (bp).

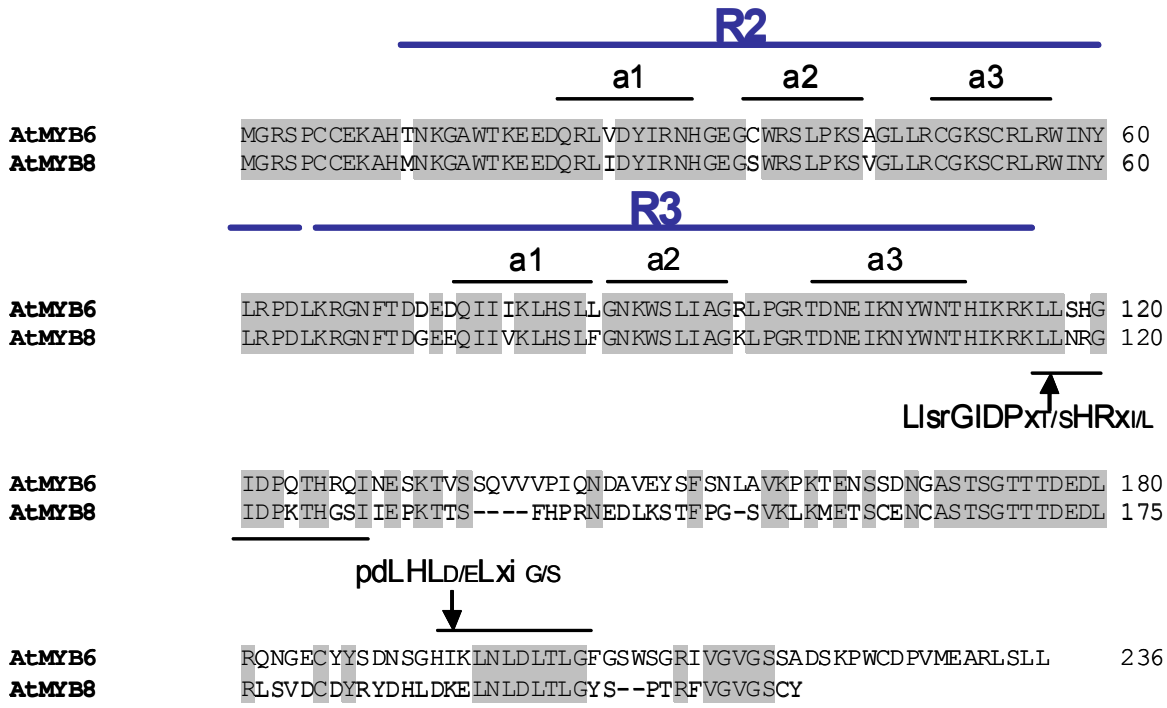


Figure 2.2. Amino acid alignment of AtMYB6 and AtMYB8

The alignment was carried out using the ClustalW program. The lines above indicate the R2 and R3 MYB repeats with their three putative α-helices. Grey background indicates 100% identity. Motifs representing the conserved putative activation (L1srGIDPxT/SHRxI/L) and the EAR repression (pdLNLD/ElxiG/S) domains are indicated. It can be seen the high conservation of the MYB domains. In contrast, their C-termini are highly divergent outside of the conserved motifs.

The *AtMYB8* gene encodes an R2R3 MYB-like protein comprising 212 amino acids with an estimated molecular weight of 23.6 kDa. *AtMYB8* contains three exons (133, 129 and 375 bp) and two introns (78 and 1621bp) which interrupt the sequences encoding the R2 and R3 repeat respectively and *AtMYB8* is present as a single copy gene on chromosome I (Fig. 2.1B). The proteins encoded by *AtMYB6* and *AtMYB8* are closely related with a 71% amino acid identity over the whole protein sequences.

2.3.2. Isolation of the T-DNA insertion mutants

To examine the possible role of *AtMYB6*, we used a reverse genetic approach to identify homozygous T-DNA insertion mutants. Several insertion lines were listed for *AtMYB6* and *AtMYB8* from The Arabidopsis Information Resource database, TAIR (www.arabidopsis.org) (Fig. 2.3A). I chose mainly those insertion lines with T-DNA insertions in exons because I wanted to identify knockout mutants at this stage. The presence of a T-DNA insert in different alleles was confirmed by PCR using gene-specific primers and a T-DNA LB border primer. The presence or absence of bands was used to distinguish between hemizygous, homozygous and wild type seedlings. Some mutant lines (shaded grey, section 2.2.2, table1), were discarded due to lack of a Mendelian ratio of segregation or due to the impossibility of detecting the T-DNA insertion. All lines were in the Columbia (Col-0) background, whereas the *hos10-1* allele was derived by T-DNA insertion in the C24 ecotype (Ishitani *et al*., 1997; Zhu *et al*., 2005). Mutant lines were grown side-by-side with wild-type controls and no phenotype was seen during vegetative growth of any of the homozygous insertion lines for *AtMYB6* or *AtMYB8* in the Col-0 background. However, the *hos10* mutant was clearly distinct from the C24 progenitor, exhibiting a reduced size phenotype (Fig.2.4A). To ensure the lack of expression of these genes in further experiments, seeds were collected from one homozygous plant of each line.

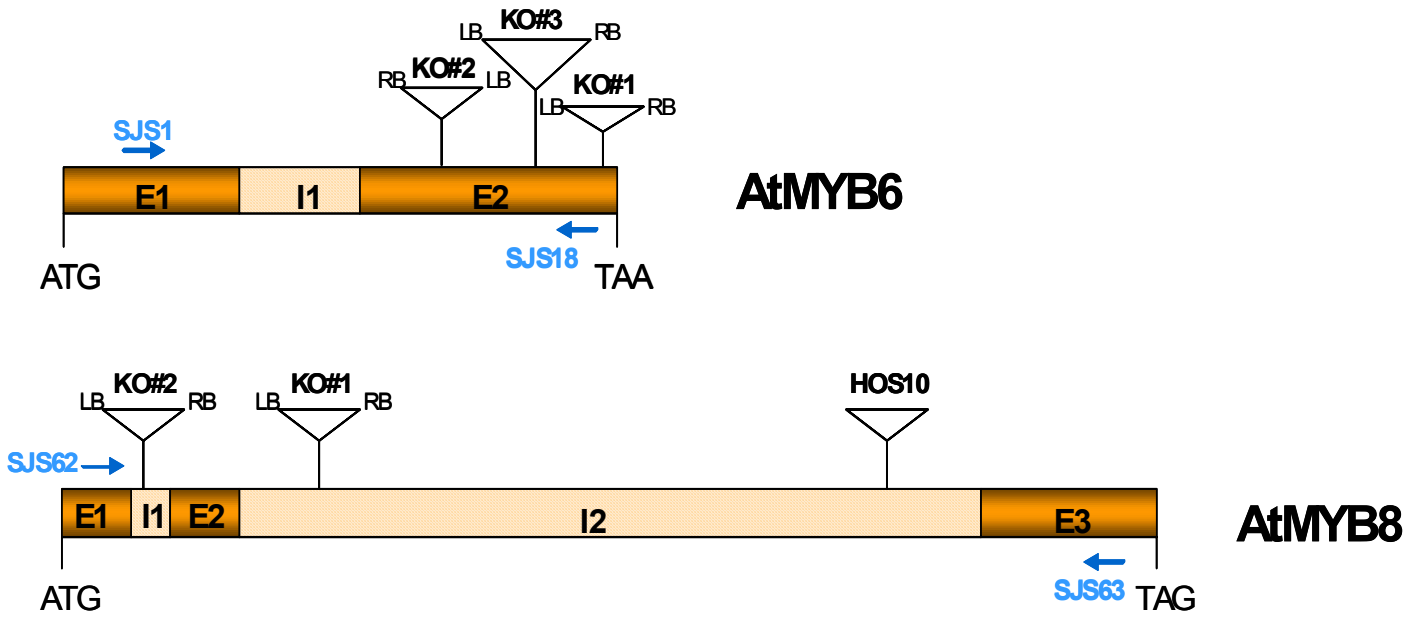


Figure 2.3. *AtMYB6* and *AtMYB8* mutants lines used in this project

Structure of the *AtMYB6* and *AtMYB8* genes and positions of the T-DNA inserts for each gene. LB and RB, represent the T-DNA left border and right border respectively. Arrows indicate primers used for the RT-PCR (see Fig. 2.4). E, exon ; I, intron.

Figure 2.4. Phenotypes and expression analysis of the mutant lines

(A) 3-4 weeks old Arabidopsis plants grown under short days photoperiod. (B) RT-PCR expression analysis of *MYB6* and *MYB8* alleles. Sizes for wild-type *AtMYB6*, *AtMYB8* and *cyclophilin* are approximately 600, 340 and 200 base pairs (see primers used in Fig.2.3).

RNA prepared from developing leaves of mutant plants was examined by reverse transcription RT-PCR using the SJS1-SJS18 oligo pair for *MYB6* and SJS62-SJS63 for *MYB8* together with primers to the cyclophilin gene as a control (Fig. 2.4B).

The three T-DNA insertions selected for the *MYB6* gene were all in the second exon. In this case, the sequence encoding the MYB domain remains intact and only the sequence encoding the C-terminus of the protein, where the regulation domain is located, would be truncated. This could potentially result in a functional protein with an altered regulatory activity. In contrast, the selected *MYB8* insertions were in introns. T-DNA inserts in intronic sequences are sometimes spliced out of the pre-mRNA with the intron, thus allowing for some gene expression.

For the *MYB6* gene, no transcript was detected in the *MYB6-KO#1* but *MYB6-KO#2* and *MYB6-KO#3* were transcribed as truncated mRNAs where the MYB domain was partially disrupted (Fig. 2.4B). With the *MYB8* insertions, there was expression of the gene in the line ARA173 (consequently, this line was discarded) but fortunately, the mutation in *MYB8-KO#1* was null as evidenced by my inability to detect *MYB8* transcript in this line. In all the samples, cyclophilin transcript was detected.

2.3.3. Abiotic stress responses in *AtMYB-KO* mutants compared to *hos10*

2.3.3.1. Freezing tolerance test

Visual phenotypes of *AtMYB6-KO* and *AtMYB8-KO* alleles showed no similarity to the dwarf phenotype of *hos10* under standard conditions. Testing the sensitivity to low temperatures, which was the other main phenotype of the *hos10* mutant, was the next step to confirm any other similarities between the different mutant lines.

This experiment was performed manually (without a programed timer inside the cold cabin) since no low temperature growth chamber was available in the Institute (see Materials and Methods in this Chapter). It was difficult to

find a manual temperature regime to set up differences in cold sensitivity between mutants and wild-type so several cold tests were carried out both in plants grown on plates and in soil as described previously (Zhu *et al*., 2004; Zhu *et al*., 2005; Paul *et al*., 2006).

Unlike the published results for the *hos10* mutant, the *AtMYB8* mutant and *AtMYB6* mutants were not hypersensitive to temperatures below -2ºC in seedlings grown in soil although they showed slight sensitivity at -6°C especially the *MYB6-KO#3* (Fig. 2.5). However, it also proved impossible to reproduce the cold sensitivity of the *hos10* mutant. Even under normal conditions, *hos10* did not grow properly and some seedlings died, making it almost impossible to assess the effect of low temperature stress on its growth. Col-0 showed similar damage at -6°C (non-acclimated) as *MYB6-KO#2* and *MYB8-KO#1* whereas *MYB6-KO#1* and *MYB6-KO#3* displayed a slight hypersensitivity at -6°C as mentioned previously above. However, the significant observation was that *hos10* showed no sign of hypersensitivity to cold compared to C24 at -6°C.

Although the reproducibility of cold treatments was not good because of the imprecision in the time taken to reach lower temperatures, I was sure that *hos10* (and the rest of the *AtMYB8* and *AtMYB6* insertion lines) were exposed to very low temperatures by checking the temperatures reached. The expected result for *hos10* hypersensitivity to cold (as suggested by the published work) was not observed. My inability to demonstrate the “hypersensitivity” to freezing temperatures of the mutant *hos10* meant that these tests, involving cold stress, were not taken further.

2.3.3.2. Abcisic acid treatment

Abscisic acid (ABA) is a plant growth regulator that participates in a wide spectrum of molecular processes.

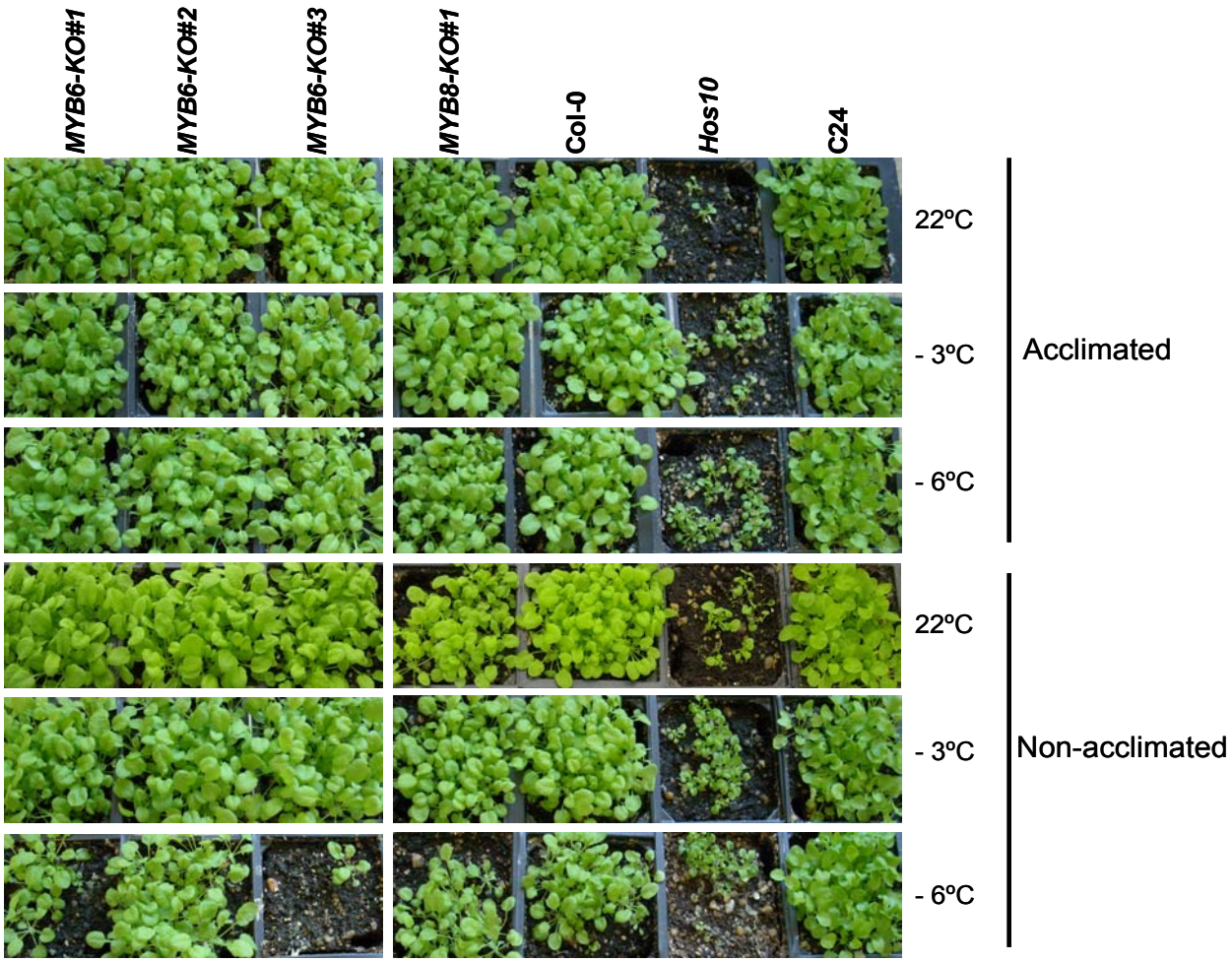


Figure 2.5. Freezing tolerance test

Wild-types and mutants treated at different temperatures. Damage assessed 2 weeks later after cold treatment.

Cold stress increases ABA levels transiently. These transient increases in ABA have been shown to mediate the transcriptional induction of some cold-responsive genes (Yamaguchi-Shinozaki and Shinozaki, 2006). The *hos10* mutant, described as highly sensitive to freezing temperatures, is impaired in abscisic acid (ABA) production in response to dehydration stress due to its inability to fully induce NCED3, the rate-limiting enzyme in the ABA biosynthesis (Zhu *et al*., 2005).

In my efforts to find a connection between *hos10* and the other alleles of *AtMYB6* and *AtMYB8,* thinking that the *hos10* phenotype could be related to an alteration in ABA biosynthesis, I did a quick test to investigate the response of the new *AtMYB6* and *AtMYB8* alleles to ABA. It is known that the addition of ABA to mature seeds inhibits their germination; ABA maintains seed dormancy (Michael *et al*., 2008). I tested if *AtMYB6* and *AtMYB8* alleles responded similarly to the application of exogenous ABA as *hos10.* By examining the inhibition of germination, I could get an idea about how ABA affects these mutants compared to *hos10* which, hypothetically, should have reduced levels of endogenous ABA, and therefore, should show less inhibition of germination by ABA than the wild-type.

Arabidopsis seeds, including wild-type ecotypes (Wt) and mutants were plated directly on solid MS medium containing 1% sucrose, supplemented with two different concentrations of ABA (Sigma) in a long day regime. Pictures were taken approximately 1-2 weeks later depending on when differences between wt and mutants were most obvious. In the control plate without the addition of hormone, seedlings grew normally (Fig. 2.6) but on ABA-supplemented medium, visible changes could be observed especially in *hos10*. With 0.5 μM ABA, most of *hos10* seedlings showed fully green expanded cotyledons whereas C24 was still germinating. Mutants in the Col-0 background, *MYB6-KO#1* and *MYB8-KO#1* looked similar to wild-type (Col-0) at that ABA concentration. When the ABA concentration was increased to 1 μM, C24 germination was almost completely inhibited but *hos10* was able to germinate and develop pale green cotyledons.

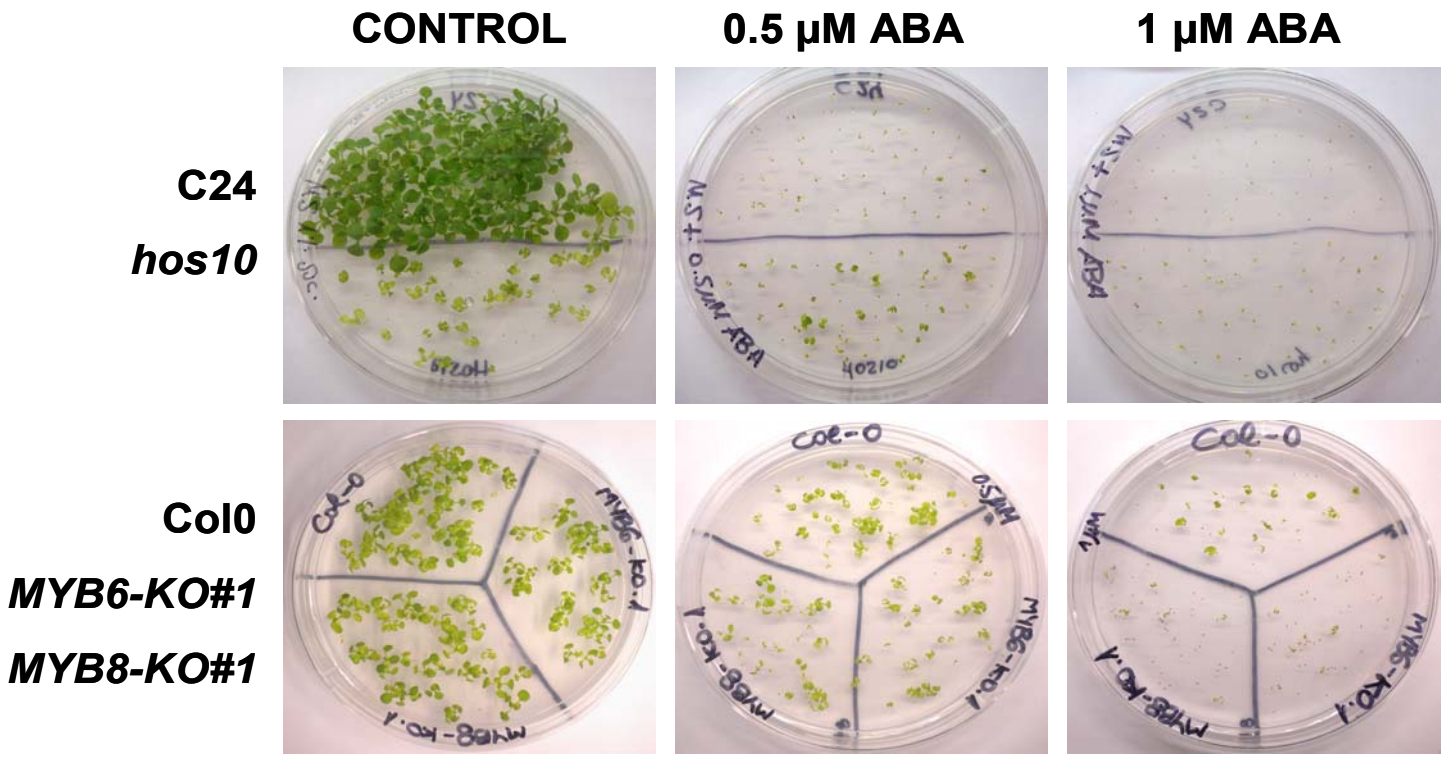


Figure 2.6. Photographs of wt, AtMYB6 and AtMYB8 knock-out seedlings grown on MS medium or supplemented with different concentrations of ABA.

The panel above shows *hos10* compared with its wild-type background C24. Pictures beneath, the mutant lines of *AtMYB6* and *AtMYB8* compared to Col-0 treated with different levels of ABA.

By contrast, for the remaining mutants of *AtMYB6* and *AtMYB8*, the differences in germination and growth were not as striking, but ABA 1 μM caused a slightly greater reduction in their growth rate compared to Col-0. These tests supported the idea that the *hos10* mutant has decreased ABA levels because it was significantly less sensitive to inhibition of germination and growth by exogenous ABA. By contrast, the mutants of *AtMYB6* and *AtMYB8* were slightly more sensitive to ABA than the Col-0 control.

Clear differences in germination were observed in *hos10* when ABA was present in the medium but this was not seen with the *AtMYB6* and *AtMYB8* alleles. I undertook another assay for the effect of ABA on seed germination. Gibberellins promote seed germination and the breaking of the dormancy maintained by ABA (Debeaujon and Koornneef, 2000). Mutants with enhanced gibberellin (GA) signaling or decreased ABA content are more resistant to inhibition of germination by paclobutrazol (PAC), a GA biosynthesis inhibitor.

AtMYB6 and *AtMYB8* alleles were grown on MS medium containing 1% sucrose and different concentrations of PAC (0.4 μM and 4 μM). Germination rate was scored day by day but no significant differences were found at either 0.4 μM or at 4 μM between Col0 and the *AtMYB6* and *AtMYB8* mutant alleles; germination was inhibited in all the seeds (photographs not shown).

2.3.3.3. Salt tolerance

Salinity is another abiotic stress to which plants respond physiologically. Because salinity stress involves lower availability of water, it is related to abiotic stresses such as drought and low temperature (Yamaguchi-Shinozaki and Shinozaki, 2006). The *hos10* mutant was claimed to be hypersensitive to NaCl, impaired in the ability to adjust to osmotic stress and to produce shorter roots than the wt when grown vertically in germination medium.

To determine whether *MYB6* and *MYB8* knockout lines are salt sensitive during early seedling development, seeds from these homozygous lines were germinated on Murashige and Skoog media. The knock-out lines were grown vertically in germination medium for 4-6 days and then were carefully

transferred to new MS plates supplemented with various NaCl concentrations with roots pointing downward. After ~15 d of cultivation on MS medium with 50 mM NaCl, wild-type (wt) seedlings and the mutant plants continued their growth normally (Fig. 2.7). The wt plants and mutants were stressed and showed anthocyanin pigmentation on MS medium with 125 mM NaCl, but at first sight only *hos10* roots were shorter than those of the C24 wt and no visual differences were seen for the remaining lines compared to Col-0.

2.3.3.4. Water loss meaurements

ABA signalling, by regulating stomatal aperture, plays a crucial role in reducing water loss under water shortage/deficit.

The *hos10* mutant plants are unable to synthesize enough ABA and lose water more easily compared to wild-type. Thus, short-term water-loss assays were performed, by evaluating the decline in fresh weight of detached leaves or detached rosettes of the wild-type and the new *AtMYB6* and *AtMYB8* mutants (Verslues *et al*., 2006). The percentage loss of fresh weight was calculated based on the initial weight of the rosette leaves.

The loss-of-function *AtMYB6* mutants #1, #2, #3 as well as *AtMYB8*#1, did not exhibit significant differences in the transpiration rates of detached leaves/rosettes compared to wild-type (Col-0) (Fig. 2.8). The different *AtMYB6* mutants responded somewhat differently between themselves within both experiments. The *AtMYB8* knockout allele showed a very similar transpiration rate to wild-type. Importantly, *AtMYB8*#1 did not show the great loss of fresh weight that was reported in the paper describing *hos10* (Zhu *et al*., 2005).

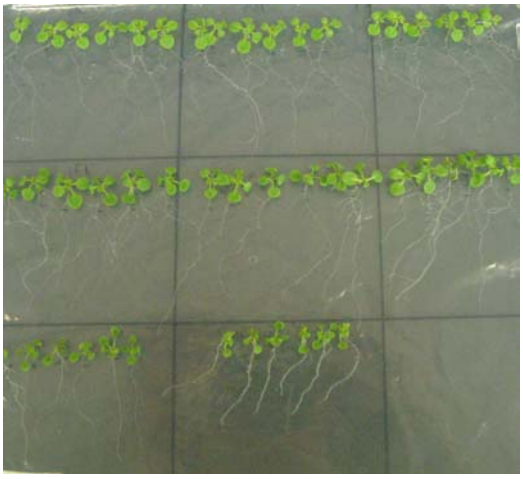
.

2.3.4. *Hos10* is not an *AtMYB8* allele

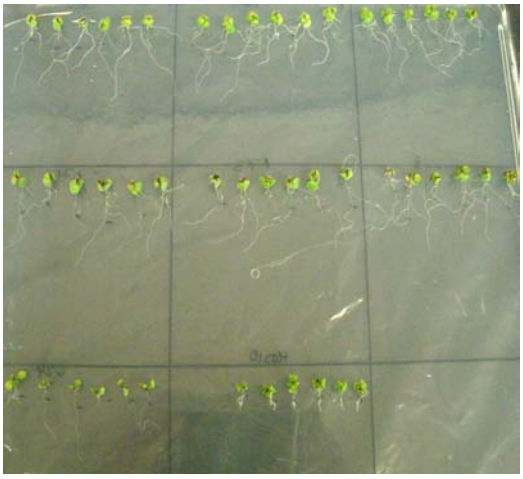
Despite all my efforts to try to find a phenotype that linked the new knockout alleles of *AtMYB8* and *AtMYB6* to the first allele of AtMYB8 described, *hos10,* using a range of growth conditions under which *hos10* had been characterised (Zhu *et al*., 2005), I found no similarities between the phenotypes of the new knockout alleles of *AtMYB6* and *AtMYB8* and *hos10*.

Col-0	*MYB6-KO#1*	*MYB6-KO#2*
MYB6-KO#3	*MYB8-KO#2*	*MYB8-KO#1*
C24	*HOS10*	

Control



NaCl 50 mM



NaCl 125 mM

Figure 2.7. Salt treatment.

Photograph of wild-type and *AtMYB6* and *AtMYB8* knockout lines after ~15 d of cultivation on MS medium + 0 mM NaCl, MS+ 50 mM NaCl and MS + 125 mM NaCl medium. The square table above shows the position on the plate of individuals of each line.

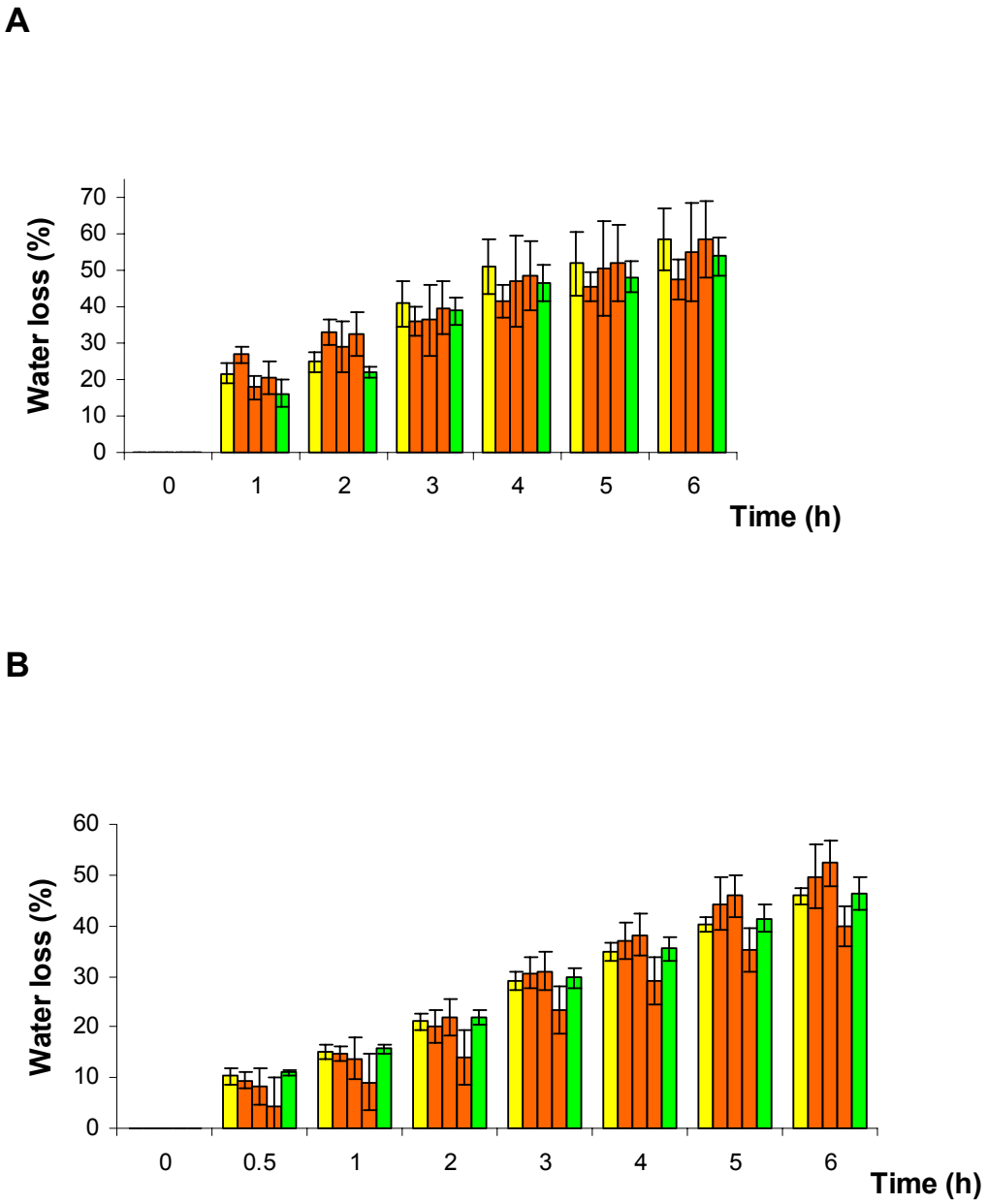


Figure 8. Short term water loss avoidance.

To monitor the water loss in the wt and mutants lines, detached leaves (graph A) or detached rosettes (graph B) were used. Graphs show decreasing in fresh weight over time after detachment. Yellow bars indicate Col-0, orange *AtMYB6* alleles #1, #2, #3 respectively and green, *AtMYB8#1* allele.

What is more significant, even the new allele, *AtMYB8*#1, did not show the striking phenotypes of *hos10*. The genetic background of the *AtMYB6* and *AtMYB8* alleles was Col-0 whereas that of *hos10* was C24. However, it seemed unlikely that differences in genetic background were the cause of a complete absence of phenotype in the *AtMYB8*#1 allele particularly because the Zhu et. al (2005) manuscript reported that other alleles of *hos10* in the Col-0 background had a "similar phenotype" in terms of sensitivity of *hos10* to abiotic stress.

For *AtMYB6,* functional redundancy with *AtMYB8* likely could provide an explanation for the lack of a *hos10*-like phenotype associated with the knockout mutations although this explanation then makes it difficult to explain the *hos10* mutant phenotype if it is due to *AtMYB8* knockout.

As a consequence of my inability to repeat the findings demonstrating the abiotic stress sensitivity of *hos10*, and my inability to confirm the clear visible phenotype of *hos10* through independent knockout alleles of *AtMYB8*, the "authenticity" of *hos10* being a knockout of *AtMYB8* was questioned and therefore, checked. PCR using primers (SJS8-SJS9) that bind either side of the hypothetical T-DNA insertion was performed (Fig. 2.9A, B). The PCR on *hos10* genomic DNA, showed a fragment of the same size as that the wild-type indicating the absence of a T-DNA insertion in the *AtMYB8* gene. To further confirm this result, RT-PCR on mutant and wild-type cDNA was also carried out, using the oligo pair SJS62-SJS63. In both cases, normal transcripts of the gene were detected demonstrating that *AtMYB8* was not interrupted in the *hos10* mutant (Fig. 2.9C). The phenotype of *hos10* must then be due to another mutation or a T-DNA insertion elsewhere in the genome. The T-DNA used in generating the *hos10* mutant line was reported to carry the gene for resistance to Basta herbicide (Glufosinate-ammonium, Sigma). Basta resistance of this line would indicate the presence of a T-DNA insertion (Fig. 2.9D). As shown in Fig. 2.9D, on MS medium with a Basta concentration of 20µg/ml, *hos10* was as sensitive as the wild-type line C24 to Basta. Plants did not develop green expanded cotyledons or a proper root system.

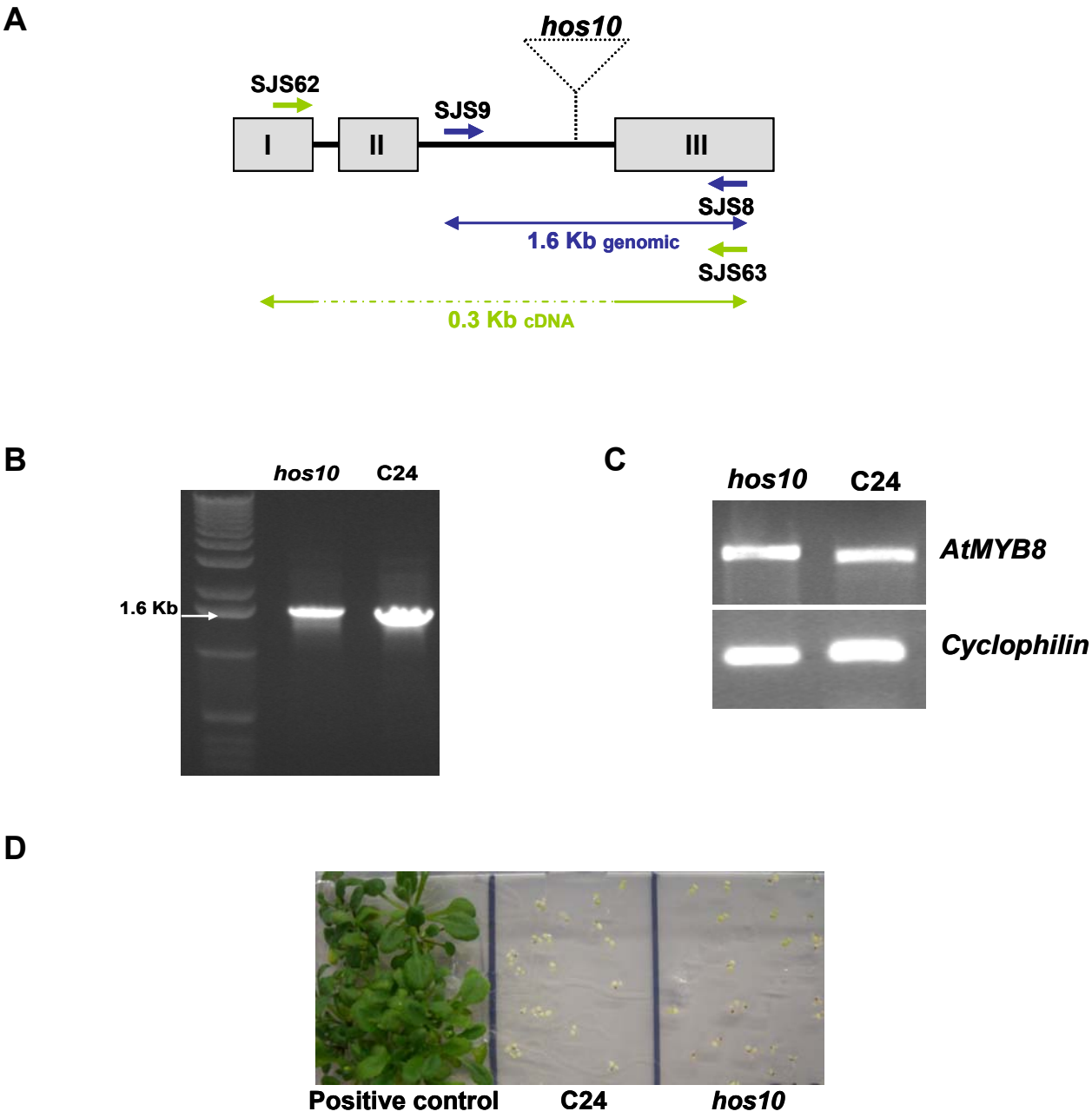


Figure 2.9. Testing whether *hos10* carries a T-DNA insertion in the *AtMYB8* gene

(A) Diagram of the *AtMYB8* gene structure and the primer positions used for the PCR and RT-PCR of the *AtMYB8* gene in the *hos10* line. The 1.6 kb size does not include the *hos10* insertion. (B) Genomic PCR of *hos10* and C24. (C) RT-PCR analysis of *AtMYB8* expression in the *Hos10* line. (D) Photograph showing the sensitivity of the different plant lines growing in MS media supplemented with glufosinate-ammonium.

However, the positive control, a transgenic line containing a T-DNA insertion with the Basta resistance gene, grew perfectly.

2.4. Discussion

Abiotic stresses such as drought, extremes of temperature, or high salinity are the most common abiotic stresses that plants encounter. From an agricultural point of view, such stresses are among the most significant factors responsible for substantial and unpredictable losses in crop production. Success in breeding for better adapted varieties to abiotic stresses depend on the use of modern molecular biology tools for elucidating the control mechanisms of abiotic stress tolerance. Hence, genetic engineering for developing stress tolerant plants, based on the introgression of genes that are known to be involved in stress response and putative tolerance, might prove to be a fast track towards improving crop varieties.

Functional analyses of plant R2R3-type MYB proteins indicate that they regulate numerous processes including responses to environmental stress (Kranz *et al*., 1998; Saibo *et al*., 2009). Curiously, just one MYB transcription factor from Arabidopsis, MYB15, has been reported to be involved in cold response, activating downstream genes that confer freezing tolerance to plants (Agarwal *et al*., 2006). Cold stress significantly constrains agricultural productivity. Studies on acquired freezing tolerance in Arabidopsis have contributed substantially towards the understanding of cold acclimation mechanisms. The *hos10* mutant was characterized as defective in the activity of an R2R3-type MYB transcription factor (*AtMYB8*), defective in the cold acclimation response (Zhu *et al*., 2005) and therefore an important component of cold signalling in plants.

This chapter describes the experimental approaches taken to elucidate a putative functional redundancy between the genes *AtMYB6* and *AtMYB8* (reportedly *hos10*) under different abiotic stress conditions.

MYB6 and MYB8 have been described as Arabidopsis transcription factors from the R2R3 MYB family (Li and Parish, 1995; Zhu *et al*., 2005). MYB

factors are a conserved and diversified family involved in different plant processes related to metabolism, development or stress (Jin and Martin, 1999). Phylogenetic studies place both transcription factors in the Arabidopsis R2R3 MYB subfamily 4 together with MYB3, MYB4, MYB7 and MYB32 (Kranz *et al*., 1998). In the subfamily which includes *AtMYB6* and *AtMYB8*, the remaining three characterized MYB members appear to have different physiological functions in Arabidopsis and only AtMYB3 has been suggested to be involved in abiotic stress responses, in addition to *hos10* (Bang *et al*., 2008).

Comparisons were made between the homozygous knockout lines of *AtMYB6*, *AtMYB8* and *hos10*. Although being in different genetic backgrounds (*hos10* is in the C24 ecotype), none of the tested lines under standard conditions showed similar phenotypes to *hos10*. Individual homozygous lines of *AtMYB6* and *AtMYB8* with single insertions were evaluated for phenotypic abnormalities. Specific tissues like roots, cotyledons, leaves and flowers were examined, as well as the overall size and form of the plants, and mutants were morphologically indistinguishable from wild-type plants. Stress treatments including salt, freezing, dehydration or ABA were examined for a further phenotypic analysis. Not even under stress conditions were visible phenotypic differences found between the *AtMYB6* and *AtMYB8* mutant lines and wild-type except for the *hos10* mutant. This kind of experiment also led me to find some discrepancies with the results reported in the *hos10* paper (Zhu *et al*., 2005). The freezing tolerance test did not reveal a striking hypersensitivity to cold in any of the lines tested: not even in *hos10* for which Zhu *et al*., 2005, showed almost complete plant death at – 2ºC. However, under salt stress, *hos10* formed shorter roots as consequence of its hypersensitivity to NaCl, a phenomenon that was reported by Zhu *et al*., (2005), but which was not observed in the *AtMYB6* and *AtMYB8* knockout mutant lines. The ABA hormonal treatment that I applied to the mutant lines indicated that *hos10* was probably impaired in ABA biosynthesis as claimed by Zhu *et al*., (2005). Germination of *hos10* on medium with ABA was not totally inhibited as happened with the wild-type ecotypes and the *AtMYB6* and *AtMYB8* mutants.

This result suggested that the *AtMYB6* and *AtMYB8* mutant lines contained normal ABA levels. Therefore, when these mutant lines were exposed to dehydration stress, they did not display significant differences compared to the controls. This supports the idea that *AtMYB6* and *AtMYB8* mutant alleles are not altered in the ABA pathway and mutants make enough ABA for their stomata to respond to water deficit.

The dwarf phenotype of *hos10* is striking. If AtMYB8 function was essential for plant development, its malfunction should affect the plant phenotype independently of its genetic background. Moreover, none of the five *AtMYB8* mutant lines genotyped showed the characteristic phenotype present in *hos10.* Although Zhu *et al*., (2005) described isolating additional alleles of AtMYB8 including insertion mutants in the Col-0 background from the SALK collection, these were not described as having dwarf phenotypes and no pictures of them were shown. The *hos10* paper indicates only that *hos10* alleles were also defective in cold acclimation and had early flowering. I saw no differences in flowering time between the *AtMYB8-KO#1* line and Col-0. It seems very unlikely that differences in genetic background could be the cause of a complete absence of phenotype in the other *AtMYB8* alleles.

These observations led me to check for the presence of a T-DNA insertion in the *hos10* mutant. PCR and RT-PCR analyses revealed that *hos10* does not have any insertion in the *AtMYB8* gene such as that described in the PNAS paper (Zhu *et al*., 2005). Surprisingly, the paper reports confirmation that loss of *AtMYB8* gene function causes the *hos10* mutant phenotype. For instance, it reported a positive complementation test with the *AtMYB8* cDNA under the control of the CaMV 35S promoter. In addition, a picture of an RT-PCR analysis of *AtMYB8* expression in *hos10* was shown to demonstrate the absence of transcript in the mutant. I have demonstrated that based on the material provided to me, that this statement is false after several tests indicating no interruption of the gene and the RT-PCR analysis showing an AtMYB8 transcript in the *hos10 mutant*. Also, I could not detect an early flowering phenotype in *hos10* as shown in the paper.

As a consequence of these discrepancies, I felt that the results reported in the *hos10* paper were not reliable at all. My concerns were communicated to Dr. Ray A. Bressan and subsequently apologies were received after they confirmed that in fact, *hos10* was not an *AtMYB8* allele. A retraction of the paper (Zhu *et al*., 2005) has been issued (Zhu et al.). To decipher what causes the *hos10* mutation, mapping of the mutant genome should be carried out since it seems that the phenotype is most likely due to a spontaneous mutation rather than to a T-DNA insertion (since *hos10* seedlings were sensitive to Basta).

This conclusion left us without a solid hypothesis about what the biological roles of AtMYB6 and AtMYB8 could be since mutants were morphologically indistinguishable from wild-type plants under standard and stress conditions The functional redundancy and multifunctionality of MYB transcription factors (Jin and Martin, 1999) complicated the establishment of a new hypothesis. Even within R2R3 MYB subgroup 4, several different functional roles of members have been characterized; mutation of AtMYB32 leads to aberrant pollen (Preston *et al*., 2004), AtMYB3, is involved in abiotic stress (Bang *et al*., 2008) whereas AtMYB4, controls production of UV light sunscreens (Jin *et al*., 2000). Nonetheless, AtMYB32 and AtMYB4 share certain functional similarities. Changes in the expression levels of AtMYB32 and AtMYB4 lead to a similar pollen phenotype, and both affect the expression of different genes in the phenylpropanoid pathway. This suggests that the phylogenetic proximity of AtMYB6 and AtMYB8 to subgroup 4 means that they may have similar or redundant functions to other members, although they could also serve specialized roles in plants.

CHAPTER 3

Chapter 3: Phenotypic analysis to identify biological functions of the *AtMYB6* and *AtMYB8* genes

3.1. Introduction

Chapter 2 describes several approaches to test the hypothesis of a possible redundancy between *AtMYB6* and *AtMYB8* genes. The *hos10* mutant (Zhu *et al*., 2005) was the reference mutant used for all these experiments although, when *hos10* was checked, I showed that it lacked any T-DNA insertions and was not (as previously described (Zhu *et al*., 2005)) an *AtMYB8* mutant. Clear, visible phenotypes were not seen in the loss-of-function alleles of *AtMYB6* and *AtMYB8* when compared to the wild-type under the different abiotic stress conditions tested.

The discovery that the *hos10* mutant was not affected in AtMYB8 activity meant that our working hypothesis that AtMYB6 and AtMYB8 regulate ABA biosynthesis and responses to abiotic stress was invalidated. The absence of visible phenotypes in the knock-out mutant plants, made a functional genomics approach to find out the possible biological functions of both MYB transcription factors essential (Jin and Martin, 1999; Stracke *et al*., 2001) .

In this chapter, I present data suggesting that *AtMYB6* and *AtMYB8* may have partially redundant functions due to their expression in overlapping patterns through the whole plant. To have an idea about what these functions could be, a bioinformatic tool, Genevestigator, an expression database system (Zimmermann *et al*., 2004), was used to retrieve information regarding the expression *AtMYB6* and *AtMYB8* under different conditions. These data were used to set up new experiments to shed light on the function of both MYB genes. In addition, results from another project carried out by Dr. Benedicte Lebouteiller, revealed a specific altered metabolic profile in plants over-expressing *AtMYB6* and *AtMYB8* that suggested a connection between these genes and phenylpropanoid metabolism.

3.2. Materials and methods

3.2.1. Cloning of *AtMYB6* and *AtMYB8* cDNAs

The coding sequence (CDS) of *AtMYB6* and *AtMYB8* genes (genomic loci At4g09460 and At1g35515, respectively) were cloned in the Gateway entry vector pDONR207 (MYB6) or pDONR201 (MYB8) and then subcloned into the pJAM1502 destination vector between the double CaMV35S enhancer and the CaMV terminator (for detailed methods see Gateway cloning in General Materials and Methods).

Constructs and Arabidopsis lines overexpressing *AtMYB6* and *AtMYB8* were provided by Dr. Benedicte Lebouteiller.

3.2.2. Cellular localization of *AtMYB6* and *AtMYB8*

3.2.1.1. Plasmid constructs

Cellular localization of both proteins was examined by transient and stable expression of the AtMYB6-GFP and AtMYB8-GFP fusion proteins. Entry clones containing the CDS of both genes were in frame with the GFP gene in the N-terminal eGFP destination vector pB7WGF2,0 (Karimi *et al*., 2002) under the control of the CaMV35S promoter. The constructs for fusion protein production were verified by sequencing using the following primers: SJS1 and SJS8 for the AtMYB6 fusion and SJS7 and SJS8 for the AtMYB8 fusion. Plasmids were introduced into *A. tumefaciens* strain GV3101 by direct transformation.

3.2.1.2. Transient expression in *Nicotiana benthamiana*

Recombinant *A. tumefaciens* strains were grown overnight at 28ºC in 10 ml of LB medium supplemented with 100 µg/ml spectomycin and 25 µg/ml rifampicin. Cells were harvested by centrifugation and resuspended to a final concentration corresponding to an optical density (OD) of 0.5 at 600 nm in a solution containing 10 mM $MgCl_2$ and 150 µM acetosyringone. Cultures were

incubated at room temperature at least 1 h before infiltration. For the agroinfiltration procedure, two zones on the opposite half-leaves of *Nicotiana benthamiana* plants were infiltrated. The upper leaves of plants were infiltrated with a 1-ml syringe without a needle. Leaves were superficially wounded with a needle to improve infiltration. Two plants were agroinfiltrated per construct. The infiltrated leaves were photographed at 2 days after agroinfiltration using a confocal scanning laser microscope Leica SP2.

3.2.1.3. Stable transformation of Arabidopsis

Stable transformation was performed by floral dipping (Clough and Bent, 1998) as indicated in "General Materials and Methods".

F1 seeds were sown in trays and sprayed with Basta. Seeds from Basta resistant plants were used for subsequent analysis. Pictures of seedlings expressing GFP were taken with a Leica SP2 confocal scanning laser microscope.

3.2.3. Tissue localization of *AtMYB6* and *AtMYB8* expression

3.2.3.1. Plasmid constructs

Promoter sequences of MYB6 and MYB8 were obtained from the Arabidopsis Information Resource, TAIR (www.arabidopsis.org/). These were obtained by PCR amplification of the *A. thaliana* wild-type genomic DNA. PCR products were recombined into pDONR201 or pDONR207 to produce an entry clone that was then used to recombine the gene promoter fragment into pBGWFS7 vector which contains the green fluorescent protein (GFP) and ß-glucuronidase (GUS) reporter genes (Karimi *et al*., 2002)

For construction of AtMYB6-GFP::GUS, a 1199 bp promoter fragment (primers SJS64 and SJS65 including an attB1 and attB2 site, respectively) of *AtMYB6* was cloned. The sequence was confirmed with the primers Att-Fw, Att-Rv and SJS71. In the same way, a 1280 bp *AtMYB8* promoter was cloned (primers SJS75II and SJS76II). The clone was verified by sequencing using the

primers Att-Fw, Att-Rv, SJS79, SJS80 and SJS81. Plasmids were introduced into *A. tumefaciens* strain GV3101 by direct transformation and Arabidopsis agroinfiltration as indicated in 3.2.1.3.

3.2.3.2. GUS histochemical assay

Promoters of the genes studied were fused to the reporter gene *uidA* (GUS, encoding the β-glucuronidase). Histochemical GUS staining of transgenic Arabidopsis plants was performed as described by (McCabe *et al*., 1988). Tissues from plants grown in Petri dishes, sterile disposable plastic boxes or soil were harvested and immersed in GUS staining solution containing 61 mM Na_2HPO_4, 3.9 mM $NaH_2PO_4.H_2O$, pH 7.0, 0.1 % Triton X-100 and 10 mM EDTA and 0.25 mg/ml X-glu (X-glu = 5-bromo-4-chloro-3-indolyl-β-D-glucuronide Na salt; Melford) and 0.3% H_2O_2 added just before using. Tissues were vacuum infiltrated for 10 minutes and incubated overnight at 37°C. Following incubation, samples were destained several times with 70% [v/v] ethanol until the blue GUS stain was clearly visible and chlorophyll completely removed. The tissue was observed under a Leica MZ16 dissecting microscope coupled to a Leica DFC420 camera.

3.2.4. Production of a double mutant by crossing *AtMYB6-KO#1* with *AtMYB8-KO#1*

Crosses were performed by standard methods using *AtMYB6* and *AtMYB8 KO* homozygous lines. Immature anthers were removed with forceps, and the pistils were fertilized two days later. Several crosses were done: *MYB6-KO#1* x *MYB8-KO#1*, *MYB6-KO#2* x *MYB8-KO#1*, *MYB8-KO#1* x *MYB6-KO#3*. The subsequent generations were analyzed by PCR. DNA was extracted from young leaves initially by a rapid method for genotyping (see General Materials and Methods) although after an initial screening, DNA was subsequently extracted with a DNeasy plant kit (Qiagen). Primer pairs for detection of insertions in *AtMYB6* locus were SJS2-SJS87, SJS18-LBGabi or SJS1-LBSalk and for *AtMYB8* were SJS78-SJS73 and SJS89-LBSalk.

3.2.5. Disease assays

3.2.5.1. Bacterial infection

Plants aged 5–6 weeks (grown in environmentally-controlled chambers at 23ºC±1ºC under short days) were used for the infection experiments. *Pseudomonas syringiae* pv. *Pto* DC3000 (or *Pto* DC3000Δ*avrPtoB*) strain was grown at 28°C on King's B medium (40 g /l proteose, 20 g /l glycerol) containing the appropriate antibiotics for selection. When the OD_{600} was between 0.6-1.0, the culture was collected and resuspended in 250 ml 10mM $MgCl_2$. Plants were sprayed with a bacterial suspension containing 1×10^8 cfu/ml bacteria with 0.04% Silwet L-77.

Leaf discs (4 from different plants of each line) were harvested 2 days after inoculation and ground in 10 mM $MgCl_2$ in a microfuge tube with a pestle. After grinding the tissue, samples were vortex-mixed and diluted 1:10 serially. Samples were finally plated on NYGA solid medium (5 g/l bactopeptone, 3 g/l yeast extract, 20 ml/l glycerol, 15 g/l agar) supplemented with rifampicin 100 µg/ml. Plates were placed at 28°C for 2 days, after which the colonies were counted.

3.2.5.2. Measurement of reactive oxygen species

Leaf discs (0.38 cm^2) were floated on water overnight and reactive oxygen species released by the leaf tissue were measured using a chemiluminescent assay. The water was replaced with 500 µl of a solution containing 20 µM luminol (Sigma) and 1 µg of horseradish peroxidase (Fluka,). In all experiments, ROS (reactive oxygen species) were elicited with 100 nM flg22 or 100 nM elf18, two peptides from bacterial proteins that act as potent elicitors of defence responses. Luminescence was measured using a Photek camera system (East Sussex, UK) over a time-series, and data were recorded at the maximum intensity of the response.

3.2.5.3. Assessment of the effects of oxidative stress following chemical treatment

Rose bengal (RB; 4,5,6,7-tetrachloro-2',4',5',7'-tetraiodofluorescein) (Sigma) or paraquat (methyl viologen, MV) were dissolved in water and filtered before using. Seedlings grown on MS medium for 5 days were transferred to new MS medium supplemented with different concentrations of RB (6 μM) or MV (3 μM).

3.2.6. Germination assay

Surface-sterilized seeds of the same age were sown on MS 0.8% agar medium. Plates were placed in darkness at 4°C for 48 h to promote and synchronize germination and then transferred to long day conditions. Radicle emergence was used as the criterion for seed germination. Germination was scored at the indicated time points (day 3 and day 5) using a dissecting microscope. Germination tests were repeated in at least triplicate experiments using > 80 seeds for each individual assay. The average germination percentage ± SE of triplicate experiments was calculated.

3.2.7. Root length measurements

Seeds were grown on MS-agar supplemented with 1% sucrose, under a photoperiod of 16 h of light, 8 h of darkness, and temperature of 23°C for 5 days. Plates were placed vertically to allow root growth along the surface of the agar and to allow the unimpeded growth of the hypocotyl into the air. Primary root length was measured from the root tip until the hypocotyl border using a plastic rule. 16 seedling roots were measured per line.

3.2.8. Microscopy

Roots were cleared in Hoyer's solution (7.5 g of gum arabic, 100 g of chloral hydrate, 5 ml of glycerol in 60 ml of water) to allow high-resolution imaging and pictures were taken using a Leica DFC360 FX light microscope (Nomarski).

3.2.9. Brassinosteroid treatment

Seeds of *AtMYB6* and *AtMYB8 KO* mutant alleles were germinated on 0.8% MS medium plus 1% sucrose in plastic disposable boxes under a long day photoperiod. For the treatment, 100nM brassinolide (Epibrassinolide, Duchefa), the most active BR, was added to the basal medium. Photographs were taken 3 and 6 weeks later.

For the etiolation test, plants were grown vertically on MS plates for 5 days in darkness.

3.2.10. Testing stomatal aperture

3.2.10.1. Stomatal conductance

These experiments were done in collaboration with Dr. Gregorio Galvez (Essex University). From five plants, three fully expanded leaves from each one (6 week-old plants grown under short days) were selected for the measurement of stomatal conductance, using a portable photosynthesis system (LI-6400).

3.2.10.2. Detached rosette experiment

(section 2.2.4.4 in Chapter 2)

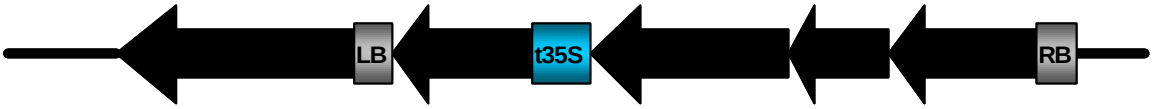
3.3. Results

3.3.1. Nuclear localization of AtMYB6 and AtMYB8

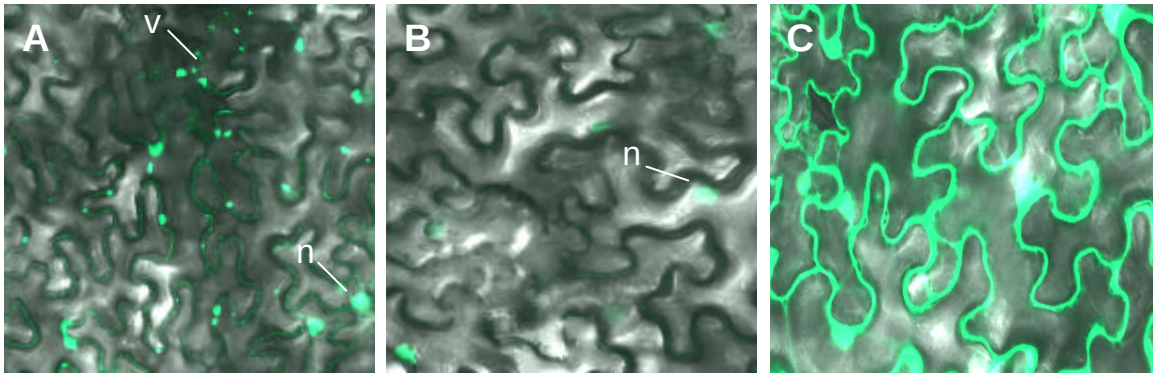
The cellular localization of both MYB proteins was investigated using N-terminal green fluorescent protein (GFP) fusions to full-length AtMYB6 and AtMYB8 proteins. Translational fusions of AtMYB6 and AtMYB8 to eGFP, driven by the 35S promoter, were constructed using the destination vector pB7WGF 2 (Karimi *et al*., 2002) (Fig. 3.1.1).

Transient expression of AtMYB6::GFP and AtMYB8::GFP fusion proteins into *N. benthamiana* resulted in nuclear localization of the fluorescence signal

1.

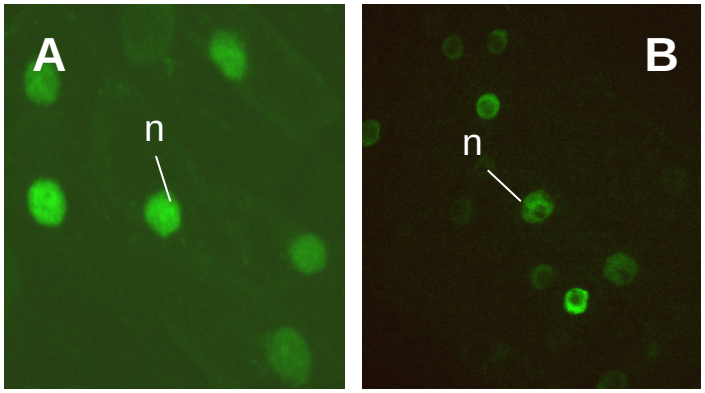


2.



AtMYB6::GFP AtMYB8::GFP 35S::GFP

3.



AtMYB6::GFP **AtMYB8::GFP**

Figure 3.1. Subcellular localization of AtMYB6 and AtMYB8

1. Diagram of pB7WGF2 Gateway plant transformation constructs corresponding to C-terminal fusions of GFP to full-length AtMYB6 and AtMYB8.

2. Subcellular localization of AtMYB6 (A) and AtMYB8 (B) in transiently transformed *N. benthamiana* leaves. (C) Distribution of the GFP protein alone as control. Pictures taken using a Leica confocal microscope (n, nuclei; v, vesicles).

3. Subcellular localization of AtMYB6 (A) and AtMYB8 (B) in stably transformed *A. thaliana*. Confocal pictures showing nuclei in ~5 day-old roots (n, nuclei).

except the control (with the 35S-GFP vector alone) where GFP was found throughout the cell, a finding that is consistent with the predicted regulatory roles for the transcription factors encoded by both genes (Fig. 3.1.2).

Although fluorescence of AtMYB8::GFP fusion was clearly localised to the nucleus, fluorescence of the AtMYB6::GFP fusion was also detected in cytoplasm and in a punctate pattern associated with the cell wall (Fig. 3.1.2A). This signal could be an artefact resulting from overexpression of the protein although the recent discovery of inter- and/or intracellular movements of transcription factors (Jorgensen, 2000; Wu *et al*., 2003; Kurata *et al*., 2005) could imply that AtMYB6 translocates from the nucleus to the cytoplasm or vice versa.

To rule out a possible aggregation induced by the GFP in the fusion protein, stable transformants in Arabidopsis with the same constructs used for the transient expression were carried out. Several seedlings for each construct grown in MS media for approximately 5 days were analyzed using confocal microscopy. AtMYB6::GFP and AtMYB8::GFP were observed only in the nuclei of from root cells (no signal was found in leaves)(Fig. 3.1.3) suggesting that the cytoplasmic and cell wall fluorescence observed in the transient assays of AtMYB6 were artifactual, resulting from the high levels of protein in the transient assays.

3.3.2. Expression patterns of *AtMYB6* and *AtMYB8*

For *AtMYB6*, a promoter-GUS::eGFP transgene construct was made by cloning a 1.2-kb PCR product comprising the genomic DNA sequence directly upstream of the predicted ATG translational start codon of GFP. In the same way, a 1.2-kb promoter fragment from *AtMYB8* was also cloned to make the reporter construct. The resulting plasmids, containing the *AtMYB6prom-GUS::eGFP* and *AtMYB8prom-GUS::eGFP* expression cassettes were used for *Agrobacterium tumefaciens*-mediated plant transformation.

Two independent *AtMYB6prom-GUS::eGFP* transgenic lines of Arabidopis with a 3:1 segregation of the basta resistance marker in the T2

generation, indicating the insertion of the T-DNA at a single locus, were used for further analyses. For *AtMYB8prom-GUS::eGFP*, it was not possible to find a homozygous line so, hemizygous plants were used for the analyses. Selected lines were analysed by light microscopy for histochemical GUS detection.

The *AtMYB6* promoter was active in almost the whole plant particularly in the vasculature (*i.e.*, roots, leaves, stems, stamen, stigma and base of the siliques). Emerging first true leaves of young *AtMYB6prom-GUS::eGFP* seedlings showed uniform expression of the reporter gene and expression was also detected in the shoot apical meristem (Fig. 3.2 Ba). As the leaves expanded, expression was restricted to trichomes, vascular system and leaf margins, more precisely in the hydathodes (Fig. 3.2 Be-Bf) although that pattern was not reproducible in each leaf type or during different developmental stages. During early flower development, GUS staining was found in floral primordia and later, only in the stigmas, stamens and abscission zone (Fig. 3.2 Bc-Bd). After anthesis, GUS staining was stronger in the abscission zone and in the stigma (Fig. 3.2 Bd). In roots, localization of GUS expression was observed in the whole system except in the root meristem (Fig. 3.2 Bh). Detailed examination of transgenic plants harbouring a translational fusion between *AtMYB6* and the reporter gene encoding the green fluorescent protein (GFP) revealed that this gene is also expressed in stomata (Fig. 3.2 D). Expression driven by the *AtMYB8* promoter was somewhat more restricted than for the *AtMYB6* promoter.

Different stages and different organs were stained and, in common with *AtMYB6, AtMYB8* was expressed in roots, floral primordia and stigmas (Fig. 3.2Ca-Cc, Ce, Cf). In contrast to *AtMYB6*, *AtMYB8* was expressed in anthers (Fig. 2Cf) and stronger GUS staining was detected in the main leaf vein and in developing petals of floral primordia (Fig. 3.2Cd-Ce). GFP expression driven by the *AtMYB8* promoter was not detected in stomata as occurs with *AtMYB6* (not shown).

These data together with the *AtMYB6* and *AtMYB8* expression profile reported by the Genevestigator expression database

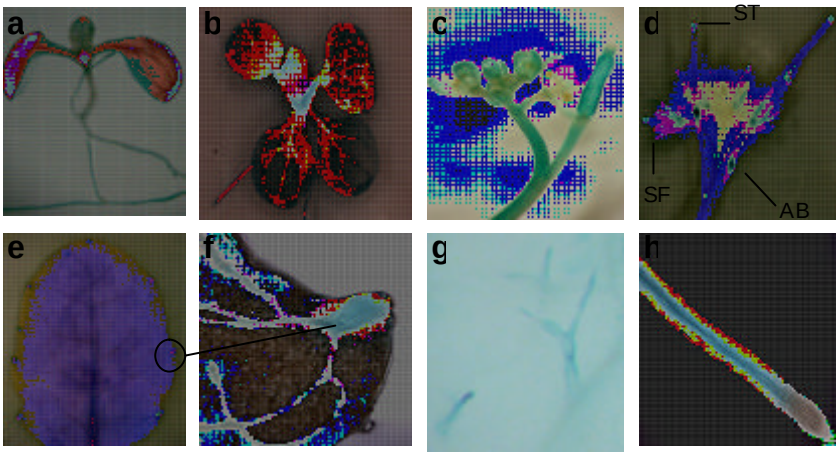
Figure 3.2. GUS and GFP expression of *AtMYB6* and *AtMYB8* promoters

Expression patterns in Arabidopsis Col-0 obtained with the *AtMYB6* and *AtMYB8* promoter: GUS fusion constructs. Images examined using a Leica MZ16 dissecting microscope. (A) Diagram of pB7WGF7 construct containg *AtMYB6* or *AtMYB8* promoter . (B) GUS expression patterns driven by the *AtMYB6* promoter at different developmental stages: a-b, young seedlings; c-d, flowers and siliques (ST, stigma; SF, stamen and filaments; AB, abscission zone); e-f, hydathodes in leaves; g, leave trichomes and h, root. (C) GUS expression patterns driven by the *AtMYB8* promoter at several stages: a-b, young seedlings roots; c and e, floral primordial; d, leaf and f, flower. (D) Confocal analysis of GFP expression in stomata in *A. thaliana* leaves. Picture on the left is the control showing autofluorescense of the stomatal aperture and picture on the right is MYB6::GFP expression in guard cells.

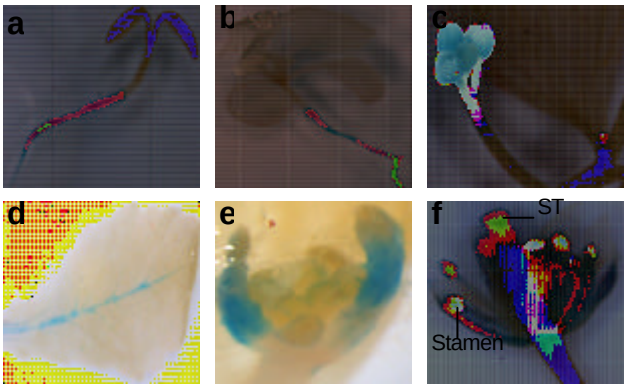
A


LB
t35S
RB

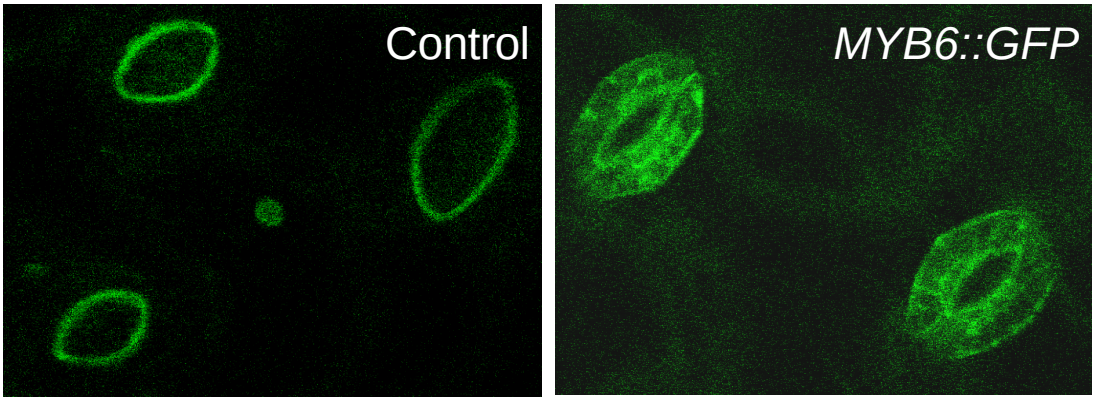
B


a
b
c
d
ST
SF
AB
e
f
g
h

C


a
b
c
d
e
f
ST
Stamen

D


Control
MYB6::GFP

(https://www.genevestigator.com/gv/index.jsp) established a basis to design new experiments that served as new strategies for the functional characterization of these two MYB transcription factors.

Digital histograms of *AtMYB6* and *AtMYB8* genes showing their expression under different conditions did not show many stimuli that provoked a significant activation or repression in both genes (Fig. 3.3). Some treatments such as cold, drought, salt, ABA or PAC, seem to alter the expression levels of these genes but as shown in the previous chapter, they did not have strong effects on the visual phenotype of *AtMYB6* and *AtMYB8* mutants. Other treatments with elicitors, bacteria or other hormones were also shown in Genevestigator to have an effect in the expression of these genes and therefore, were used for further analysis to find phenotypic differences in the lines with altered activities of *MYB6* and *MYB8*.

3.3.3. *AtMYB6* and *AtMYB8* and plant innate immunity

From the analysis of the *AtMYB6* and *AtMYB8* expression profiles reported by the Genevestigator expression database and as a part of the characterization of these genes of unknown function, the response of the *AtMYB6* and *AtMYB8 KO* mutants to biotic challenge was evaluated.

It has been reported that some MYB transcription factors are involved in responses to pathogen attack (Cheong *et al*., 2002; Vailleau *et al*., 2002; Mengiste *et al*., 2003).

An experiment was designed to check if *AtMYB6-KO* mutants were affected by biotic stress, in this case by *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto DC3000*), a virulent strain able to infect Arabidopsis and to cause disease symptoms (Hirano and Upper, 2000).

Plants at the young rosette stage were sprayed with a bacterial suspension and disease lesions were scored every 2 days. Pictures were taken one week after the inoculation (Fig. 3.4A). There were no detectable differences in the damage of the *AtMYB6* knock-out alleles compared with wild-type plants. Growth of the *Pto DC3000* pathogen on the same plants, was also

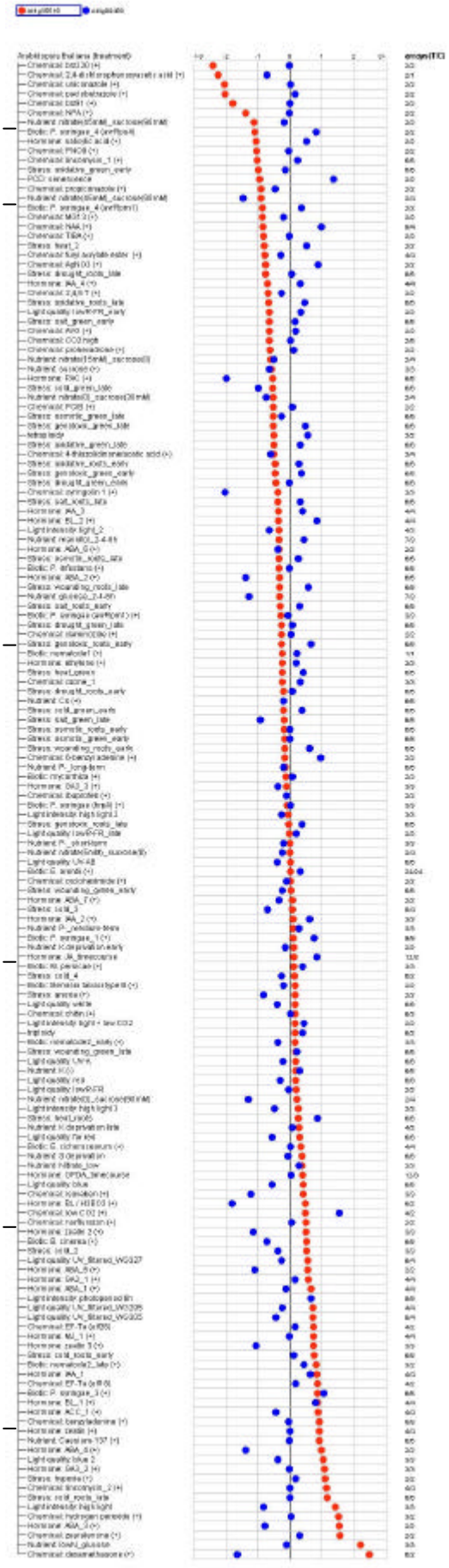


Figure 3.3. Genevestigator stress-specific expression profile

Diagram shows the stimulus tool of the Genevestigator which provides a summary of gene expression responses to external stimuli or perturbations. In general, the responses are compilations of several repetitions of the same treatment within a single experiment. If the same stimulus is tested in several experiments, this stimulus will appear multiple times in the list of stimuli. The gene expression responses are calculated as ratios between treatment and control samples. The resulting values thus reflect up- or down-regulation of genes.

In list view, stimuli are displayed as a linear list and ordered by the expression ratio of the gene marked with a blue rectangle in this case, AtMYB8 (red dots). AtMYB6 expression is showed with blue dots. Black arrows indicate those treatments with *Pseudomonas syringae* and brasinosteroid. The chip type is an ATH1 (22k full genome Arabidopsis Affymetrix GeneChip).

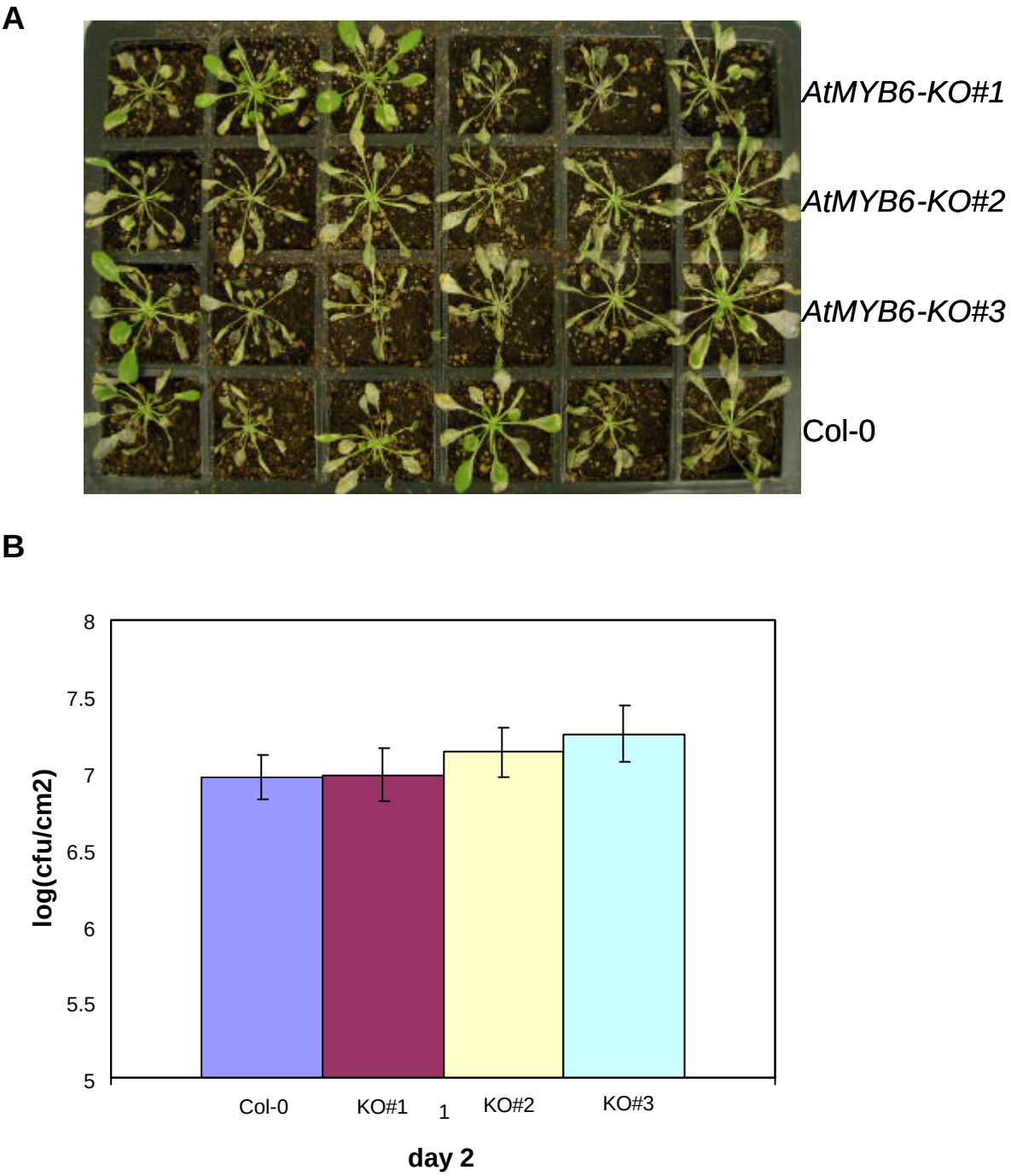


Figure 3.4. *Pseudomonas syringae pv. tomato DC3000* infection

(A) Disease symptoms on Col-0 and *AtMYB6* mutant (KO) lines after spray inoculation with *Pto DC3000* bacteria. The picture was taken 1 week after the infection. (B) Growth of *Pto DC3000* on wild-type and *AtMYB6* mutant plants 2 days after spray treatment. Error bars indicate SEM.

assayed by counting the colony-forming units (cfu) in 2 day-old infected leaves and no significant differences in pathogen numbers between wild-type and *AtMYB6* mutants were observed (Fig. 3.4B).

Another way to assess sensitivity of plants to necrotrophic pathogen attack is with the application of ROS producing chemicals to the medium because these pathogens promote host cell death through the generation of these chemical species (Lamb and Dixon, 1997). Rose bengal and paraquat are two different reagents that generate ROS. The first one generates singlet oxygen species when exposed to light and the second one inhibits electron transport in the reduction of NADP to NADPH during photosynthesis so generating ROS from the chloroplast.

AtMYB6 and *AtMYB8* mutant alleles together with the wild-type were sown in MS media supplemented with rose bengal (6 μM) or paraquat (3 μM). After 2 weeks growing vertically on plates, all the seedlings were affected to a similar degree, they were chlorotic and damage was uniformly distributed between wild-type and knock-out mutants (pictures not shown).

When *AtMYB6* and *AtMYB8* overexpression lines became available, new pathogen assays were carried out. Plants were infected with *Pto* DC3000 and also with another less virulent *Pseudomonas* strain, *Pto* DC3000Δ*avrPtoB*. An Arabidopsis mutant Col-0 *fls2* was used as a positive control. This mutant can not perceive flagellin, an elicitor of the plant defence response, and therefore, is more susceptible to bacterial infection (Gómez-Gómez and Boller, 2000). A few days after the inoculation (with *Pto* DC3000 or *Pto* DC3000Δ*avrPtoB*), some chlorotic leaves were observed in plants in the two trays to which the different strains were applied (Fig. 3.5A-B). Larger disease lesions were visible in plants treated with *Pto* DC3000 due to its higher pathogenity but, in both cases, there were no differences when overexpression lines were compared to the wild-type. Only the positive control, *Col-0 fls2,* showed signs of hypersensitivity when infected with either bacterial strain.

A

Col-0 *MYB6-KO#2* *MYB8-KO#1* *MYB6-OX* *MYB8-OX* Col-0 *fls2*

B

Col-0 *MYB6-KO#2* *MYB8-KO#1* *MYB6-OX* *MYB8-OX* Col-0 *fls2*

Figure 3.5. Disease symptoms on *AtMYB6* and *AtMYB8* mutants lines after spray inoculation

Pictures showing symptoms on knock-out (KO) mutants and overexpression (OX) lines compared to Col-0 and Col-0 fls2, one week after spraying. (A) Treatment with *Pto* DC3000?*avrPtoB.* (B) Inoculation with *Pto* DC3000 (MYB6-KO#1 was previously assayed, Fig. 3.4).

Elicitors are molecules that activate defence responses, including the generation of reactive oxygen species. Measuring ROS can give a quantitative indication of plant damage. Control and mutants plants were treated with flg22 and elf18, two bacterial peptides that elicit the generation of ROS.

Measurements of ROS in Arabidopsis leaf tissue were assayed by measuring the H_2O_2-dependent luminescence of luminol (Lourdes *et al*., 1999). Luminescence was measured with a Photek camera system (East Sussex, UK) over time (Fig. 3.6). Luminescence indicating damage in plants inoculated with elf18 was slightly less than in the plants treated with flg22 but comparisons between wild-type and mutants did not show obvious differences in the levels of ROS induced in response to either elicitor.

3.3.4. Brassinolide application in *AtMYB6* and *AtMYB8 KO* mutants

Brassinosteroids (BRs) are unique steroidal phytohormones that exhibit a broad spectrum of physiological functions of plant growth, differentiation, and development. Brassinolide (BL) is the most active BR. Application of brassinolide to the medium promotes hypocotyl elongation in light-grown Arabidopsis plants (Clouse and Sasse, 1998; Asami *et al*., 2005). Although the mutants used in this study did not show the classic characteristics of the BR phenotype in the light (dwarfism, dark green, rounded leaves or delayed development), bioinformatic data from Genevestigator suggested changes in expression of *AtMYB6* and *AtMYB8* in response to BL hormone treatment.

Arabidopsis seeds were treated with 100 nM of BL to assess the effects of this naturally active BR on early growth of seedlings. Exogenous BL in the culture medium clearly repressed the elongation of the inflorescences in *MYB6-KO#1*. BL promoted slightly stem elongation in Wt (Col-0) and also in the *MYB6-OX* line (Fig. 3.7A,B). Pictures taken few weeks later, almost at the senescent stage, showed relatively longer inflorescences in *MYB6-OX* and Col-0 compared to *MYB6* knock-out in plants which had been treated with the hormone (Fig. 7B). In untreated plants, the inflorescences of *MYB6-KO#1* were longest (longer than those of Wt) and the inflorescence stems of *MYB6-OX* were visually shorter than Wt.

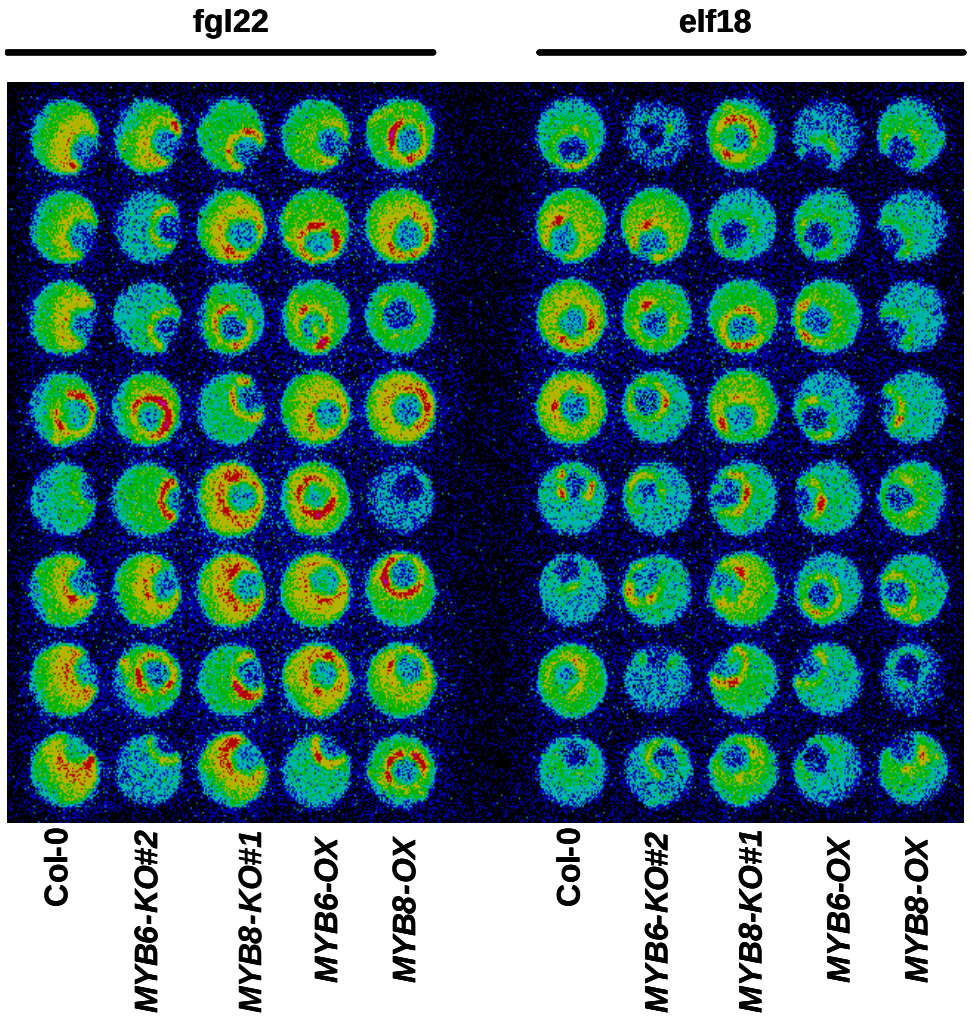


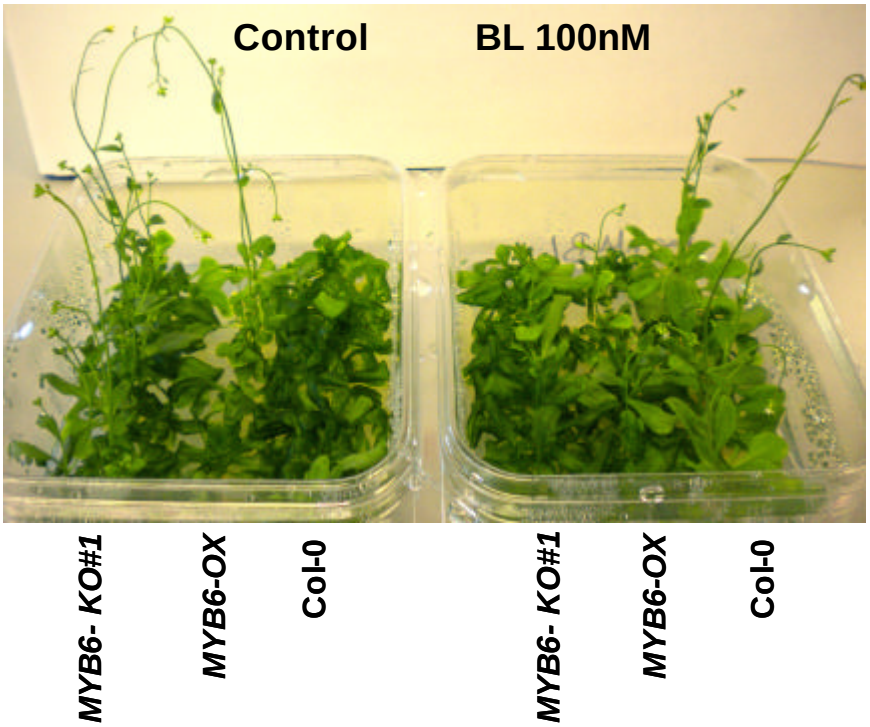
Figure 3.6. Oxidative burst in leaf tissues of *AtMYB6* and *AtMYB8* mutants

Luminescence of *A. thaliana* leaf slices from knock-outs (KO) and overexpression (OX) lines in a solution with luminol and peroxidase after treatment with 100 nM flg22 or 100 nM elf18. There were eight biological replicates per line for each treatment. Red colour indicates highest production of ROS and therefore, more damage to the plant. Quantitative results wer not made since no striking differences were visually observed.

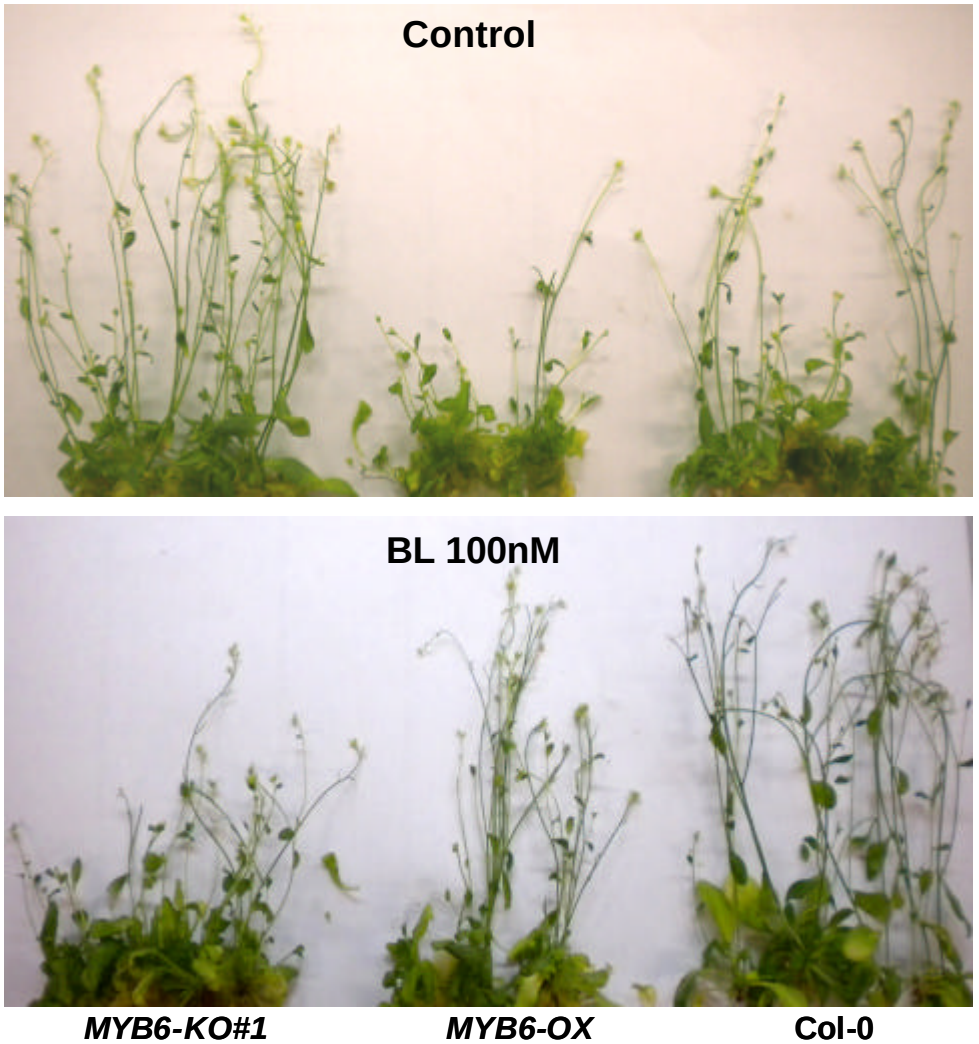
Figure 3.7. Treatment with the brassinosteroid hormone, brassinolide

(A) ~ 3 week old seedlings of *AtMYB6* knock-out and overexpression lines without hormone treatment or 100 nM brassinolide. (B) The same plants as in (A) but ~ 6 weeks old and removed from the media to facilitate photography. (C) ~ 3 week old seedlings of *AtMYB8* knock-out and overexpressed lines without hormone treatment or 100 nM brassinolide. (D) the same plants as in (C) but ~ 6 weeks old and removed from the media.

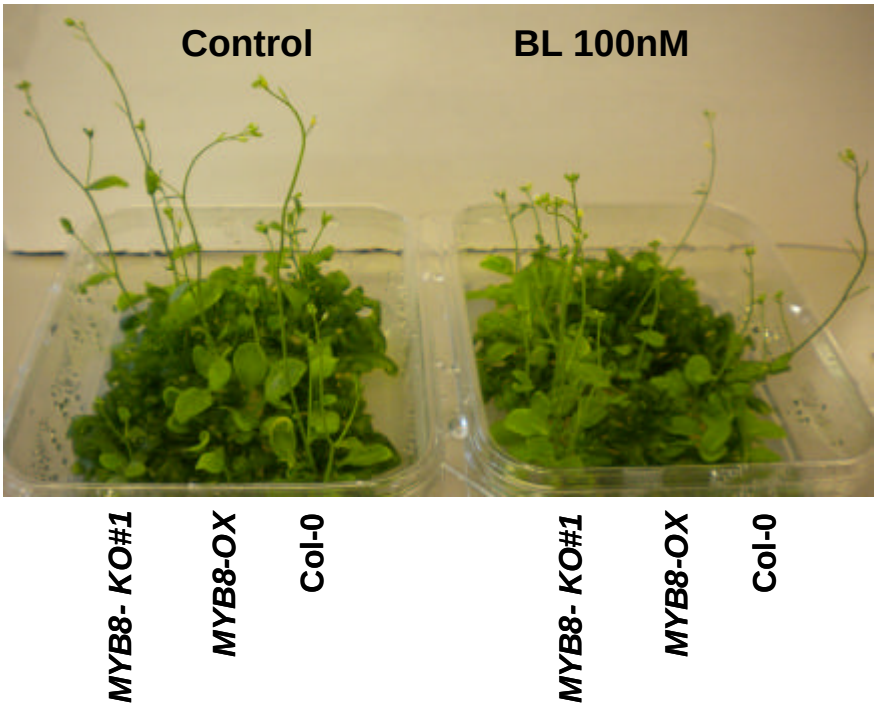
A


Control
BL 100nM
MYB6- KO#1
MYB6-OX
Col-0
MYB6- KO#1
MYB6-OX
Col-0

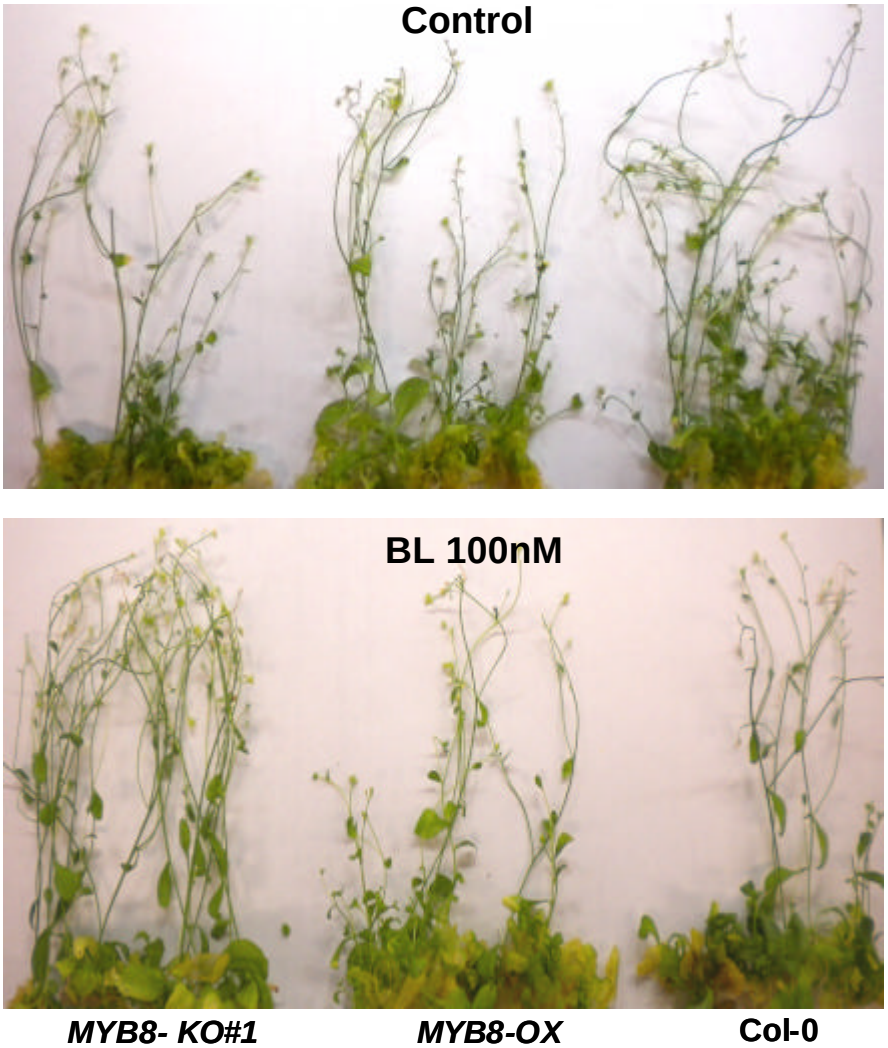
B


Control
BL 100nM
MYB6-KO#1
MYB6-OX
Col-0

C


Control
BL 100nM
MYB8- KO#1
MYB8-OX
Col-0
MYB8- KO#1
MYB8-OX
Col-0

D


Control
BL 100nM
MYB8- KO#1
MYB8-OX
Col-0

Thus, BL appears to repress inflorescence elongation in the *MYB6-KO#1* line, but enhance it in Wt and *MYB6-OX* lines. This could be a direct effect on inflorescence elongation, or an indirect effect on flowering time, which would, in turn, impact on inflorescence elongation (bolting). Either way, mutation of *AtMYB6* seems to impact on the sensitivity of plants to brassinosteroids. Differences were not that obvious between the BL treatment and the control for the *MYB8-KO#1* and *MYB8-OX* lines. No real differences in inflorescence elongation were observed when the mutant and wild-type lines were compared under the same conditions. There was almost no change in the overexpression line *MYB8-OX* treated or non-treated (Fig. 3.7C-D). BL application showed some effects in younger seedlings (Fig. 3.7C). All the *MYB8-KO#1* plantlets had bolted at the time when only a few *OX* plants or wild-types had started inflorescence elongation in presence of the hormone in the media. This effect was consistent with observations of the same plants 3 weeks later (Fig. 3.7D), although the stimulation of inflorescence growth by BL was not very obvious in these lines compared to the experiment with the *MYB6-KO* and *OX* lines.

Because brassinosteroid mutants fail to undergo hypocotyl etiolation in the dark (Bishop and Koncz, 2002), the effects of BL on hypocotyl elongation of *AtMYB6* and *AtMYB8-KO* mutant alleles were tested. A typical etiolation response produces elongated hypocotyls, apical hooks and closed cotyledons in dark-grown seedlings. Dark-grown, 5 day-old *AtMYB6-KO* and *AtMYB8-KO* mutants showed similar elongation of their hypocotyls in response to BL as Wt (not shown).

3.3.5. Germination assay

When grown *in vitro*, mutant seeds looked to have different germination rates therefore, knockout and overexpression lines of *AtMYB6* and *AtMYB8* were screened for alteration in germination in order to establish whether any differences in germination rates were statistically significant.

Wild-type, knockouts and overexpression lines of *AtMYB6* and *AtMYB8* were grown in MS and their efficiency of germination was scored, as the time when radicles emerged. Figure 3.8, shows the differences in germination rate measured. After 3 days, significant differences were observed in the percentage germination of seeds of the different lines. Both *AtMYB6-KO* lines showed higher levels of germination after 3 days than controls, while for the *AtMYB8-KO* line there was a slightly higher level of germination, but this was not statistically significant. Both *AtMYB6-OX* and *AtMYB8-OX* lines showed lower levels of germination than Col-0. After 5 days the *AtMYB6* and *AtMYB8 KO* lines showed overall higher levels of germination than Col-0. The *AtMYB6-OX* line still had lower levels of germination than Col-0 at 5 days, but the *AtMYB8-OX* line had higher levels of germination than Col-0 indicating that the likely effect of *AtMYB8-OX* was to slow the rate of radicle emergence post imbibition.

3.3.6. Developmental aspects of *AtMYB6* and *AtMYB8*

AtMYB6 and *AtMYB8* knockout lines did not show obvious differences in appearance in comparison to wild-type plants. Only the KO allele of *AtMYB8*, *AtMYB8-KO#1*, displayed a weak phenotype being slightly taller than the Col-0 used as a control when plants were grown on soil under a long day photoperiod (Fig. 3.9A). Overexpression lines were generated by Dr. B. Lebouteiller where *AtMYB6* or *AtMYB8* cDNAs were expressed in Arabidopsis under the control of the strong CaMV 35S promoter. The expression level of both overexpressing lines was examined by quantitative RT-PCR analyses. Homozygous transgenic lines with highest levels of expression for each transgene were taken for further experiments. *AtMYB6* and *AtMYB8* overexpression produced phenotypic alterations especially in the 35S::*AtMYB6* transgenic Arabidopsis plants. When 35S::*AtMYB6* plants were grown on soil, they developed normally although rosettes and leaves from the transgenic line were smaller than those from the wild-type (Fig. 3.9D) although this became less obvious when mutant plants got older.

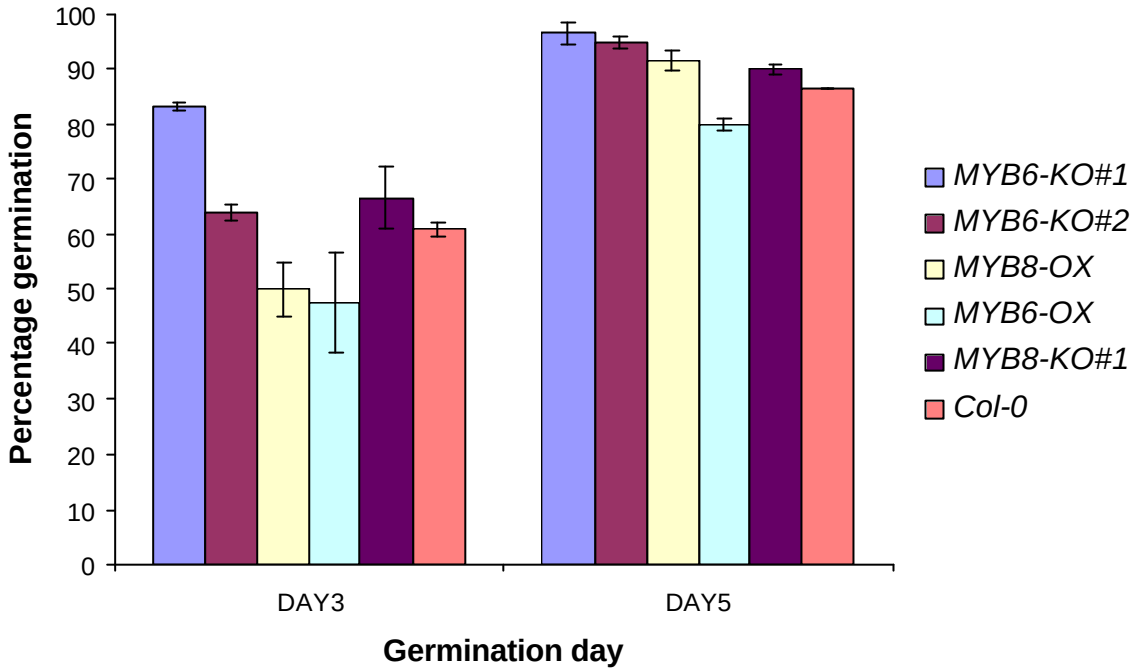


Figure 3.8. Germination efficiencies of wild type and *AtMYB6 and AtMYB8* mutants

Wild-type, knockout and overexpressed mutant seeds scored at 3 and 5 days after germination (radicle emergence). The result shows the average of three independent experiments and at least 80 seeds were used in each experiment. The average germination percentage ± SE of triplicate experiments was calculated.

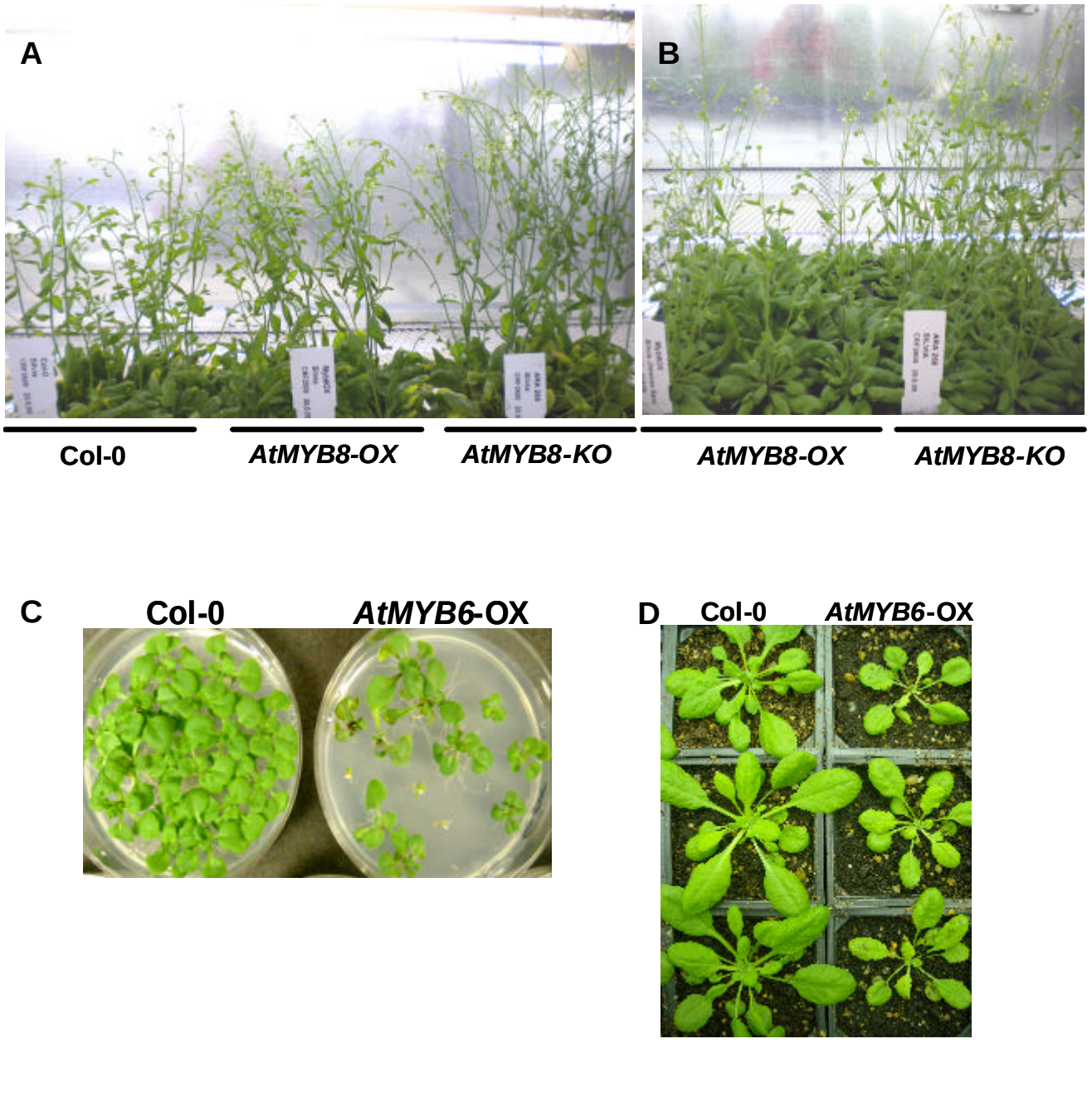


Figure 3.9. *AtMYB8* and *AtMYB6* phenotypes

(A) *AtMYB8* mutant lines showing differences in height compared to wild-type plants. (B) *AtMYB8-KO#1* developed the main inflorescences before than the *AtMYB8-OX*.plants. (C) Overexpression of *AtMYB6* gene altered the development of seedlings grown in vitro. (D) On soil, overexpression of *AtMYB6* produced slightly smaller rosettes compared to wild-type seedlings.

However, when the overexpression line AtMYB6-OX was grown on MS media, it displayed a range of phenotypes; from normal appearance, to seedlings that did not develop more than the first true leaves or those whose leaves, were malformed (Fig. 3.9C). Contrary to *AtMYB6* overexpression line, 35S::*AtMYB8* was phenotypically very similar to Col-0 plants (Fig. 3.9A and 3.9B) although this transgenic line developed its main inflorescence slightly later than the *AtMYB8-KO* line where more stem biomass could be observed (Fig. 3.9B).

The gain-of-function lines of *AtMYB6* and *AtMYB8* also showed a root phenotype when grown *in vitro*. *AtMYB6-OX* transgenic plants displayed a stronger phenotype with a reduction in root length as well as dramatically altered root morphogenesis compared to wild-type (Fig. 3.10A). The abnormal radial expansion of the roots of the *AtMYB6-OX* line appeared just above the meristematic zone which is consistent with the GUS expression of the *AtMYB6* gene along the root except in the meristem (Fig. 3.10B). In addition, the *AtMYB8-OX* line also showed a root phenotype; most of the *AtMYB8-OX* roots measured (n=16) had shorter distance from the root tip to the root hair initiation zone (Fig. 3.10C). The *AtMYB6-KO#1* allele did not show abnormal root morphology when observed under the microscope but displayed slightly shorter roots than Col-0 (Fig. 3.11A). A similar effect on root length was also shown by the *AtMYB8-KO#1* and and also by the overexpression lines where there was a decrease in root length in the overexpressers when compared to the wild-type (Fig. 3.11A). Consistent with the observed results is the statistical analysis of the root length which suggested slightly reduced root growth especially in the *AtMYB6-KO* (Fig. 3.11B) and demonstrated a very significant reduction in root growth in the *AtMYB6-OX* line. A reduction in root growth in the *AtMYB8-OX* line was also observed although, it was not as striking as *AtMYB6-OX*.

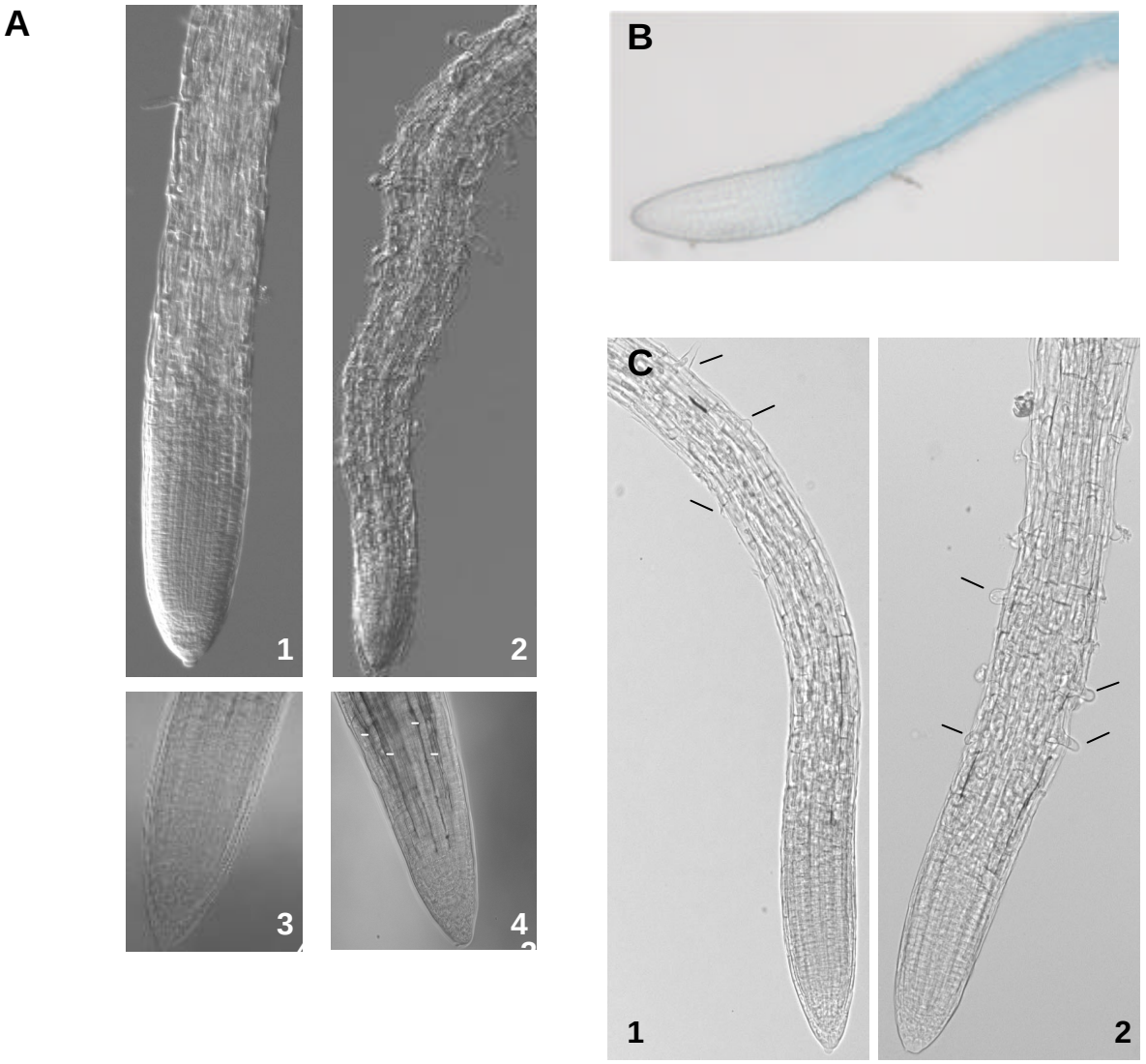
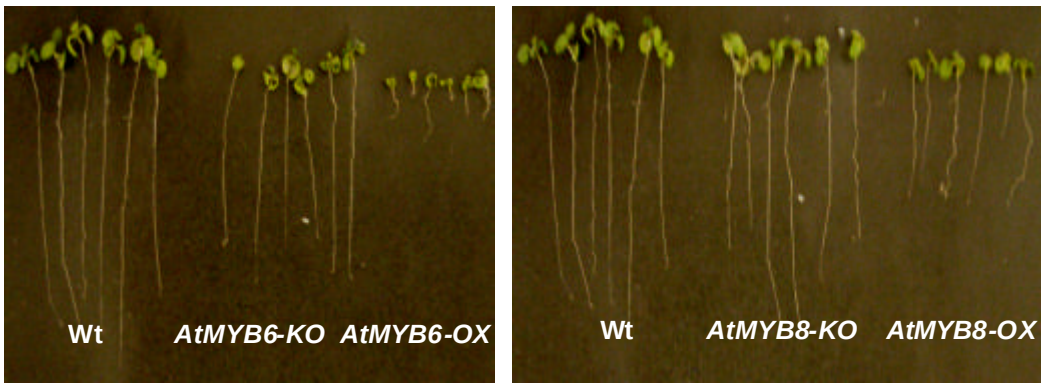


Figure 3.10. *AtMYB6* and *AtMYB8* root phenotypes

(A) Nomarski microscope images of ~3-5 day old roots cleared in Hoyer's solution. 2 and 4, images of the *AtMYB6-OX* root with 10x and 40x magnification. 2, distortion of the cells present in elongation zone. 4, possible (non tested) lignin accumulation indicated by arrows. 1 and 3, images of Col-0 displaying a normal development of the root. (B) Expression pattern of *AtMYB6prom-GUS* in root except in the meristematic zone. (C) Roots bright field pictures (10X) of 1, Col-0 and 2, *AtMYB8-OX* indicating less distance from the root tip to the root hair (arrows) initiation.

A



B

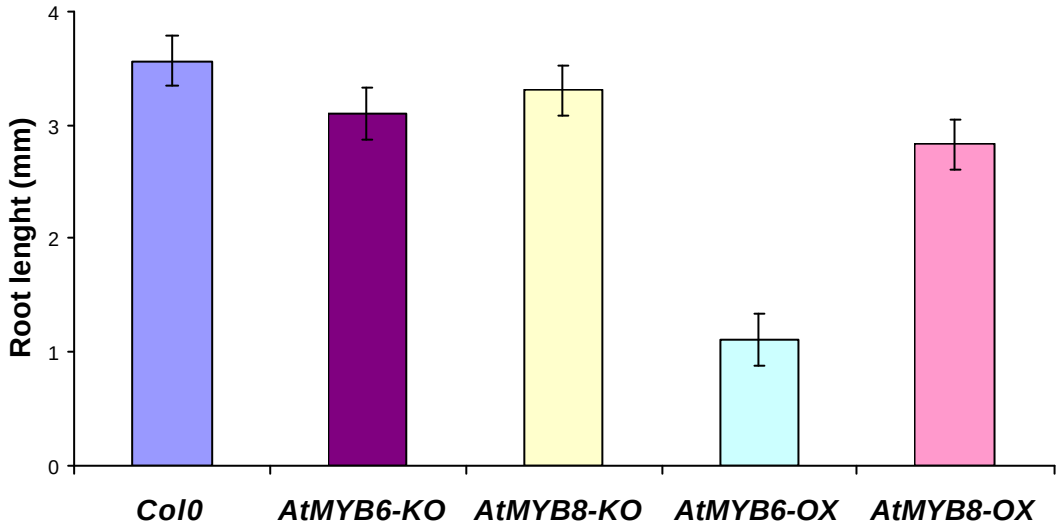


Figure 3.11. *AtMYB8* and *AtMYB6* root length

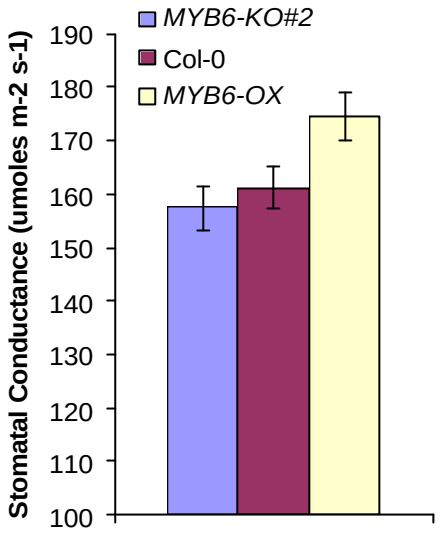
(A) Five day old seedlings grown vertically under long days photoperiod in MS supplemented with 1% of sucrose. Pictures show the knockout and overexpressed mutant lines from *AtMYB8* and *AtMYB6* genes. (B) Root length from seedlings grown under the same conditions as the plants from the above picture. Primary root length was measured from the root tip until the hypocotyl border using a plastic rule. Error bars represent SE; $n = 16$.

3.3.7. Evidence that *AtMYB6* maybe involved in determining stomatal movement

AtMYB6 expression was detected in stomata (section 3.3.2). Stomata regulate gas exchange and are distributed across the leaf epidermis.

Expression of *AtMYB6* in guard cells raises the possibility of this transcription factor as a putative modulator of the stomatal activity. Three other MYB transcription factors involved in stomatal movements have been described (Cominelli *et al*., 2005; Liang *et al*., 2005; Jung *et al*., 2008). To test this hypothesis, several approaches were taken. Measuring the stomatal conductance can give an idea of the relative size of the stomatal aperture; the higher the evaporation rate, the higher the conductance of the leaf. A portable leaf porometer was used for measuring conductance in different individual plants, mutants and wild-type lines. Results showed clearly that stomatal conductance in the *AtMYB6* overexpression line was significantly higher than in the wild-type; the pore aperture in the overexpression plants was bigger than in controls (Fig. 3.12A). In contrast, the *AtMYB6#2* loss-of-function allele did not show a significant difference in pore size compared to the control. The experiment was set up also with the knock-out *AtMYB6-KO#1* but seedlings died before making the assay and there was no time to repeat the assay. The observation of an effect of *AtMYB6-OX* on pore size, but the lack of effect of mutation of *AtMYB6* on pore size could be the result of redundancy in function between *AtMYB6* and *AtMYB8* meaning that the loss of *AtMYB6* functionality in the *MYB6-KO#2* (which also was not the best allele to test due its partial expression) can be complemented by AtMYB8 activity. Incomplete stomatal closure can lead to transpirational water loss. In Chapter 2, water loss measurements were described. However, this time, I also compared the results with the *AtMYB6* overexpression line. I measured water loss from detached mutant and wild-type rosette leaves at room temperature every two hours. No significant differences in water loss were found between the *AtMYB6* mutants/overexpression lines and the wild-type lines (Fig. 3.12B) despite the measured differences in stomatal conductance in unstressed leaves.

A



B

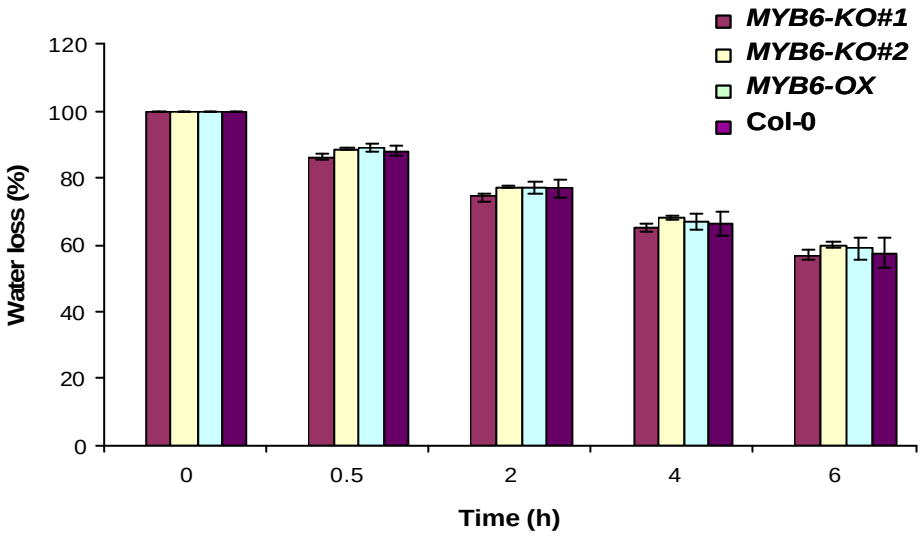


Figure 3.12. AtMYB6, stomatal guard cells and rate of water loss

(A) Stomatal conductance measured with a porometer. The results are means ± SE of 35 measurements per line. (B)Time course of water loss from rosettes expressed as a percentage of the initial fresh weight at indicated intervals. Each bar represents the mean of 4 measurements with standard errors.

3.3.8. Crossing of *AtMYB6* and *AtMYB8 KO* mutants

To test the hypothesis that *AtMYB6* and *AtMYB8* may play redundant roles in the control of expression of their target genes, crosses were performed between plants homozygous for the *MYB6* and *MYB8* single mutants: *MYB6-KO#1* x *MYB8-KO#1*, *MYB6-KO#2* x *MYB8-KO#1*, *MYB8-KO#1* x *MYB6-KO#3*.

The doubly hemizygous F1 progeny were grown to maturity and allowed to self-pollinate. Siliques containing the segregating F2 embryos were obtained. Because the *MYB6* and *MYB8* genes are located on separate chromosomes and are hence genetically unlinked, it was expected to find a 15:1 ratio of F2 progeny with other genotypes to that of the *MYB6* and *MYB8* doubly homozygous mutant.

Candidates from F2 plants were genotyped by PCR. At least 60 plants from each cross were analyzed, (in some cases 120); visual inspection of the plantlets did not reveal any novel phenotypes.

As both the *MYB6* and *MYB8* mutants, contained a T-DNA insertion with selection markers (kanamycin and sulfadiazine resistance genes, respectively), the kanamycin resistance marker was used for further genetic characterizations and segregation analysis. Unfortunately, a probable silencing effect on the *nptII* gene did not allow me to retrieve any information from the segregation analysis. Many seedlings were partially green and did not show a clear sensitivity to kanamycin making it almost impossible to derive a segregation ratio.

On the other hand, technical problems with the PCR-based genotyping complicated my analysis. Oligo pairs which had worked well for genotyping the single mutant lines did not work for screening for the double mutant, presumably because of the similarity between the single gene-specific primers, which, when used in combination with the T-DNA primers could not distinguish between insertions in the two genes. Consequently, new and more specific primers were designed to avoid interference between detecting the T-DNA insertions in both genes. Even so, PCRs still gave problems to reproduce results and when these were analyzed, ratios did not fit any Mendelian law. I

was unable to recover any homozygous double mutant lines despite several screening attempts of the F2 and F3 generations.

To detect possible lethality, siliques of double heterozygous plants were inspected for the presence of aborted embryos or arrested ovules. No defective seeds showing the expected segregation ratios (1:15 lethal to viable) were observed.

3.4. Discussion

To decipher the possible roles of *AtMYB6* and *AtMYB8* several approaches, using reverse genetics, have been taken. Multifunctionality within the MYB transcription factor family (Jin and Martin, 1999), makes it difficult to predict the function of these two transcription factors, especially when there was no clear altered phenotype in the KO mutants that could indicate in which biological processes they could be involved.

Consistent with their role as transcription factors, *AtMYB6* and *AtMYB8* full-length proteins were localized to the nucleus in stable transformants of Arabidopsis. On the other hand, expression analysis of these genes by promoter fusions to GUS and GFP suggested differences in the cells in which they are active. This could indicate differences in the biological processes they regulate.

AtMYB6prom::GFP was observed in guard cells of the stomata leading to the hypothesis of a putative role of this transcription factor in stomatal movements. To my knowledge, only three other MYB factors are known to be involved in stomatal aperture/closure. *AtMYB60* is involved in light-induced opening of stomata, *AtMYB61* works in dark-induced stomatal closure whereas *AtMYB44* is responsible for ABA-induced stomatal closure (Cominelli *et al*., 2005; Liang *et al*., 2005; Jung *et al*., 2008). By measuring the stomatal conductance, an altered stomatal function was detected in the *AtMYB6* gain-of-function line, where a high stomatal conductance implied that the overexpression line had increased stomatal aperture. In contrast, the loss-of-function mutant did not show a significant difference in stomatal aperture

compared to wild-type. This could be due to functional redundancy, such that AtMYB8 can functionally complement the loss of AtMYB6 function in regulating stomatal aperture. However, I detected no expression of *AtMYB8* in stomata using an AtMYB8:GFP fusion. This result points to *AtMYB6* as a possible regulator of the transpiration rate in plants. In agreement with that idea, is the fact that *AtMYB6* is strongly expressed in hydathodes (GUS assay), structures present in the leaf margins of Arabidopsis that mediate the guttation process related to the water status of the plant (*i.e.* stomatal opening and leaf water potential) (Takeda, F. 1989). However, no significant changes were detected between mutants (note that *MYB6-KO#2* was not the best allele to test due its partial expression), overexpression lines and wild-type lines when water loss was measured in rosettes.

The examination of GUS staining revealed that *AtMYB6* and *AtMYB8* were expressed at all stages of plant development analyzed from young seedlings to flowering plants. Thus, at the tissue and organ levels, *AtMYB6* and *AtMYB8* had similar expression patterns. The overlapping expression of the two genes in roots, some vascular tissues, floral primordia and stigmatic tissues further suggests that the two proteins might function redundantly in many different tissues of the plant. To address this possibility, I attempted to produce a double mutant by crossing the *MYB6-KO* and *MYB8-KO#1* lines. Although several crosses were genotyped, no double homozygous *AtMYB6- AtMYB8 KO* mutants were identified. Some plants homozygous for the *MYB6-KO* and hemizygous for *MYB8-KO#1* allele were retrieved. However, progeny of these plants did not show the expected segregation ratios (Mendelian ratios) for these alleles. The fact that I was unable to retrieve the homozygous *MYB8-KO#1* allele in the homozygous MYB6-KO background suggests that the double mutant may be lethal, and that together, the two genes serve an essential function. However, the predicted ratio of 1:15 lethality of embryos was not observed in examination of developing embryos in siliques.

Some treatments such as cold, drought, salt, ABA or PAC, seem to alter the expression levels of *AtMYB6* and *AtMYB8* genes as predicted by the

Genevestigator database. However, as described in the previous chapter, they did not have any effects on the visual phenotype of *AtMYB6* and *AtMYB8* mutants. More expression data available through this database, offered some vague insight into the possible functions of *AtMYB6* and *AtMYB8*. To characterize further these possibilities, the effects of application of biotic stresses to the mutants were investigated.

The roles of some MYB transcription factors in pathogen infection response have been reported (Vailleau *et al*., 2002; Mengiste *et al*., 2003). *MYB6-KO#1* and *MYB8-KO#1* did not show any hypersensitivity to the application of different *P. syringae* pathogen strains nor to elicitor treatment with flg22 and elf18 suggesting that these transcription factors are unlikely to be regulating pathogen defence responses, even though they are induced by biotic stress. Consistent with this finding, oxidative treatments that generate ROS (reactive oxygen species), chemical components produced in plants by necrotrophic pathogens, showed no differential impact on mutants compared to wild-type plants.

Genevestigator analysis also suggested up or down-regulation in AtMYB6 and AtMYB8 expression depending on the BL treatment applied, so an analysis of mutant responses to BL treatment was undertaken. *AtMYB6* and *AtMYB8* knock-out and overexpression lines were grown in media supplemented with brassinolide, the most active brassinosteroid (BR). *AtMYB6* and *AtMYB8* mutants did not display the characteristic dwarfism of the BR-deficient or –insensitive mutants (Asami *et al*., 2005) under light or darkness. The test showed phenotypic variability in the mutants, differentially altering stem elongation in knockout and overexpression lines when BL was in the medium suggesting a possible role of these MYB factors in the BR signaling pathway. Absence of a clear BR-related phenotype may indicate that AtMYB6 and AtMYB8 are not involved directly in the BR pathway but may be somehow related to it by regulating indirectly components of the BR signal transduction pathway.

In addition, the later availability of the MYB6 and MYB8 overexpression lines allowed me to observe new developmental effects of *AtMYB6* and *AtMYB8* overexpression on germination and growth. The phenotypes of these lines have been included here as part of a general characterization of the *AtMYB6* and *AtMYB8* phenotypes although further explanations about the phenotypes are included in the discussion of the next chapter. Changes in germination rate, growth differences and root morphogenesis can be brought about in plants by multiple factors or conditions. It is difficult to draw conclusions about the causes of these phenotypes resulting from overexpression of *AtMYB6* and *AtMYB8* transgenic lines without additional experiments. Results reported in Chapter 4 offer an explanation/hypothesis for the developmental phenotypes described in this chapter and therefore, are discussed in Chapter 4.

CHAPTER 4

Chapter 4: Putative roles of *AtMYB6* and *AtMYB8* in phenylpropanoid metabolism

4.1. Introduction

The result of my attempts to clarify the function of *AtMYB6* and *AtMYB8* in Arabidopsis using the *hos10* mutant was the demonstration that the *hos10* mutant phenotype was not the result of loss of function of *AtMYB8* as published by Zhu *et al*., (2005). Consequently it was necessary to identify the biological functions of these two transcription factors *de novo*. Analysis of the visible phenotypes of the knockout and overexpression lines for the two genes suggested that these transcription factors impacted root growth and germination rates and also had an effect on sensitivity to brassinosteroid-influenced elongation of the inflorescence stem (Chapter 3). However, none of these observed phenotypic effects of loss of AtMYB6 or AtMYB8 function or their ectopic expression provided real insight into the target genes controlled by either transcription factor or the primary biological processes they regulate.

Both AtMYB6 and AtMYB8 proteins align in subfamily 4 of the R2R3 MYB transcription factor tree (Kranz *et al*., 1998, Stracke *et al*., 2001). Other members of this subfamily of proteins have been shown to play roles in regulating phenylpropanoid metabolism (Jin *et al*., 2000; Preston *et al*., 2004), Generally, R2R3 MYB proteins that cluster in the same subfamily have related biochemical functions due to the strong structural similarities in their DNA binding domains (meaning that they recognise the same target motifs) and the conserved motifs in their C-terminal domains (such as the conserved EAR transcriptional repression domain in members of R2R3 MYB subgroup 4). In addition, analysis of *AtMYB6-OX* and *AtMYB8-OX* lines in a related project ongoing in the laboratory, using NMR fingerprinting, had revealed differences in the phenolic profiles of leaf tissue between these lines and control plants (B. Lebouteiller and L. Shintu personal communication). These two lines of evidence suggested that AtMYB6 and AtMYB8 might be involved in regulating

phenylpropanoid metabolism, and suggested that these possible functions should be investigated further.

Phenylpropanoids are plant secondary metabolites with a wide range of functions including mechanical support (lignins), protectants against biotic and abiotic stress (phytoalexins, antioxidants, and UV-absorbing compounds), pigments and co-pigments (anthocyanins, flavonols and flavones), and signals (for example, flavonoid nodulation factors). Besides these physiological functions, there is a growing interest in these secondary metabolites due to their potential benefits for human health (Lepiniec *et al*., 2006; Du *et al*., 2009). Phenylpropanoids derive from the aminoacid phenylalanine, a product of primary metabolism, which serves as the precursor for the general phenylpropanoid pathway (consisting of three enzymes; phenylalanine ammonia lyase [PAL], cinnamate 4-hydroxylase [C4H] and p-coumaroyl CoA ligase [4CL]) which supplies the precursors for the many branches of the pathway that synthesise these different phenylpropanoid metabolites (Chapple *et al*., 1992; Dixon and Paiva, 1995).

Flavonoid biosynthesis is one branch of the general phenylpropanoid pathway, (that also leads to lignin and sinapate ester biosynthesis) and is itself composed of a number of branches which lead to isoflavonoids, flavonols, proanthocyanidins (condensed tannins), and anthocyanins (Fig. 1.7 and 1.8 from Chapter 1). The chalcone synthase gene encodes the first enzyme committed to flavonoid biosynthesis (Weisshaar and Jenkins, 1998). Over 6000 different flavonoid compounds have been described, and their best known roles are as pigments and UV light protectants. The enzymes responsible for the formation of these natural products are responsive to a range of environmental stimuli (particularly stress signals) and expression of the genes that encode these enzymes is usually under the control of MYB transcription factors. In the case of anthocyanin and proanthocyanin biosynthesis the R2R3MYB transcription factors that control expression of the biosynthetic genes interact in a regulatory complex with bHLH transcription factors and WD40 repeat proteins (Ramsay and Glover, 2005).

In this chapter, I report variations in phenylpropanoid levels resulting from alterations in the activity of the *AtMYB6* and *AtMYB8* genes. Results suggest that both genes can modify this biosynthetic pathway in Arabidopsis. Overexpression of either gene in tobacco plants also produced changes in the phenylpropanoid profiles. These changes could be responsible for some of the visible phenotypes described in Chapter 3.

4.2. Materials and methods

4.2.1. Tobacco plant material

4.2.1.1. Tobacco transformation

Nicotiana tabacum var. Samsun NN was used for transformation. Constructs containing the ORFs of *AtMYB6* and *AtMYB8* were the same as those used for Arabidopsis transformation (Chapter 3, 3.2.1 section) but transformation was mediated using *Agrobacterium tumefaciens* strain LBA4404 in sections cut from leaves of 4-5 week-old tobacco plants.

Five young, healthy, fully-expanded leaves were used per transformation. Leaves were first sterilized with a 10% bleach solution for 10 minutes and afterwards rinsed 3 times with sterile distilled water. A 100 ml culture of *Agrobacterium tumefaciens* containing the construct was grown o/n at 28ºC in a shaker with the corresponding antibiotics and then centrifuged at 7000 rpm for 20 min. The pellet was resuspended in 25 ml of liquid MS and poured into a 140 mm Petri dish. Tobacco leaves were chopped using a scalpel, avoiding the main vein, and placed into the Petri dish containing the Agrobacterium-MS suspension. Leaf strips were transferred to 90 mm Petri dishes with MS 0.8% agar (15-20 plates/construct) ensuring that the surfaces of the tissue contacted the agar. Plates were sealed with surgical tape, and incubated for 48 hours in darkness at 25ºC. The strips were then transferred to new MS 0.8% agar plates supplemented with BAP 1 µg/ml, NAA 0.1 µg/ml, cefotaxime 500 µg/ml and kanamycin 100 µg/ml. The plates were taped and left

in the growth cabinet in the light. Callus formed on the cut edges of the leaf sections after 3-4 weeks. Subsequently numerous shoots started to form from the regions of callus. The calli with shoots of greenest appearance were cut and placed in new plates with MS medium agar (0.8%) with all the supplements previously described and under the same growth conditions until they formed shoots large enough to transfer to rooting medium. These explants were transferred to individual containers containing half-strength MS medium without hormone supplements (to encourage rooting) but with kanamycin (100 µg/ml) and after a month, plantlets which had developed roots were transferred to potting compost and grown on in the greenhouse.

4.2.2. Analysis of soluble phenolics

4.2.2.1. Phenolic extraction

For the analysis of soluble phenolics, Arabidopsis seedlings were grown in MS supplemented with 1% sucrose in disposable plastic boxes or in soil at 23°C with a 16 h photoperiod in a growth chamber. The whole plant, including roots, was collected when the first flowers opened. A mix of 6 plantlets per line was used for each extraction. For tobacco, potential transformants were identified by their ability to produce roots and grow on kanamycin-containing media. Putative transgenic plants that formed roots were grown on in soil with a 16 h photoperiod in the glasshouse and 2 young leaves from several independent transformants were collected from 6 week-old plantlets.

Tissue was freeze-dried and ~ 10 mg of powdered plant tissue was extracted with 300 µl of 70 % methanol HPLC grade. The sample was vortexed and 95 µl distilled H_2O plus 5 µl of 10mM methyl vanillate (Sigma: dissolved in methanol and used as internal standard) were added. The extract was clarified by centrifugation at high speed for 10 sec and filtered through a Costar Spin-X centrifuge tube with a 0.22-µm nylon membrane (Corning, NY) to remove cell debris before injection.

4.2.2.2. LC-MS analysis

Analysis of soluble phenolics was carried out by liquid chromatography-mass spectrometry (LC-MS). Each sample (50µl) was mixed with 50µl water in small glass inserts in HPLC vials before injecting 10 µl. The instrument was a DecaXPplus ion trap mass spectrometer attached to a Surveyor HPLC system (Thermo Fisher). Phenolics were separated on a 100×2mm, 3µm particle size Luna C18 reverse phase column (Phenomenex) using the following gradient of methanol versus 0.1% formic acid in water, run at 30ºC and 250µl.min^{-1}. Gradient conditions were: time 0.2% MetOH; 30 min, 70% MetOH; 30.5 min, 2% MetOH; 38 min, 2% MetOH.

Detection was by UV at 520 nm for anthocyanins and 214 nm for everything else (bandwidth 9 nm), but full spectra were also collected from 200-600 nm, from which extracted wavelength chromatograms could be made for any other analytes. Detection was also by electrospray MS in positive mode, with data-dependent MS2 and MS3 in which the most abundant ions were chosen automatically for fragmentation at 35% collision energy and an isolation width of 4.0 (this combination will fragment most normal covalent molecules, and will trap most things, but will occasionally trap nearby ions and isotope peaks if the masses differ by only 1 or 2). Dynamic exclusion was used to ensure that once an ion had been selected twice within 0.5min, it would be ignored for 0.5min while the next most abundant ions would be analysed instead. Spray chamber conditions were 5.2 kV spray voltage, 50 units sheath gas, 5 units auxiliary gas, and 350ºC capillary temperature.

4.2.3. UV-B treatment

Seeds were sown *in vitro* on MS medium, placed 2 days at 4°C, then transferred to the growth chamber (23°C, 16 h light 8 h dark) for 10 days. The UV treatment was performed 20 cm distance with the lids of the Petri dishes removed under a short wave UV lamp (Transilluminator 220 V, model TM-20, UVP Upland, CA) for 20 min in a flow hood cabinet. The plantlets were then transferred back to the growth chamber for 8 days before observation.

4.2.3.1. Chlorophyll measurement

Seedlings from plates (UV-treated) were separated from roots, weighed and placed in a 2 ml centrifuge tube. To obtain a homogenous mix, tissue was ground in 1 ml 80% acetone using a ball mill (GenoGrinder, BT&C/OPS Diagnostics). To remove cell debris, the samples were centrifuged for 5 min at 14,000 g. For chlorophyll quantification the absorbance of the supernatant was measured at wavelengths of 645 and 663 nm (Spectra Max 340 PC, Molecular Devices). Total chlorophyll was calculated as follows:

Total chlorophyll concentration (µg/ml) = OD_{645} x 20.2 + OD_{663} x 8.02

The resulting value was divided by the fresh weight of tissue to obtain the chlorophyll concentration as ng/mg fresh weight (Arnon, 1949).

4.2.4. Quantitative RT- PCR

RNA was extracted from two central leaves of 5-week old rosettes grown on soil under a long day photoperiod.

For analysis of gene expression, first-strand cDNA was synthesized from 1 µg of RNA using SuperScript III Reverse Transcriptase (Invitrogen) and an oligo dT primer, according to the manufacturer's instructions (General materials and methods section).

For real time PCR, 3 µl of cDNA (5 ng/ µl) was combined with SYBR Green JumpStart master mix (Sigma). Expression levels were recorded for the genes PAL1 (*At2g37040*), C4H (*At2g30490*), C3H (*At2g40890*), 4CL3 (*At1g65060*) and CHS (*At5g13930*). Cyclophilin was used as endogenous control. Thermocycling was initiated by 4 min incubation at 95°C followed by 40 cycles (95°C for 10 s; 64°C for 30 s; 72°C for 15 s). PCRs were performed in three technical replicates (for each of the three biological replicates) with a PTC-200 Peltier Thermal Cycler (MJ Research, Waltham, MA), and the data were collected and analyzed with Chromo 4 Continuous Fluorescence detection system.

4.2.5. Protein array

4.2.5.1. Plasmid constructs

AtMYB6 and *AtMYB8* were cloned in frame in the Gateway ENTRY vectors pDONR207 and 201 respectively (see section 3.2.1.). An LR recombination reaction for each was performed as described in the Gateway manual (Invitrogen/Life Technologies) using pDEST-TH1 as destination vector (Hammarstrom *et al*., 2002). This vector is a Gateway version of the original pMAL-c2a (New England Biolabs). The cloned gene is inserted downstream from the *malE* gene of *E. coli*, which encodes maltose-binding protein (MBP), resulting in the expression of an MBP fusion protein.

Resulting expression clones were transformed into DH5α cells and plated on 100 mg/ml ampicillin. Colonies were confirmed by sequencing with specific primers, SJS56 for both proteins and SJS1 and SJS62 for *AtMYB6* and *AtMYB8* respectively to ensure that proteins were in frame with the MBP tag.

4.2.5.2. Protein expression

Expression was carried out in the strain BL-21 (DE3) Codon Plus. Freshly grown single colonies were picked and grown overnight at 37°C, transferred to 50 ml of L medium (supplemented with glucose) with ampicillin (100 µg/ml) and cloramphenicol (10 µg/ml) in a 250 ml flask, and cultured until the A_{600} reached approximately 0.8. Recombinant protein expression was induced by adding IPTG to 1 mM and growing on for 4 h.

The cells were harvested by centrifugation (6,000 rpm, 20 min, 4°C), and the pellets were resuspended in lysis buffer containing 20mM Tris-HCl pH 7.4, 200mM NaCl, 1mM EDTA and lysozyme 1 mg/ml. The sonicate was clarified by centrifugation at 15,000 *g* for 20 min at 4°C, and the supernatant as well as the pellet were recovered. The pellets were suspended in SDS 2% final concentration. Then, the samples were prepared for Bio-Rad SDS-PAGE analysis to evaluate the protein composition of each preparation. Gels were stained with Coomassie Brilliant Blue (InstantBlue, Expedeon).

Crude proteins extracts were analyzed directly for their preferred binding motifs using the array facility in the Research Centre CNB-CSIC of Madrid (Dr. R. Solano, Dr M. Godoy). Dr Marta Godoy performed the binding of the MYB proteins to the array following the protocol described by Berger and Bullyk (2009).

4.2.5.3. Western blot

Protein samples were separated by SDS-PAGE under reducing conditions using the 1X MOPS buffer (1 M Tris base, 2% SDS and 20 mM EDTA were mixed for a 20X buffer). The resolved proteins were blotted onto a 0.45 µm nitrocellulose membrane (Schleicher&Schuell), blocked in PBS buffer (1.37 M NaCl, 14.7 mM KH_2PO_4, 78.1 mM Na_2HPO_4 and 26.8 mM KCl were mixed for a 10X MOPS buffer with pH adjusted to 7.2-7.3) with 2% dried milk powder, and probed with a 1:10000 dilution of primary antibody Anti-MBP (Sigma), followed by incubation with the anti-rabbit AP conjugated secondary antibody (1:5000, Sigma). NBT/BCIP Western blotting detection reagents were used for detection.

In the pull-down experiment that was undertaken proteins were instead electrotransferred with a BioRad system onto Hybond PVDF 0.45 µm (GE Healthcare). Membrane blocking with antibody incubation was performed in TBST buffer (20 mM Tris–HCl, pH 7.6, 137 mM NaCl, 0.1% [v/v] Tween-20) containing 5% (w/v) non-fat dry milk. A 1:1000 dilution was used for the GFP (B-2) sc-9996 (Santa Cruz Biotechnology, inc.), a mouse monoclonal antibody raised against amino acids 1-238 representing full length GFP (green fluorescent protein) of *Aequorea victoria* origin. Proteins were detected using a chemiluminescence detection kit (ECL Plus Western Blotting Detection System, GE healthcare).

4.2.6. Protein-protein interactions: AtMYB6 and AtMYB8 pull-down assays

Leaves of *Nicotiana benthamiana* agroinfiltrated with the different constructs (expressing AtMYB6::GFP and AtMYB8::GFP fusion proteins, see 3.2.1.1) were collected after 2 days and ground into fine powders in liquid nitrogen.

For each construct, 4 g of frozen tissue were suspended in 15 ml extraction buffer (150 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10mM EDTA) with 2% PVP (polyvinylpyrrolidone), 10mM DTT (dithiothreitol), 1% Igepal (octylphenyl-polyethylene glycol) and a protease inhibitor mix containing 2 mM orthovanadate, 20 mM molybdate, 20 mM NaF, a few crystals of PMSF plus a protease inhibitor cocktail tablet. The mix was then incubated for 15 min at 4ºC with constant shaking. Total protein crude extracts were prepared by centrifuging the suspension at 20,000 *g* for 20 min at 4°C. The resulting supernatants were filtered through miracloth and mixed in a pre-equilibrated column (with equilibration buffer; 150 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM EDTA) containing Anti-GFP fused to GST beads (a gift from Dr. Vardis Ntoukakis). After 2 hours incubation at 4ºC with gentle shaking, the extract was subsequently passed through the column and followed by a 10 ml wash with equilibration buffer plus 10 mM DTT. Proteins were eluted with elution buffer containing 150 mM Tris-HCl, pH 7.5, 1 M NaCl and 10mM DTT. The concentration of the eluted proteins (10 ml) was estimated by adding 30 µl of StrataClean Resin (Stratagene) with gentle shaking for 30 minutes at 4ºC. After centrifugation at 6,000 rpm for 10 min, the beads were resuspended in 1 ml sterile deionized water and centrifuged again to recover the beads and to separate them from the eluted proteins. Eluates (~ 40 µl) were mixed with sample buffer and boiled for 10 min prior to loading the SDS-PAGE polyacrylamide gel (Invitrogen).

4.3. Results

Preliminary data from Dr. B. Lebouteiller suggested that AtMYB6 and AtMYB8 transcription factors may be involved in phenylpropanoid metabolism,

a complex pathway origin giving rise to multiple and varied secondary metabolites (Weisshaar and Jenkins, 1998) (see Fig. 1.7 and 1.8 from Chapter 1). These data gave similar metabolic profiles for the Arabidopsis lines overexpressing *MYB* subfamily IV genes (*AtMYB4, AtMYB6, AtMYB7, AtMYB8*). The NMR analysis showed changes between overexpression lines and the wild-type in levels of certain compounds synthesized by the phenylpropanoid pathway (Fig. 4.1). In addition, qPCR analysis of the AtMYB6 and AtMYB8 overexpression lines compared to the wild-type demonstrated that the sinapoyl malate content was correlated with the expression level of AtMYB6 and AtMYB8. However, kaempferol content was correlated only in AtMYB8 overexpression lines (Fig. 4.2) An overview of these results suggested that the transcription factors from subfamily IV were negative regulators of sinapoylmalate synthesis but positive regulators of synthesis of kaempferol derivatives indicating that these roles should be further investigated for *AtMYB6* and *AtMYB8* genes.

4.3.1. UV-B light treatment

Since flavonoid compounds (which are phenylpropanoids) have been implicated in protecting plants from the damaging effects of UV-B radiation (Arvind and Jitendra, 1997), I tested the response of *AtMYB6* and *AtMYB8* knockouts and overexpressed lines to UV-B light.

To determine whether the lack of epidermal UV-light absorbing metabolites could be visualized *in vivo*, the mutant seedlings were treated with UV-B radiation. As controls, the wild-type Col-0 and the *AtMYB4*-overexpressing line which has been characterized by having an enhanced sensitivity to UV-B (Jin *et al*., 2000) were used. Ten day-old mutant seedlings and Col-0 were irradiated with the short wave UV-B source at 20 cm distance from the seedlings with the lids of the Petri dishes removed for 20 minutes. After exposure, the Arabidopsis seedlings were grown for one week more prior to scoring and photography.

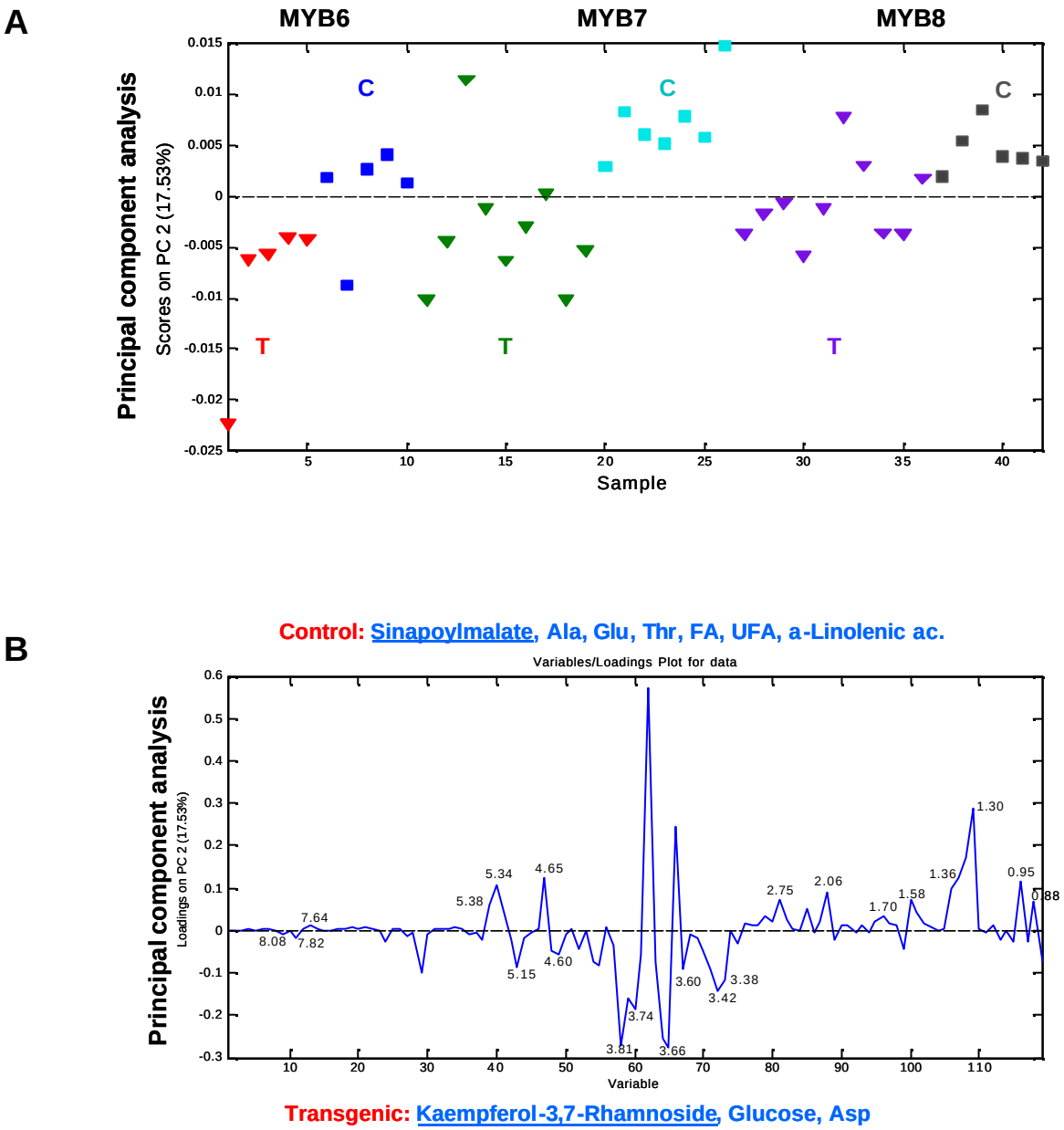
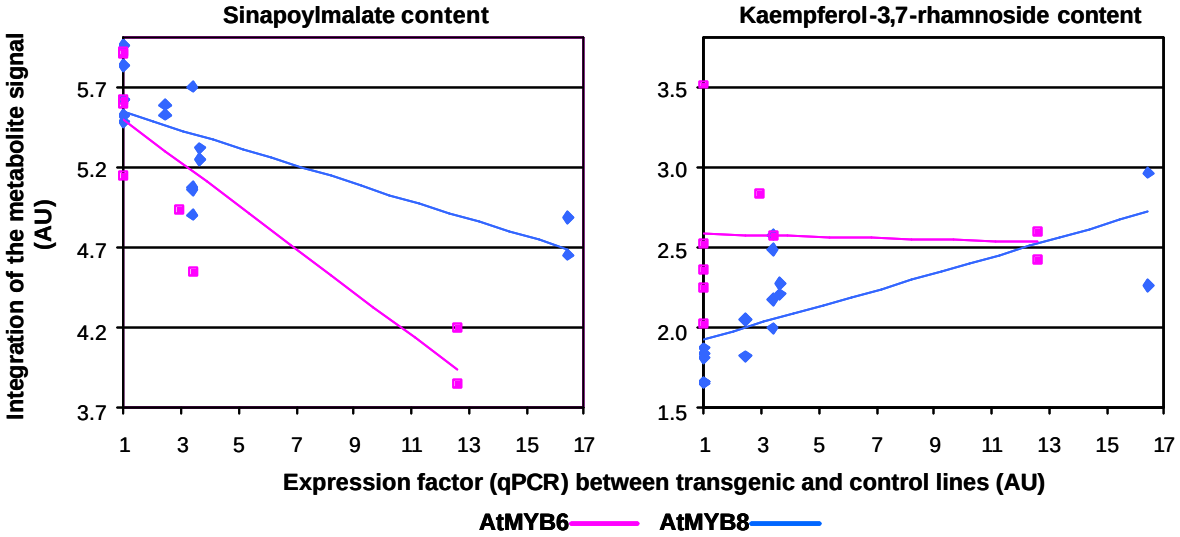


Figure 4.1. NMR analysis of some members from subfamily 4

(A) PCA metabolite analysis of heterozygous plants overexpressing *AtMYB6*, *AtMYB7* and *AtMYB8*. The second principal component enables to separate most of the transgenic samples T (triangles) from their control C (circles). There are a common set of markers allowing the discrimination between transgenics and controls for these three genes. (B) The study of the PC2 loadings shows that two secondary metabolites are significant; sinapoyl malate, in lower concentration in the transgenic, and kaempferol-3,7-rhamnoside both involved in the phenypropanoid pathway.

Figure courtesy of Dr. Shintu and Dr. Leboutillier (2008).



	Correlation coefficient of Pearson		Degree of freedom	Threshold for p<0.01
	Sinapoylmalate	Kaempferol-3,7-Rhamnoside		
AtMYB6	-0.877	-0.054	7	+/-0.798
AtMYB8	-0.737	0.686	14	+/-0.623

Figure 4.2. Correlation between *AtMYB6* and *AtMYB8* expression level and metabolite content

The sinapoyl malate content was correlated in AtMYB6 and AtMYB8 with the expression level however; kaempferol content was correlated only in AtMYB8. Sinapoyl malate and kaempferol-3,7-rhamnoside, are respectively the major soluble phenolic in *A. thaliana* leaves and the major flavonoid in plants. These data suggest that AtMYB6 and AtMYB8 could be regulators of phenylpropanoid metabolism although with some specificities for each transcription factor.

Figure courtesy of Dr. Lebouteillier (2008).

AtMYB6 loss-of-function lines showed visually slightly more damage than Col-0 seedlings, especially the mutant KO#1. The lesser effect for KO#2 could because the insertion in this line does not cause complete elimination of the *AtMYB6* transcript (i.e: KO#2 is not a null mutant) (Fig. 4.3A). In contrast, however, *AtMYB6OX* was hypersensitive to the UV-B light treatment like the *AtMYB4* overexpression line. In contrast, the *AtMYB8* knockout seedlings were slightly less affected by the UV-B treatment than the Col-0 seedlings, showing an increased resistance to the UV-B exposure in the assay. The overexpression line, *AtMYB8-OX,* was more sensitive than the wild-type to UV-B exposure although it was more resistant than *AtMYB4-OX*. The assay was repeated several times and results were consistent.

To quantify the damage caused by UV-B light in the different lines, chlorophyll levels were measured, since levels of this pigment decrease in response to enhanced UV-B radiation due to the repression of chlorophyll synthesis following UV-B exposure (Choi and Roh, 2003). The *AtMYB8-KO*#1 seedlings showed slightly higher levels of chlorophyll than the wild type in contrast with the mutant lines of *AtMYB6* whose pigment content was lower than that of the wild-type following exposure to UV-B (Fig. 4.3B). The reduced growth of the AtMYB6-OX seedlings following UV-B exposure could have been due to the fact that overexpression of AtMYB6 causes slower growth (Chapter 3 section 3.3.6). However, the reductions in growth can not explain the severe reduction in chlorophyll measured in this line following UV-B treatment, confirming that it is hypersensitive to UV-B. These results suggest that like AtMYB4, AtMYB6 and AtMYB8 negatively regulate the accumulation of UV-B absorbing phenylpropanoids (flavonoids and sinapoyl esters). The damage following exposure of AtMYB6-OX seedling to UV-B was great, suggesting that when overexpressed AtMYB6 significantly reduces the levels of these UV-B protecting phenylpropanoids, a conclusion supported by the metabolite fingerprinting analysis of this line (Fig. 4.1 and 4.2). The effect of overexpression of *AtMYB8* was considerably less pronounced. Metabolite fingerprinting did detect modifications in phenylpropanoid distribution in this

line. In addition, I detected a reduction in sinapoyl malate and flavonols in this line compared to wild type by LC/MS.

Unlike the situation for AtMYB4 knockout lines (Jin *et al*., 2000), AtMYB6KO1 and KO2 and AtMYB8KO1 did not show significant reductions in damage in response to UV-B irradiation, although the AtMYB8-KO mutant did show slightly more chlorophyll than wild type following exposure to the same levels of UV-B irradiation. This could be due to functional redundancy between AtMYB6 and AtMYB8 and possibly other members of subgroup 4 in mediating response to UV-B light stress.

My results suggest a possible inverse relationship between tolerance to UV-B irradiation and the activity of AtMYB6 and AtMYB8 which might work through these two transcription factors negatively regulating general phenylpropanoid metabolism and so influencing the levels of sinapoyl malate and flavonols available for absorbtion of UV-B. UV-B attenuation is mainly attributed to flavonoids and related phenolic compounds (sinapoyl esters) in Arabidopsis (Landry *et al*., 1995; Booij-James *et al*., 2000). The precise level of protection afforded by altering the activity of AtMYB8 and AtMYB6 will depend on their relative effects on flavonols and sinapoyl esters since these compounds are believed to contribute differentially to protection against UV-B light stress.

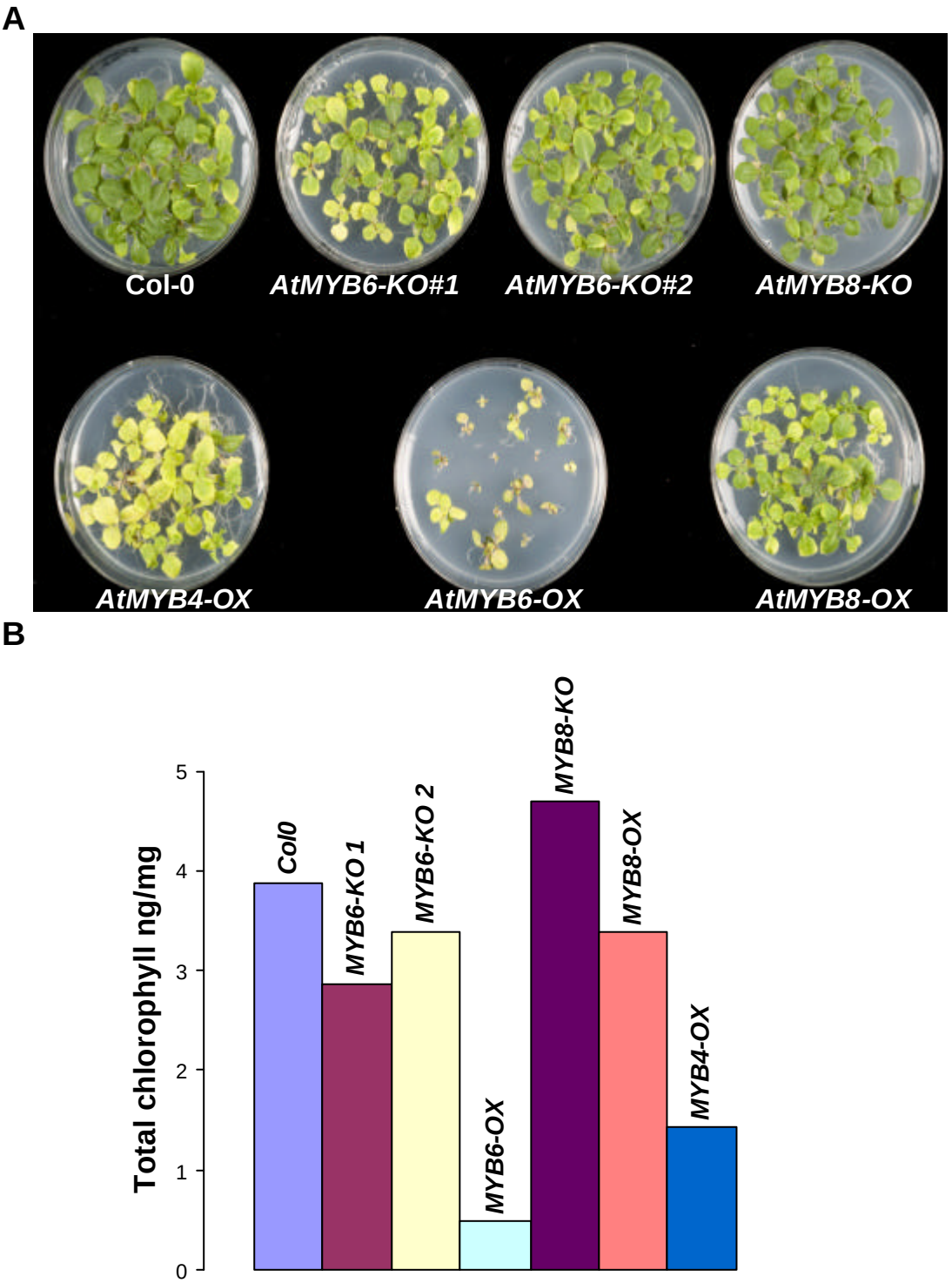


Figure 4.3. UV-B treatment in mutant Arabidopsis seedlings
(A) Tolerance to UV-B light exposure after 1 week treatment. Col-0 and *AtMYB4-OX* line were used as controls. (B) Quantifying UV-B damage by measuring chlorophyll. Chlorophylls were extracted from UV-B treated seedlings from the picture above. Error bars not present since measurement was carried out only once

4.3.2. *AtMYB6* and *AtMYB8* are involved in the phenylpropanoid metabolism

To further evaluate the putative role of *AtMYB6* and *AtMYB8* genes on phenylpropanoid metabolism, the soluble phenolics of whole Arabidopsis bolting plantlets (including roots) were analyzed by liquid chromatography-mass spectrometry (LC-MS).

Replicates of the experiment were grown on soil and on MS medium, although clearer differences were observed in plants grown on plates. Phenolic levels in transgenic plants (mutants and overexpression lines) grown *in vitro* were compared to those in wild type plants.

Individual phenolics were identified by examination of their fragmentation patterns, and comparison, where possible, to fragmentation patterns reported in the literature. In the absence of NMR data, these identities should be considered as probable rather than proven. Chromatographic and spectroscopic methods, and comparison with published data allowed the identification of significant peaks of phenolic compounds in all the samples, three kaempferol derivatives (kaempferol 3-*O*-[(rhamnosyl) glucoside]-7-*O*-rhamnoside; kaempferol 3-*O*-glucoside-7-*O*-rhamnoside and kaempferol 3-*O*-rhamnoside-7-*O*-rhamnoside) and sinapoyl malate. Many of these compounds are positional isomers or homologues, which may be distinguished on the basis of different light absorption spectra in the UV–Vis range.

The overlay of UV chromatograms at 214 nm of wild type (Col-0) and of *AtMYB6* mutant alleles (*AtMYB6-KO# 1, 2, 3)* showed an increase in kaempferol derivatives compounds surprisingly, in a very similar way to that seen in the overexpression line which is contradictory (Fig. 4.4A). Overexpression of the *AtMYB6* gene, raised levels of flavonols in the three major kaempferol derivatives identified, particularly those with retention times (RT) of 18.28, and 23.76 A1 and A4 respectively and sinapoyl malate (A2) with an RT of 20.47 (Fig. 4.4C).

The *AtMYB8* knockout line profile showed slight differences in phenylpropanoid levels compared to controls (Fig. 4.4B). There were higher

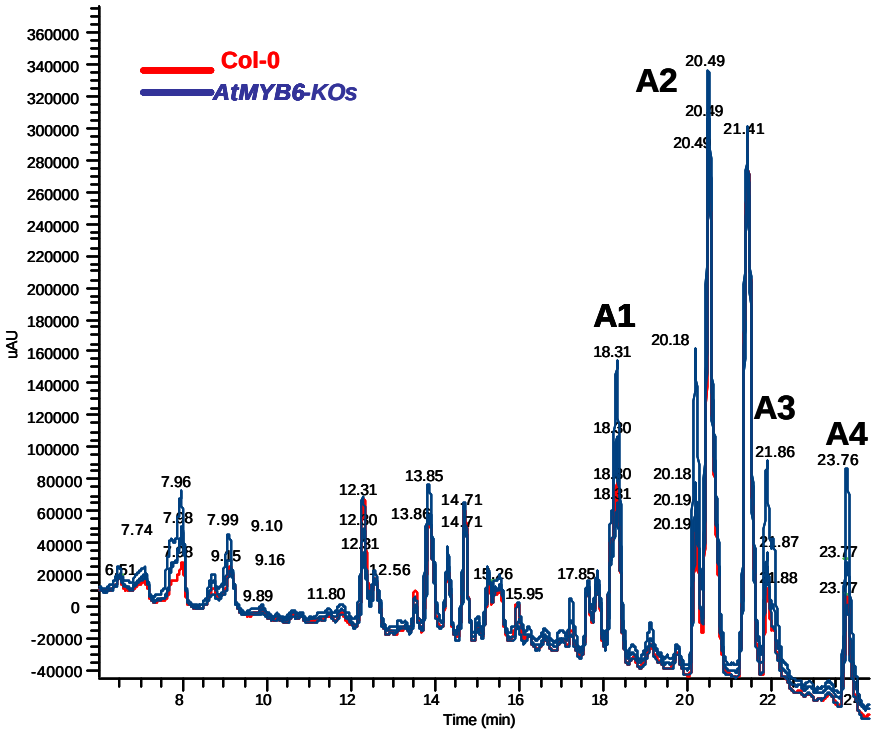
levels of sinapoyl malate (A2) in the *AtMYB8-KO#1* line compared to the Col-0 control, and higher levels of an unidentified compound at RT 12.31. Otherwise the levels of flavonol derivatives appeared similar in the *AtMYB8-KO#1* line to the control (Col-0). An opposite effect was produced by *AtMYB8* overexpression line where the levels of phenolics (particularly sinapoyl malate, A2) decreased compared to those in wild-type plants (Fig. 4.4D).

To determine whether AtMYB6 and AtMYB8 could work in the same way in other species, we expressed the *AtMYB6* and *AtMYB8* cDNAs in tobacco (*Nicotiana tabacum* var Samsun NN) under the control of the strong constitutive CaMV 35S promoter. Of more than 15 independent primary transformants, three lines (lines 6.2, 6.7 and 6.10) for *AtMYB6* and two lines for *AtMYB8* (lines 8.3 and 8.1) were investigated further.

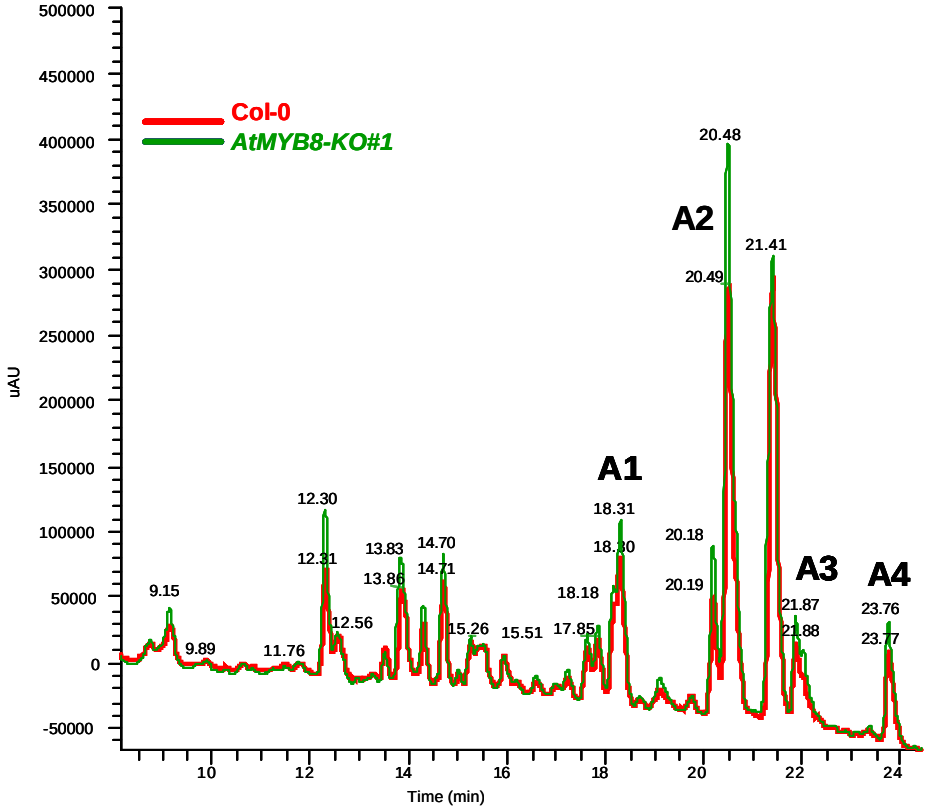
Figure 4.4. LC-MS analysis of extracts from *AtMYB6* and *AtMYB8* mutant plants

(A, B) LC-MS profile of *AtMYB6-KO* alleles (blue lines) and *AtMYB8-KO#1* (green line) compared to the wild-type Col-0 (red line). Phenylpropanoid related compounds are kaempferol-rhamnosyl-glucoside-rhamnoside (A1); sinapoyl malate (A2); kaempferol-rhamnosyl-glucoside (A3); kaempferol-rhamnoside-rhamnoside (A4). (C, D) LC-MS profile of *AtMYB6* and *AtMYB8* overexpression lines (same colour code as above mentioned).

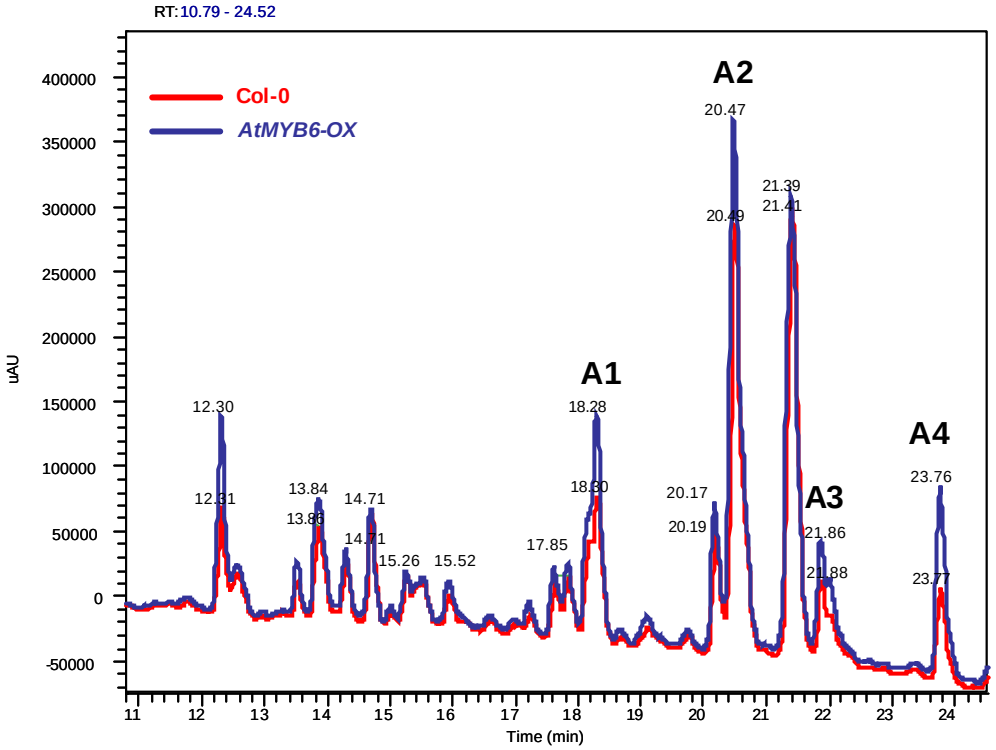
A


Col-0
AtMYB6-KOs
A1
A2
A3
A4
uAU
Time (min)

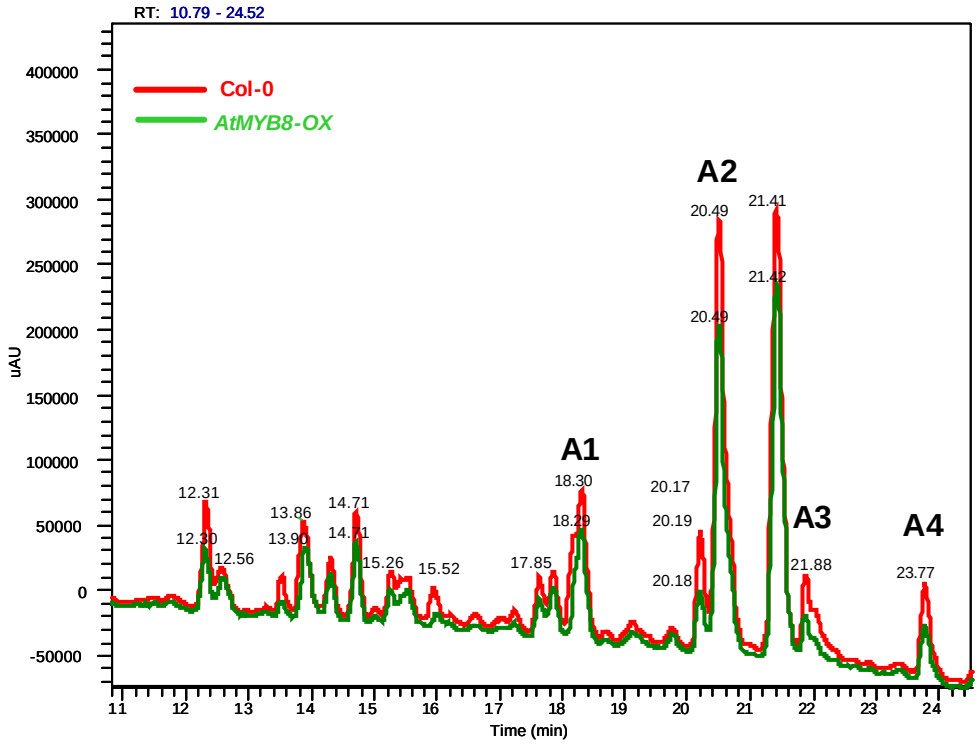
B


Col-0
AtMYB8-KO#1
A1
A2
A3
A4
uAU
Time (min)

C


RT: 10.79 - 24.52
Col-0
AtMYB6-OX
A1
A2
A3
A4
uAU
Time (min)

D

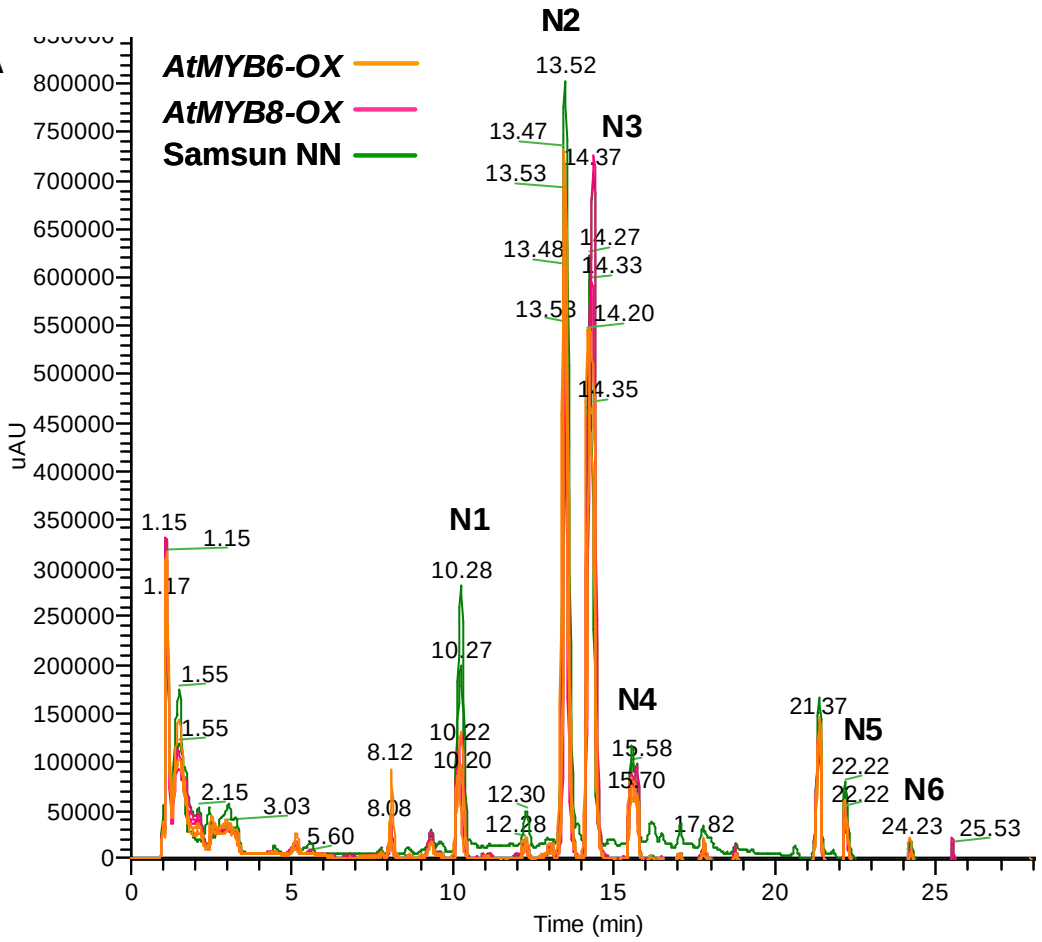

RT: 10.79 - 24.52
Col-0
AtMYB8-OX
A1
A2
A3
A4
uAU
Time (min)

The tobacco transgenic lines did not show visible phenotypic differences when compared to the control (*N.tabacum* Samsun NN transformed with an empty vector). Phenolic compounds in the leaves of tobacco transgenic plants were analysed by LC-MS and compared to those in leaves of wild-type (transformed with empty vectors). The phenolic composition of tobacco is not the same as that of Arabidopsis, so many of the compounds have different elution times.

Comparison of UV data showed that chlorogenic acid (CGA) was by far the most abundant phenolic in tobacco leaves. This was established by comparing the retention time and the UV spectra of the peaks of the extracts with the genuine standards and further confirmed by analyzing their fragmentation patterns using LC-MS (data not shown). Four peaks with masses appropriate for CGA were identified (N1-N4 in Fig. 4.5A; different isomers) although CGA1 (N1) and its isomer CGA2 (N2), with retention times (RT) of 10.28 and 13.52 min, respectively, had notably higher levels in control leaves compared to leaves of tobacco lines expressing either *AtMYB6* or *AtMYB8*. In addition, low levels of quercetin glycosides and kaempferol glycosides were also detected but no significant differences between the control and the transgenic lines were found. Quantification of the major peaks revealed that expression of *AtMYB6* and *AtMYB8* genes reduced approximately twofold the levels of the main CGA isomers detected in the transgenic lines (Fig. 4.5B).

To identify possible targets of AtMYB6 and AtMYB8 operating in the phenylpropanoid pathway, transcript levels of certain genes from the metabolic pathway (*PAL*, *C4H*, *C3H*, *4CL3* and *CHS*) were measured by real-time RT-PCR in the mutants of Arabidopsis (Fig. 4.6). Results did not show great changes in the expression levels of the genes measured in the different mutants. Only the expression levels of *C4H* (encoding cinnamate 4-hydroxylase) were statistically significantly altered in the knockout mutants of *AtMYB6* and *AtMYB8*, where the *C4H* was found to be slightly down-regulated.

A



B

Compound	*AtMYB6*	*AtMYB8*
N1	2.1	1.8
N2	1.2	1.9
N3	1.1	1.3
N4	1.3	1.2
N5	1.5	1.5
N6	1.4	1.4

Figure 4.5. LC-MS analysis of leaf extracts from *N. tabacum* expressing *AtMYB6* and *AtMYB8* genes

N1-N4, chlorogenic acid isomers (CGA); N5, quecertin and N6, kaempferol, both with rhamnose and glucose attached. (A) UV chromatogram (200-600 nm) showing the major peaks. Controls are indicated by the green lines, transgenic plants expressing *AtMYB6* in orange and in pink, *AtMYB8* transformants. (B) Phenolic compounds indicating fold difference to wild-type tobacco.

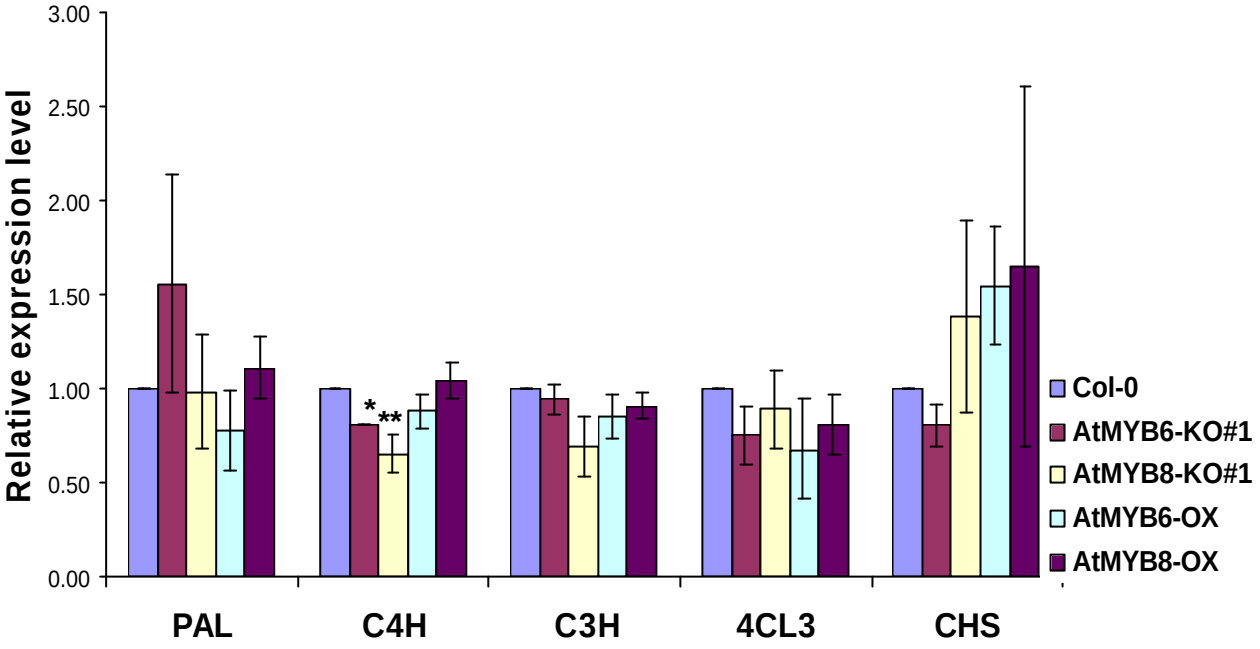


Figure 4.6. qPCR expression analysis of certain phenylpropanoid pathway genes in AtMYB6 and AtMYB8 KO and OX mutants

Relative transcript levels in control and AtMYB6 and AtMYB8 (KO and OX) mutants. Expression of the following genes of the phenylpropanoid pathway was measured; PAL, phenylalanine ammonia-lyase, C4H, cinnamate 4-hydroxylase; C3H, 4-coumarate 3-hydroxylase; ; 4CL3, 4-coumarate CoA ligase and CHS, chalcone synthase. Data represents the mean ± SE of three biological replicates. Significance of differences between control and mutants was determined using the two-sample *t*-test, as indicated by asterisks (*P < 0.01; **P < 0.05).

Curiously, *C4H* is the main target of AtMYB4 (from the same subfamily IV as AtMYB6 and AtMYB8) and this act as a negative regulator of hydroxynammic acid metabolism

4.3.3. Protein chip assay

To narrow the search for target genes of the AtMYB6 and AtMYB8 transcription factors (TFs) I undertook experiments to identify their preferred binding site motifs using a protein-oligonucleotide chip assay. Transcription factors interact with specific DNA regulatory sequences to control gene expression. Using oligonucleotide microarrays it is possible to determine the binding specificities of any TF from any organism.

The compact, synthetic DNA sequence design for the protein binding microarray (PBM) experiments allows the representation of all possible DNA sequence variants of 11 bp (4.194.304) just in 240.000 probes on a single, universal microarray. The design is possible because in each probe (35 mer) there are 25 different compact sequences (11 bp) that cover all binding sites by converting high-density single-stranded oligonucleotide arrays to double-stranded DNA arrays (Berger *et al*., 2006) (Fig. 4.7).

AtMYB6 and AtMYB8 proteins were fused to the maltose binding protein (MBP) using pDEST-TH1 as destination vector and expressed in *E. coli* strain BL-21 (DE3) Codon Plus. The predicted sizes of the fusion proteins were 77.1 kDa (MYB6) and 74.4 kDa (MYB8). Total protein extracts were detected by SDS-PAGE followed by staining with Coomassie blue. The results are shown in Figure 4.8A and 4.8B. The protein sizes deduced from the SDS gel were in the range of the predicted values. A Western blot analysis using an Anti-MBP antibody confirmed the presence of the fusion proteins in the extracts. Although the signal was low due to the quality of the picture, I can ensure the presence of both proteins their induction which was clear in the original Western membrane. Since I was interested in the crude extract, to find an important amount of protein in the soluble fraction as well in the insoluble did not matter.

Crude extracts of AtMYB6 and AtMYB8 fusion proteins were bound to the oligonucleotides on Agilent DNA chips and bound proteins were identified with an anti-MBP antibody to detect MBP-tagged proteins on the chip. All bound proteins revealed signals (yellow) over background indicating that they possess a MBP-tag when treated with the secondary antibody, anti-rabbit-IgG-Cy3-conjugate (Fig. 4.8C). In the microarray image, those spots with more intense signal were those containing the binding motif sequence. That sequence was deduced through statistical analysis (Berger and Bulyk, 2009). The result was represented as a logo with the web-based tool “enologos” (http://biodev.hgen.pitt.edu/cgi-bin/enologos/enologos.cgi). A consensus sequence only for AtMYB6 was obtained because the results from the AtMYB8 chip were not conclusive due to the high background produced in the chip which did not allow us to recover reliable results for bound protein. The DNA sequence recognized by the MYB6 transcription factor, ACCTACC, and its variants are showed in (Fig. 4.8D). That result is consistent with the data published by Feng and Parish (1995) who identified the AtMYB6 target sequence by mobility retardation assays and demonstrated that MYB6 binding motif resembles more the binding site of the maize *P* gene product CCTACCA than the animal c-Myb DNA-binding site (James et al.), CTAACTG.

Simple pattern matching searches (TAIR Patmatch web) for the consensus binding sequence reported for the *AtMYB6* protein (ACCTACC) upstream of annotated *A. thaliana* genes, allowed me to identify more than 1000 genes with the motif within 1000 base pairs of the translation start site as possible *AtMYB6* targets. From the 141 loci identified that are involved in phenylpropanoid metabolism, only eight contained the ACCTACC motif within 1000 base pairs (Fig. 4.8E).

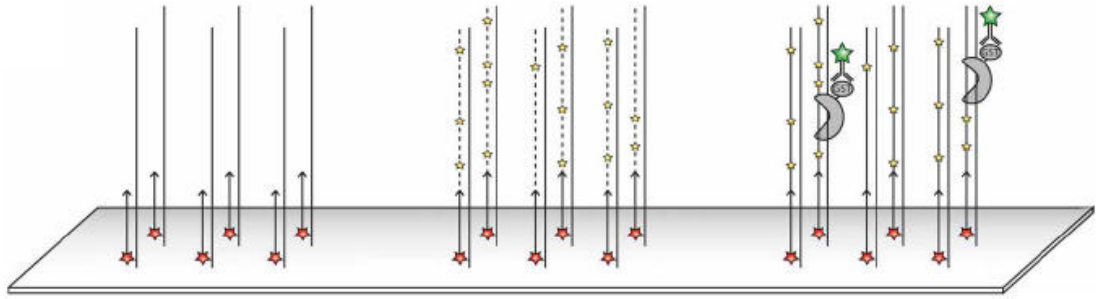
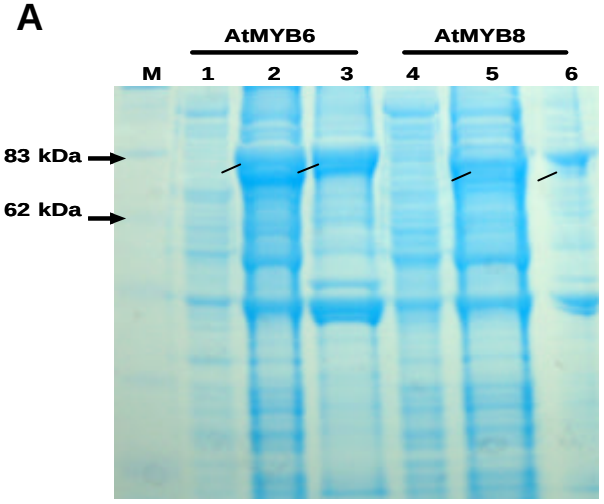


Figure 4.7. Universal protein binding microarray containing all possible 11-mer binding sites. The picture (from Michael F Berger et al. 2006) shows the main stages of the experiment. First, the double chain is synthesised from a binding sequence to a primer present in all the probes. The primer extension is done using unlabeled dNTPs and a small quantity of fluorescently labeled dUTP-Cy3 to provide a measure of the relative amount of dsDNA at each spot for subsequent data normalization. Next, the protein of interest is fused to an MBP tag is added and then detected by an anti-MBP antibody.

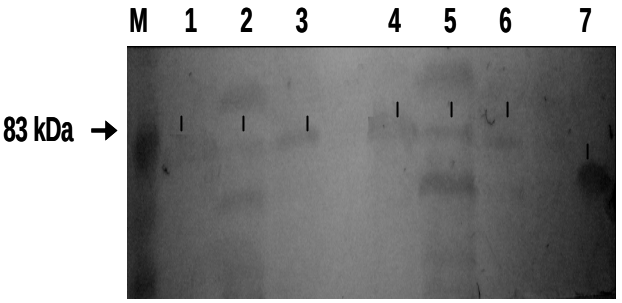
Figure 4.8. *AtMYB8* and *AtMYB6* protein expression using MBP tag

(A) Coomassie blue staining of SDS-PAGE gel showing AtMYB6-MBP (77.1 kDa) and AtMYB8-MBP (74.4 kDa) fusion proteins in crude extracts of *E. coli*. Arrows indicate proteins of expected sizes. Uninduced cell extract (lanes 1 and 4), induced soluble extract (lanes 2 and 5) and induced insoluble proteins from extracts (lanes 3 and 6). (B) Western blot analysis with the same samples shown in the stained gel (A; sample order is the same). Membrane probed with Anti-MBP antibody directed against the MBP tag. Lane 7 is purified MBP protein (42.48 kDa) used as control for the Anti-MBP. (C) Array containing the probes with the binding sites of AtMYB6 (yellow spots). (D) Consensus sequence recognized by AtMYB6 after the bioinformatic analysis. Result represented using the "ENOLOGOS" server. The size of the individual symbols within each column reflects the frequency of the corresponding nucleotide at this position. (E) Table listing genes encoding enzymes from the phenylpropanoid pathway containing the AtMYB6 binding motif ACCTACC in their promoters.

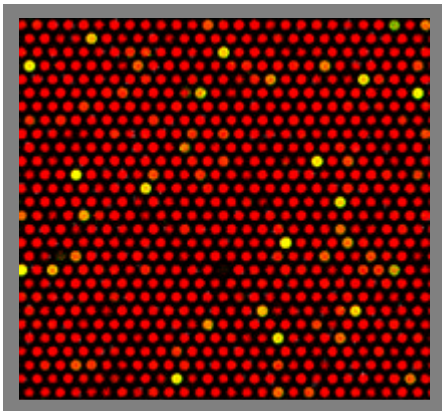
A



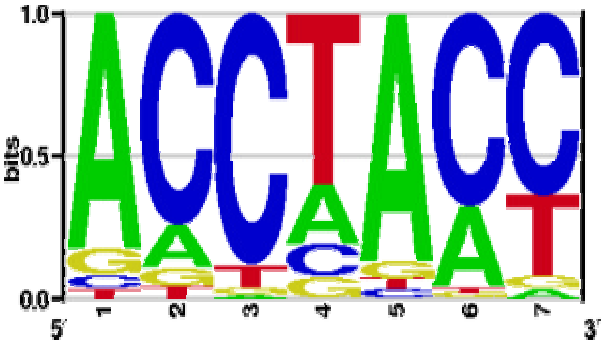
B



C



D



E

LOCUS	ENZYME
AT1G65060	4-CoumarateCoA ligase (4CL3)
AT1G30530	UDP-glycosyltransferase (UGT78D1)
AT1G06000	UDP-rhamnose:flavonol-3-O-rhamnosyltransferase
AT2G37040	Phenylalanine ammonia-lyase activity (PAL1)
AT2G30490	Cinnamate-4-hydroxylase (C4H)
AT3G55120	Chalcone flavone isomerise (CHI)
AT4G36220	Ferulate 5-hydroxylase (F5H)
AT5G13930	Chalcone synthase

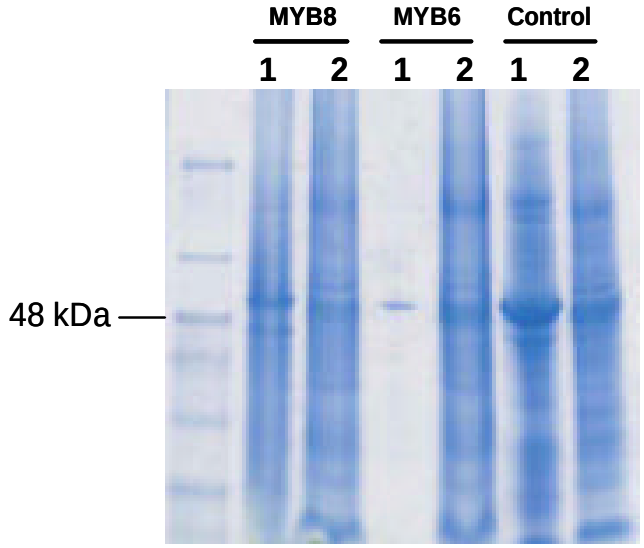
4.3.4. Protein-protein interactions: Inmunoprecipitation

Transcription factors (TFs) can modulate gene expression by binding to motifs in the promoters of their target genes. This may, or may not involve interactions with other TFs (Walhout, 2006). Inmunoprecipitation experiments to test interactions between AtMYB6 or AtMYB8 and other proteins could provide evidence for their participation in protein complexes for transcriptional regulation. If such interaction studies proved positive, finding some of the proteins associated with these MYB transcription factors could shed light on their possible biological functions.

Nicotiana benthamiana leaves were agroinfiltrated with the proteins AtMYB6 and AtMYB8 fused to the green fluorescent protein (GFP) (the same constructs used for the subcellular localization of these proteins, section 3.2.1.1). Pull-downs from agroinfiltrated plant extracts were analyzed looking for heteromeric binding interactions. Protein crude extract and proteins fused to the GFP-GST beads were separated by SDS-PAGE, and after blotting on a membrane, the protein complexes were detected with Anti-GFP (B-2) sc-9996 (Santa Cruz Biotechnology, inc.), a monoclonal antibody raised in mouse against GFP. In the gel stained only with Coomassie blue, no Anti-GFP immunoprecipitates could be detected. In fact, only bands corresponding to the complex GST-GFPbeads-MYB protein (~ 48 KDa) were clearly visible; the other bands that were present in the transgenic lines were also present in the control (Fig. 4.9A). The western-blot result did not show any differences either, and no trace of a differential protein complex was detected. All the bands recognized by the Anti-GFP antibody in the MYB6-GFP or MYB8-GFP samples were also present in the control line extracts (Fig. 4.9B).

The experiment was repeated twice and similar results were obtained. However, this does not mean that AtMYB6 and AtMYB8 transcription factors do not bind other proteins but rather that IP may not allow detection of some protein–protein interactions that are either too weak or too transient to be detected by traditional co-immunoprecipitation methods.

A



B

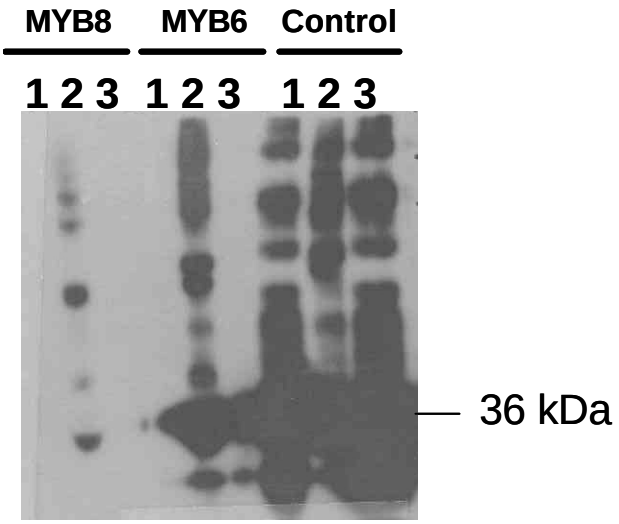


Figure 4.9. AtMYB6 and AtMYB8 immunoprecipitation

(A) Protein extracts from *N. benthamiana* expressing transiently AtMYB6-GFP, AtMYB8-GFP or GFP alone were run on a SDS-PAGE gel and stained with Coomassie (lane 1, eluates; lane 2, protein crude extract). The presence of protein complexes were detected by protein gel blot analysis using antibodies against the GFP (B) (lane 1, eluates; lane 2, protein crude extract; lane 3, unbound fraction).

4.4. Discussion

Analysis of *A. thaliana* mutant lines for the *AtMYB6* and *AtMYB8* genes suggests that these MYB transcription factors modify phenylpropanoid levels of this complex pathway. I have confirmed that they have a likely role in the regulation of phenolic production against UV-light radiation and I hypothesize that these changes in phenolic content may be the cause of some developmental phenotypes observed in the different Arabidopsis *AtMYB6* and *AtMYB8* transgenic lines reported in the previous chapter.

MYB proteins have a major role in the regulation of secondary metabolism such the phenylpropanoid pathway. Metabolites from this pathway function as structural components of the cell walls, as pigments, as signalling molecules, in defence against hervibores or microbes, wounding or as UV light protectants, a role developed by the flavonoids and sinapate esters, compounds from the branches of the general phenylpropanoid pathway (Dixon and Paiva, 1995).

A correlation between flavonoid content and sunscreen protection has been already cited (Li *et al*., 1993; Landry *et al*., 1995). High UV-B directly damages DNA, membranes and proteins altering the normal development of the plant and triggering the biosynthesis of UV-absorptive secondary metabolites amongst other signaling pathways (Jansen *et al*., 1998). Several Arabidopsis mutants related to protection against UV radiation, have been reported as defective in the general phenylpropanoid pathway (Arvind and Jitendra, 1997).

AtMYB6 and *AtMYB8* seem to be involved in the production of sunscreen protectants in response to UV-light and several cis-regulatory light inducible elements that I found in their promoters support this hypothesis. Curiously, another MYB gene (*AtMYB4*) from the same subfamily group as *AtMYB6* and *AtMYB8*, participates in the transcriptional regulation of the synthesis of the UV-protectant compound sinapoyl malate (Jin *et al*., 2000; Hemm *et al*., 2001) highlighting once more the functional redundancy in the plant MYB family (Jin and Martin, 1999). Knockout and overexpressed mutant

lines from the *AtMYB6* and *AtMYB8* genes showed different sensitivity responses to UV-light irradiation that indicated likely alterations in the phenylpropanoid profiles of these plants.

The relationship between phenylpropanoids and relative UV-B tolerance was very clear for *AtMYB8* mutant and overexpression lines. The *AtMYB8-OX* line had a decreased content in flavonoids and sinapoyl malate (as previously demonstrated by the NMR results) that made it more vulnerable to the UV-B radiation received. Consistenly the *AtMYB8-KO* showed a higher resistance against the UV-B light probably as a result of its higher level of sinapoyl malate. which is consistent with the assesment that sinapate esters are more effective UV-B protectants than flavonoids (Landry *et al*., 1995). However, results for *AtMYB6* mutants and overexpression lines were contradictory. Contrary to the NMR analysis, the content of flavonoid derivatives and sinapate esters in the *AtMYB6* overexpression line were higher than normal but this fact did not increase the tolerance of the line to UV-B as shown by the pictures. This increased UV-B sensitivity, together with the preliminary results from NMR analysis of the *AtMYB6-OX* line, suggests that sinapoyl malate and flavonoid content should be lower contrary to what was observed by LC-MS analysis. The most likely interpretation of these data is that the increase in phenylpropanoids was an aberration of the sample since replicates from similar experiments did not show clear differences in phenolic contents, which could have due to redundancy of AtMYB6 function with that of AtMYB8. In addition, the *AtMYB6-KO* lines, #1 and #2 (which is expressed partially), were slightly hypersensitive to UV-B treatment (especially *AtMYB6-KO* #1) although the levels of phenylpropanoids in leaves of these lines were very similar to those in the wild-type. However, UV-B tolerance was measured in seedlings much younger than those assayed for their metabolite profiles. In future experiments metabolites in younger plants should be assayed to have more consistent results and correlate the phenylpropanoid metabolite profile with the UV-B tolerance.

I think it most likely that AtMYB6 is a negative regulator of genes in general phenylpropanoid metabolism. Certainly, this is the also the conclusion from Dr. Lebouteiller's NMR analysis. My failure to reproduce some data by LC/MS could be due to variations in several conditions such as age of the plants, developmental stages sampled, *etc*.

On the other hand, the results from tobacco overexpressing *AtMYB6* and *AtMYB8*, where decreases in the levels of phenolics in leaves were observed compared to controls reinforce the idea that both genes are negative regulators of the general phenylpropanoid pathway like the *AtMYB4* gene from the same subfamily. AtMYB4 represses the expression of the gene encoding the enzyme that is a primary limiting step in the synthesis of sinapate ester sunscreens, cinnamate 4-hydroxylase (C4H). Mutants of *AtMYB4* show enhanced levels of sinapoyl esters in their leaves and, therefore, are more tolerant of UV-B irradiation than wild type plants (Jin *et al*., 2000) in a very similar way to what I observed for *AtMYB6* and *AtMYB8*. Together, these results indicate a general role for members of subfamily IV in controlling phenylpropanoid metabolism and again show the degree of functional redundancy between similar members of this family of transcription factors.

To determine whether *AtMYB6* and *AtMYB8* genes regulated certain enzymes of the general phenylpropanoid pathway a qPCR analysis was carried out. PAL, C4H, C3H, 4CL3 and CHS expression levels were measured on the KO and OX mutants for AtMYB6 and AtMYB8 but relevant changes on the expression of these putative target enzymes were not found. Probably, more biological replicates should be necessary to find statistically and biologically significant differences (plants sampled for qPCR were not at the stages tested for biological responses).

To date, there is only one recent reference in the literature that relates one of the proteins of this study, *AtMYB6*, to the phenylpropanoid metabolism supporting the idea that it plays a role as a transcriptional repressor of the phenylpropanoid biosynthetic pathway (Rowan *et al*., 2009).

It is known that most plant MYB proteins recognize the cognate site TAACTAAC (MBSII) whereas others have affinity for other target sites recognition, T/CAACG/TGA/C/TA/C/T (MBSI), common for mammalian cells invertebrates and cellular slime moulds. Some MYBs in plants, have flexibility to bind both MBS sites (Romero *et al*., 1998). In an attempt to identify the possible targets for AtMYB6 and AtMYB8 in the phenylpropanoid pathway, I was able to identify the binding DNA sequence only for the AtMYB6 transcription factor. My data point to the seven base pairs motif, ACCTACC, as the consensus sequence recognized by AtMYB6 MYB domain which coincides practically with the MBSII site (TAACTAAC) and with the sequence deduced and already published by Li and Parish in 1995. The consensus binding sequence reported for the AtMYB6 protein was identified in more than 1000 genes with the motif within 1000 base pairs of the translation start but only eight genes identified for the phenylpropanoid metabolism contained the motif. These eight genes encode enzymes involved in general phenylpropanoid metabolism, which provides precursors for the different branches of phenylpropanoid metabolism. Although data for AtMYB8 could not be retrieved, it is likely that it has the same binding domain as AtMYB6 because AtMYB8 is very similar to AtMYB6 in its DNA binding domain.

Phenotypes of the *AtMYB6* and *AtMYB8* mutants reported in the previous chapter may be explained through some of the results described in this chapter. Coordinating flavonoid regulation is important for normal plant growth and development (Broun, 2005; Ringli *et al*., 2008). Aberrations in development due to a malfunction of phenylpropanoid metabolism have been already reported although few architectural phenotypes of flavonoid pathway mutants are known. Stunted growth, reduced pollen viability or altered flower morphology and pigmentation are some of the effects produced by the overexpression of the phenylalanine ammonia-lyase (PAL) gene in tobacco plants (Elkind *et al*., 1990); in petunia and maize chalcone synthase mutants are defective in pollen tube growth (Mo *et al*., 1992) and shoot and root morphological phenotypes for the *transparent testa* (*tt*) mutants (altered in

enzymatic steps affecting flavonoid synthesis) (Buer and Djordjevic, 2009) are some examples associating development and phenylpropanoids. Some of these morphological phenotypes for the *tt* mutants like increased lateral root density (*tt5*, *tt6*, and *tt10*), long hypocotyl phenotype (*tt6*), differential effects in root hair density or the distance from the root tip to the root hair initiation zone or the increased inflorescence production (*tt8, tt10,* and *ttg*) are similar features to those that I found in AtMYB6 and AtMYB8 mutants.

Flavonoids have been found as natural inhibitors of auxin transport and are localized to the tissues that transport this hormone (Murphy *et al*., 2000; Brown *et al*., 2001; Peer and Murphy, 2007). Auxin (IAA), is a plant hormone implicated in diverse aspects of plant growth and development and numerous mutants deficient in auxin response exhibit overall dwarfism or altered organ morphology (Friml, 2003; Mockaitis and Estelle, 2008). Relationship between flavonoid levels and phytoregulators during plant development has been established through reports of plant phenotypes caused by the flavonoid effects on auxin (IAA) movement. I hypothesize that the alterations in flavonoid content reported in this chapter could cause some of the phenotypes reported in the previous chapter. This idea is consistent with the reduced size and the observed defects in roots and leaves of *AtMYB6-OX* seedlings grown on plates (Chapter 3 Figs. 3.9, 3.10 and 3.11). The cell walls of *AtMYB6-OX* seedlings also displayed accumulation of lignin (a phenylpropanoid-derived compound) indicating that increased levels of phenylpropanoids could boost lignin production. In several studies, high levels of lignin are associated with a reduction in growth suggesting that the process of cell elongation may be linked to the regulation of lignification. That is the case for *eli1* mutants, which are ectopic lignification mutants with abnormal cellular organisation and altered cell expansion (Cano-Delgado *et al*., 2000) or other mutants that have also been shown to lignify ectopically such as *kor*, *rsw1* and *det3* (Ana *et al*., 2003; Newman *et al*., 2004), impaired in various aspects of cell expansion. The *root swelling 1* (*rsw*1) mutant resembles more the *AtMYB6-OX* phenotype. The *det3* mutant provides a link between R2R3MYB activity and lignin deposition.

AtMYB61 was found to be responsible for the ectopic accumulation of lignin in the *det3* mutant indicating that *AtMYB61* may function as a factor in a spatial and temporal regulation of lignin deposition in Arabidopsis. *AtMYB58* and *AtMYB63* have also been recently found as MYB transcription factors that activate lignin biosynthetic genes during secondary wall formation (Newman *et al*., 2004; Zhou *et al*., 2009). Alternatively, other factors could cause the characteristic short and swollen primary root phenotype of *AtMYB6-OX*. Some mutants with swollen growth phenotype such as *sabre*, *radially swollen 4* or *pleiade* mutants, have an altered cell wall composition and microtubule organisation (Aeschbacher *et al*., 1995; Muller *et al*., 2002; Wiedemeier *et al*., 2002).

The reduced levels of flavonoids in the *AtMYB8-OX* line could possibly increase the efflux of auxin from cells, so promoting lateral root hair formation (Fu and Harberd, 2003; Weijers and Jürgens, 2004). This idea supports the root phenotype in the *AtMYB8-OX* line which presented increased and precocious lateral root hair initiation (Fig. 3.10). A cis-regulatory element found in the *AtMYB8* promoter adds strength to this theory. The regulatory sequence KCACGW, has been described as a conserved root hair–specific module in genes from diverse angiosperm species with different hair distribution patterns (Kim *et al*., 2006). *AtMYB6-KO* as well as *AtMYB8-KO* lines did not show as altered and short roots as the overexpression lines but in both cases the mutants displayed slightly shorter roots than the wild-type. The fact could be explained by the role of the auxins as inhibitors of root elongation except for the transgenic plants *AtMYB6-OX* and *AtMYB8-KO* which I have previously suggested had decreased levels of auxin. It is possible that these mutants may have reduced response to auxin and obtain growth stimulation at concentrations of auxin that normally would inhibit wild type root growth (Ruzicka *et al*., 2007) or alternatively that flavonols are playing a role in auxin-independent cell growth and development as suggested by Ringli. *et al*., (2008).

Strong root phenotypes of *AtMYB6*OX and *AtMYB8*OX could be indicative of possible overlapping/complementary functions of these genes, in fact, the genes have a very similar expression pattern in roots as shown in Chapter 3. Gain- and loss-of-function mutations in the Arabidopsis transgenic lines affected root phenotypes differentially, suggesting *AtMYB6* and *AtMYB8* genes as putative regulators of root system development. The role of the MYB transcription factors in control of cell morphogenesis and pattern formation is well known (Du *et al*., 2009). Some root epidermal patterns in Arabidopsis are due to MYB transcription factors such as *WEREWOLF* and *CAPRICE* (*CPC*), involved in the mechanism of cell fate determination of root-hair cells or *MYB59*, which regulates cell cycle progression and root growth (Mu *et al*., 2009).

Alternatively, AtMYB6 and AtMYB8 transcription factors could be involved in germination. Germination is a process controlled by hormone interactions and environmental factors (Kucera *et al*., 2005). Structure/pigmentation of the seed coat (testa) has been reported as an important determinants of radicle emergence (Debeaujon *et al*., 2000; Koornneef *et al*., 2002). Flavonoid pigments confer seed coat colour and are positively correlated with the permeability of the seed coat and chemical-mechanical protection of the seed. The role of flavonoids in germination has been demonstrated in the *transparent testa* (*tt*) and *banyuls* (Bang et al.) mutants whose lack of pigmentation in the seed coat appears to be related to high germination rates. Although *AtMYB6* and *AtMYB8* mutants (KO and OX lines) showed differences in their flavonoid metabolic profiles in vegetative tissues, their testa colour was not altered. The alternative hypotheses that link the content of pigments in seed coat to reduced germination (in *ban* mutants), Albert *et al*., (1997) or the suggestion that the lack of pigments reduces dormancy (Debeaujon *et al*., 2000) are not in agreement with my results from the germination tests because all the *AtMYB6* and *AtMYB8* mutants (except the *AtMYB6-OX*), germinated earlier than the wild-type, independently of their differential contents of phenolics.

The findings presented here, suggest that AtMYB6 and AtMYB8 proteins are negative regulators of the complex phenylpropanoid pathway and when the expression of these genes is altered it affects the plant phenotype indirectly, presumably as a result of changing the regulation of transcription of downstream genes.

CHAPTER 5

Chapter 5: Functional characterisation of two MYB genes in the moss *Physcomitrella patens*

5.1. Introduction

Arabidopsis thaliana has become the model system of choice for research in plant biology over the past 25 years but interest in the moss, *Physcomitrella patens,* started to grow in the late nineties although it remains unknown to many plant researchers as an experimental organism (Fig. 5.1A) (Meinke *et al*., 1998; Schaefer and Zryd, 2001).

Mosses and flowering plants diverged more than 400 million years ago and, although their morphologies and life cycles are very different, they shared a common ancestor. In fact, sequence comparisons of *A. thaliana* and *P. patens* show that at least 66% of *Arabidopsis* genes have homologues in Physcomitrella (Nishiyama *et al*., 2003). The division Bryophyta which comprises the mosses, hornworts and liverworts has been considered the simplest and earliest diverging group of land plants having originated in the early Silurian (430 million years ago). This places them in an ideal evolutionary position for studying the mechanisms that allowed plants to conquer the land (Fig. 5.1B) (Rensing *et al*., 2008). Genes which encode proteins with similar sequences that are present in both flowering plants (Angiosperms) and bryophytes are likely to have shared a common ancestral gene 400 Mya. Conservation of function in different lineages over 400 Mya implies that such genes are likely to have been important to the survival of plants on land.

The genome of *Physcomitrella patens* was the first bryophyte genome sequenced and *P.patens* is the first land plant to be successfully transformed by homologous recombination (Fig. 5.2). The high gene targeting efficiency in Physcomitrella is comparable to that observed in yeast (*S. cerevisiae)* making this bryophyte the only plant able to integrate foreign DNA at high efficiency to allow the generation of targeted knockouts with high precision. In addition, *P. patens* has a well-studied life cycle and a rapid generation time which makes it

A

CHARACTERISTICS	*ARABIDOPSIS THALIANA*	*PHYSCOMITRELLA PATENS*
Family	Brassicaceae	Bryophytaea
Dominant generation	Sporophyte (diploid)	Gametophyte (haploid)
Genome sequence (year)	2000	2005
Genome size	~125 Mb (5 chromosomes).	~ 511 Mb (27 chromosomes).
Gene number	~ 25,500 genes.	~25,500 putative transcripts.
MYB gene number	~126 R2R3-MYB	~ 62 R2R3-MYB
Transformation	Floral dipping	(PEG)-mediated direct gene transfer to protoplasts
Some advantages as a model system	-easily manipulated, - genetically tractable, small genome size.	- interesting phylogenetic position - efficient homologous recombination - dominant haploid phase - small size, ease of cultivation -rapid generation time

B

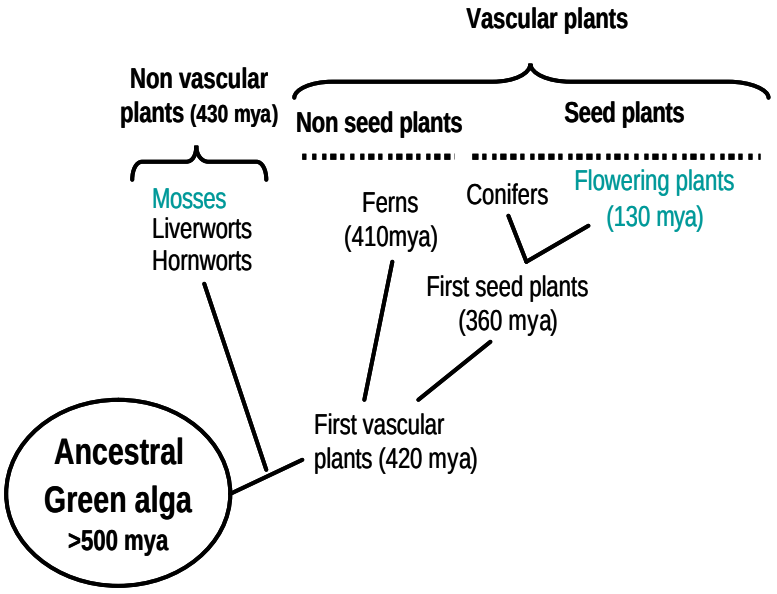


Figure 5.1. (A) Main characteristics between *A.thaliana* and *P.patens* as plant model systems. (B) Diagram of the relationships between basal groups of plants and their origin indicated in million of years (mya).

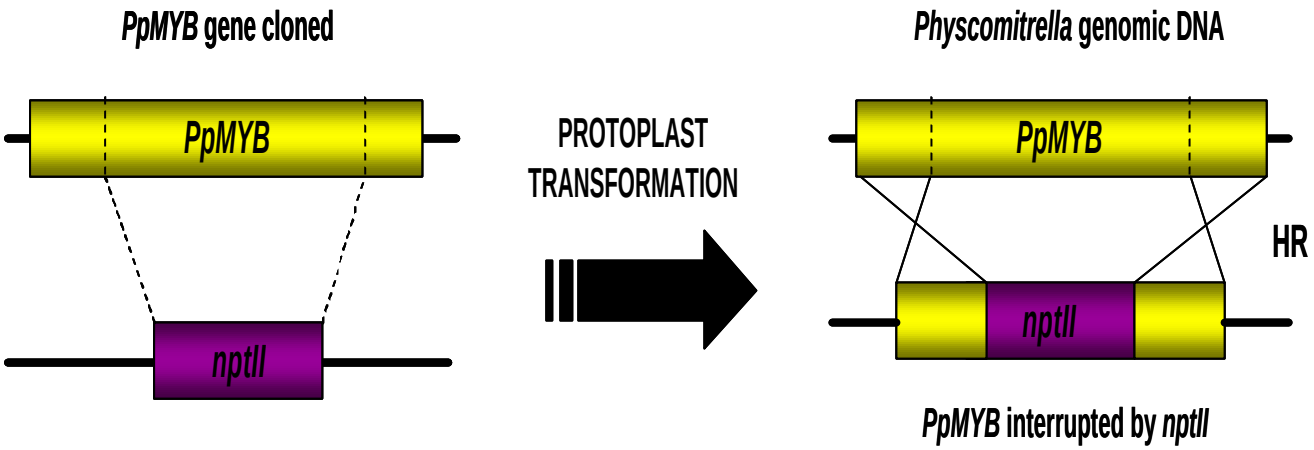


Figure 5.2. Homologous recombination to generate knockout constructs of MYB genes in *P.patens*

Gene disruption constructs for the generation of targeted knockout plants can be obtained by the insertion of a selection marker cassette carrying the *nptII* gene into genomic DNA of the gene of interest. The transgene can be integrated into a precise locus therefore the transgene has to be flanked with similar sequences to that of the insertion site desired.

attractive for genetic studies. It has a dominant gametophyte generation, which is haploid, and produces gametes in the gametangia (called archegonia and antheridia) which fuse to form the sporophyte, which is reduced, diploid and produces haploid spores through meiosis. The haploid spores germinate to form the gametophytic generation which has filaments of cells (chloronemata and caulonemata) called protonemata. Within a few weeks, the protonemata produce leafy shoots (gametophores) that will produce the sporophyte (Fig. 5.3) (Schaefer and Zryd, 2001; Hohe *et al*., 2004; Cove, 2005).

Very little is known about MYB transcription factors in bryophytes, in fact, the MYB family in Physcomitrella was originally estimated to be small with fewer than 3 members (Martin and Paz-Ares, 1997), although the PlantTFDB database estimates 60 genes encoding MYB proteins. This estimate includes genes encoding MYB proteins which do not have the R2R3 MYB DNA binding domain. A recent screen of the *P. patens* genome sequence by Paul Bailey at JIC has identified 48 proteins encoded by R2R3 MYB genes which can be aligned with the Arabidopsis R2R3 MYB proteins (Fig 5.4). From these comparisons it would appear that most *R2R3 MYB* genes exist as highly similar gene pairs, suggesting a very recent duplication of the Physcomitrella genome as predicted recently (Rensing *et al*., 2008). Prior to publication of the *P. patens* genome sequence only two R2R3 MYB proteins, Pp1 and Pp2 had been isolated from the moss Physcomitrella (Leech *et al*., 1993) and characterized as possible regulators of protonemal growth.

Comparative genetic analysis between Arabidopsis and Physcomitrella is an important tool that should help to understand the evolution of plants and hence more complex systems benefiting the field of evolutionary developmental genetics or “evo-devo” as well as offering to the plant research community the opportunity to understand the early evolution of land plants.

Through the study of selected homologous MYB transcription factors in the moss *Physcomitrella patens*, which are structurally similar to the Arabidopsis proteins AtMYB6 and AtMYB8 I attempted to elucidate if these

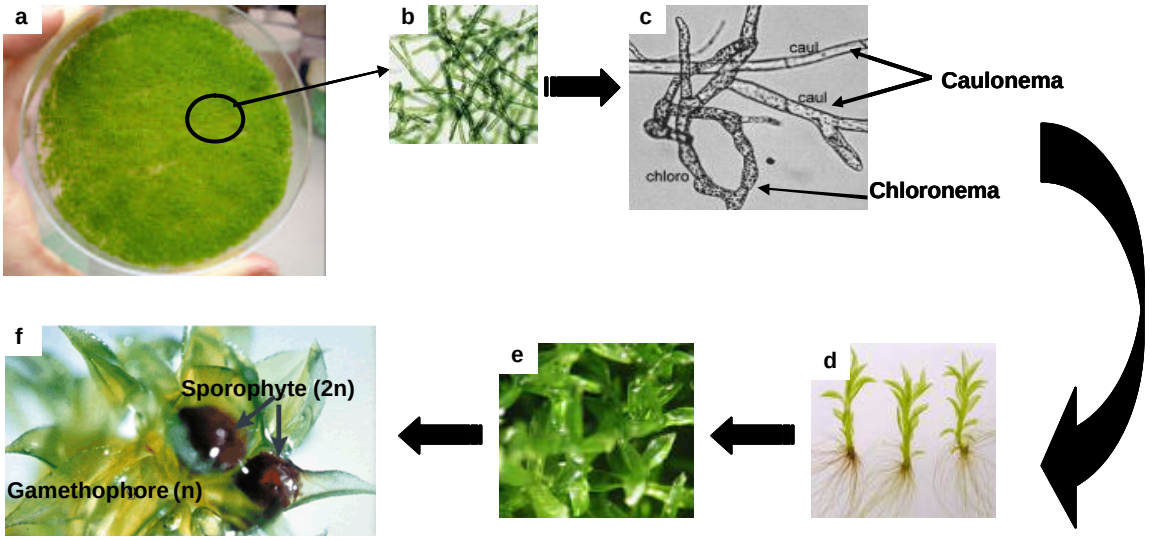


Figure 5.3. Brief *Physcomitrella patens* life cycle

(a) *Physcomitrella patens* protonemata growing on plate. (b,c) Spore germination produces a photosynthetic filament of cells called the protonema which is formed by two different types of cells the chloronema, with large and numerous chloroplasts and the caulonema, with fewer chloroplasts and faster growth. (d) Leafy buds are produced by the protonema as it matures developing the leafy shoot of the gametophore n (haploid). (e) Colony of gametophores (f) The leafy gametophores produce the gametangia (antheridia or archegonia) after cold treatment. After fertilization, it develops a diploid (2n) sporophyte containing the spores (haploid). The life cycle takes approximately 3 months under optimal conditions.

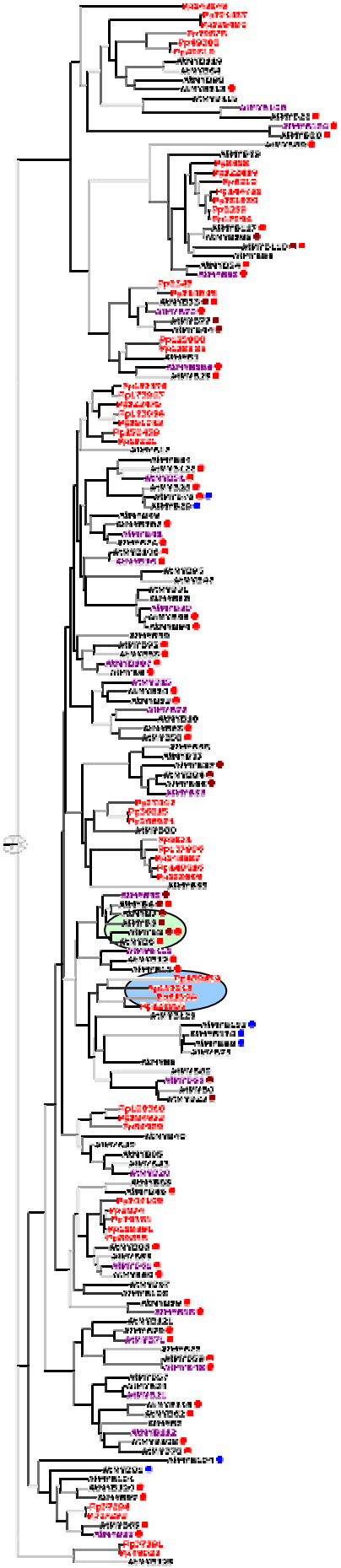


Figure 5.4. A view of the main phylogenetic tree showing the Arabidopsis R2R3 MYB family with a single query sequence added (name in red). Genes residing on duplicated regions of the genome are indicated: red or dark red for recent or old duplication events, respectively.

The Physcomitrella names of the genes in the tree correspond to the names given by the JGI sequencing project:

Pp189472 = PpMYB4

Pp126855 = PpMYB5

and PpMYB4 and PpMYB5 duplicates (not shown in the tree):

Pp61394 = XP-001783614-1

Pp113218 =XP-001752936-1

The red dots represent genes residing on regions of the genome that have been duplicated in an ancestor so may exhibit some functional redundancy. Green circle contains AtMYB6 and AtMYB8 whereas the blue circle contains PpMYB4 and PpMYB5.The blue dots correspond to tandemly duplicated genes. The moss sequences are in red because they were the query sequences added to the AtMYB tree in the web page.

Courtesy of Dr. Paul Bailey

proteins had conserved functions despite being separated by 400My of evolution, or whether they diverged functionally, to regulate different pathways or processes. The evolutionary position of *P. patens* makes it ideal for the study of the evolution of metabolic pathways, including phenylpropanoid biosynthesis. In this chapter, I describe the characterisation of the *PpMYB4* and *PpMYB5* genes as close relatives of *AtMYB6* and *AtMYB8* in Arabidopsis, and investigation of their functions as putative regulators of the phenylpropanoid pathway in lower plants.

5.2. Materials and methods

5.2.1. Plant Material and Growth Conditions

The Gransden strain of *P. patens* (Hedw.) Br. Eur.[*Aphanorhegma patens* (Hedw.) Lindb] was used as wild-type. Protonemal tissue was grown in 9-cm Petri dishes on cellophane disks (A.A. Packaging Limited, Liverpool Road, Walmer Bridge, Preston Lancashire PR4 5HY) on KNOP medium which contains 0.8 g/l $CaNO_3.\ 4H_2O$, 0.25 g/l $MgSO_4.7H_2O$, 0.0125 g/l $FeSO_4.7H_2O$, 0.25 g/l KH_2PO_4 (pH 7 adjusted with KOH), and trace element solution (0.055 mg/l $CuSO_4.5H_2O$, 0.055 mg/l $ZnSO_4.7H_2O$, 0.614mg/l H_3BO_3, 0.389 mg/l $MnCl_2.7H_2O$, 0.055 mg/l $CoCl_2.6H_2O$, 0.028 mg/l KI, 0.025 mg/l $Na_2MoO_4.2H_2O$). The medium was also supplemented with 0.5 g/l ammonium tartrate, 5 g/l glucose and solidified with 0.8 % (w/v) agar (DIFCO bacto agar) (termed Pp-NH_4 media). Cultures were grown in a controlled environment room at 23ºC±2ºC under a long photoperiod (16 h light/ 8 h dark). Tissue was propagated vegetatively by homogenisation and subculture every 5-7 days.

5.2.2. DNA and RNA isolation

For DNA plasmid isolation Qiaprep Miniprep Kits (Qiagen) were used. For a quick screening of the putative stable transformants, genomic DNA from Physcomitrella was extracted with NaOH and Tris-HCl as described for Arabidopsis in the General Material and Methods section.

To confirm the stable transformants and for the Southern blot, *P. patens* DNA was extracted by a modified CTAB method according to the protocol described by Knight *et al*. (2002). Seven day-old protonemal tissue from one plate was ground to a fine powder in liquid nitrogen using a mortar and pestle. Genomic DNA was extracted using CTAB extraction buffer and DNA was recovered by ethanol precipitation and dissolved in TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA).

For RNA extraction, moss tissue at different stages was harvested and gently squeeze-dried with Whatman paper. Total RNA was obtained using the RNeasy plant mini kit (Qiagen). RT-PCR was carried out using SuperScriptIII Reverse Transcriptase (Invitrogen).

5.2.3. Cloning *PpMYB4* and *PpMYB5* for Arabidopsis transformation

PpMYB4 and *PpMYB5* ESTs were provided by the BRC Resource Centre (BRC resource number pdp07793 and pdp12446 respectively) (Nishiyama *et al*., 2003). Sequencing of both strands confirmed that both cDNA clones were full length by comparison with the annotated genome sequences.

The *PpMYB4* and *PpMYB5* cDNAs were subcloned in the Gateway vector pDONR207 using the primer pairs SJS3III/SJS4II and SJS5II/SJS6 respectively and then subcloned into the destination vector pJAM1502 which has a Gateway destination recombination site between the double CaMV 35S promoter and the CaMV Terminator sequence in the binary vector pBin19 (for methods describing Gateway cloning see the General Materials and Methods section 1). The resulting plasmids PpMYB4cds (CaMV35S: PpMYB4: CaMV Term) and PpMYB5cds (CaMV35S: PpMYB5: CaMV Term) were used to transform Arabidopsis (Col-0) as described in the General Materials and Methods section 11.

5.2.4. Construction of the Physcomitrella targeting vectors

The genomic regions of the *PpMYB4* and *PpMYB5* genes were amplified from total cellular DNA by PCR with the Gateway primers SJS13-SJS14II and

SJS15-SJS16 respectively. The amplified DNA fragment for each gene (~3 kb) was cloned into the pDONR201 entry vector to give pDONR201:PpMYB4 and pDONR201:PpMYB5.

To create gene knockouts in *P. patens,* the *nptII* gene cassette, which consists of the CaMV 35S promoter, the *nptII* coding region, and the CaMV gene 6 terminator, was used. For *PpMYB4*, the *nptII* gene cassette from pMBL8b (*nptII* in antisense orientation) was excised as a ~ 2 kb *Bam*HI*/Hin*dIII fragment and cloned into the *Bam*HI and *Hin*dIII sites of pDONR201:*PpMYB4* using T4 DNA ligase (Roche) such that the *nptII* gene cassette replaced almost the complete *PpMYB4* CDS to leave ~1 kb of *PpMYB4* gene sequence flanking each side of the *nptII* gene cassette. Similarly, for *PpMYB5*, the *nptII* gene cassette from pMBL10b was digested with *Sma*I and *Pvu*I (~2 kb) and cloned into the *Xmn*I site of pDONR201:*PpMYB5* to leave ~1 kb of *PpMYB5* gene sequence flanking each side of the *nptII* gene cassette.

Gene disruption was confirmed in each case by sequencing and then inserts in the pDONR vectors were recombined into the destination vector pMBL6attR (accession number DQ228132), a gift from Dr. A Cuming.

5.2.5. Physcomitrella transformation

5.2.5.1. Transformation

The sterile transformation procedure followed the method of Didier Schaefer described in http://www2.unil.ch/lpc/docs/pdf/Transfo_moss.pdf with some modifications.

Collected protonemata from 4-6 plates grown for 5-6 days were transferred to a 9 cm Petri dish containing a sterile solution of 0.48M mannitol with 250 mg Driselase (Sigma D-9515) for 30 min at RT with occasional gentle mixing. The protoplasts were filtered through a 100 µm nylon mesh (Millipore NY1H04700) and left for an additional 15 min to complete digestion. Then, they were collected by centrifugation for 5 min at 700 rpm and resuspended in 10 ml 0.48M mannitol. The wash was repeated two more times. Protoplasts were

counted with a hematocytometer and resuspended in MMM media (0.48M mannitol, 15mM $MgCl_2$ and 0.1% [w/v] MES-KOH [pH 5.6]) at a concentration of $1.2x10^6$ cells/ml. *Not*I linearised DNA of each plasmid (~ 15 µg) containing the disrupted gene construct, was resuspended in 30 µl dH_2O and placed into a polypropylene Falcon tube. Then 300 µl of the protoplast suspension and 300 µl of PEG solution (0.38M mannitol, 0.1 M $Ca(NO_3)_2$, 35-40% [w/v] PEG 4000, 10mM Tris-HCl [pH 7.2]) were added. The mixed suspension received a heat shock for 5 min at 45ºC and was then left at room temperature for 10 min with mixing. After this, protoplasts were progressively diluted to 6.5 ml with proto liquid media (KNOPS media without agar plus 66 g/l mannitol) and kept in darkness o/n. The protoplasts were collected by centrifugation for 5 min at 700 rpm and resuspended in 7.5 ml proto liquid media and 7.5 ml top agar (0.48M mannitol and 1.4% [w/v] agar; this medium was warmed to 45ºC to prevent gelling). This mix was poured into 4-5 Petri dishes (9-cm), containing protoplast regeneration medium (proto liquid media supplemented with 10mM $CaCl_2$ and 0.8% [w/v] DIFCO agar) overlaid with cellophane discs. Plates were placed for 6-7 days in a growth chamber at 23ºC±2ºC with a light regime of 16 h light/8 h darkness. Regenerating colonies were subjected to G-418 (MELFORD) selection (50 µg/ml) on Pp-NH_4 plates for 10 days. Stable resistant clones were selected after a second round of growth on non-selective Pp-NH_4 medium followed by transfer on Pp-NH_4 medium containing G-418 (50 µg/ml).

5.2.5.2. Molecular analysis of the transformants

Targeting events were analyzed by PCR. For this purpose, genomic DNA of wild-type and transgenic plants was extracted and the 5′ and 3’ integration events were confirmed by PCR experiments with the primers shown in Table 1, which annealed to the genomic sequence of *P. patens* (surrounding each *MYB* gene) and to the *nptII* selection cassette (Fig. 5.7) PCRs were carried out in a total volume of 25 µl as described in the General Materials and Methods section 6.

Gene/Oligo pair	5' extreme	3' extreme
PpMYB4 (pdp07793)	*SJS61*/ SJS54	*SJS33*/ SJS55
PpMYB5 (pdp12446)	*SJS33*/ SJS52	*SJS34*/ SJS53

Table 1. Primers used to detect stable transformants

Oligo pairs used to detect the insertion of the *nptII* cassette in the *PpMYB* genes. Oligos in italics anneal in the *nptII* cassette.

5.2.6. Southern blot analysis

P. patens genomic DNA from wild-type and transgenic plants (2 µg) was digested with *HindIII* and fragments were separated by electrophoresis in 0.7% agarose. The gel was then treated with 0.25M HCl for 10 min and with 0.4M NaOH for 30 min and transferred to a nylon membrane (Hybond-Nx, Amersham) in 20 × SSC (3 M NaCl, 0.3 M sodium citrate, pH 7) o/n. The probe, a fragment from the pMBL10b nptII cassette was PCR amplified with the primers nptIIFw/nptIIRv and radiolabeled with the RediprimeII Random Prime Labelling System (Amersham). Unincorporated radiolabeled nucleotides were removed with the Illustra Probe Quant G-50 (Amersham). All hybridizations were at 60°C. The membrane was prehybridized with Church buffer (0.5 M phosphate buffer, 7% w/v SDS, 1 mM EDTA) and 10 mg/ml salmon sperm DNA for 1 h. Hybridization occurred in the same solution with the addition of the α-^{32}P-dCTP-labelled probe. Following hybridization, the membrane was washed 3 times in 0.1xSSC and 0.5% SDS. Detection was accomplished with a Phosphoroimager.

5.2.7. Extraction of phenolics

Phenolic extraction and LC-MS was carried out as indicated in Materials and Methods in Chapter 4 with slight modifications. For Physcomitrella, 7 day-old moss tissue was grown under standard conditions. After harvesting, the tissue was gently squeeze-dried with Whatman paper and then freeze-dried for 2-3 days. Three biological replicates from two independent transformants for each knockout line were used for the metabolic analysis. 50µl of sample was

mixed with 50µl water, centrifuged, and 80µl supernatant was transferred to small glass inserts in vials for analysis. Injections were 15µl. The experiment was repeated three times in similar conditions and the results presented show a summary of the compounds likely to phenylpropanoids in Physcomitrella and changes in the levels of these compounds in the *PpMYB4* and *PpMYB5* knockout mutants. Because there are no published data on phenylpropanoids in *P. patens*, any identification of particular peaks as representing phenylpropanoids is tentative, based on their water solubility, mass and absorbtion at 215nm and 325nm. UV differences found at 215 and 325nm were analyzed and a brief “metabolomic” approach was carried out using xcms to find peaks and multiple t-tests in Excel to check which are most likely to have changed. The t-tests here are used exclusively to narrow down which peaks were most worthy of further attention, not to demonstrate that they have genuinely changed. Xcms is a data-analysis package written in the statistical language “R”. Xcms opens MS data, finds peaks, corrects for variations in retention time between runs, and makes a basic comparison of two treatments peak-by-peak to find those peaks that differ the most. Since there were multiple treatments, xcms was used only as a peak-picking and retention time correcting software. Each peak is a mass eluting at a specified retention time, and they are named by mass and retention time (in seconds) with names such as M360T540, M for mass, T for time). One chemical can appear as several masses (the base peak but various adducts, isotope peaks, and fragments). Before carrying out any analyses on all the peaks found by the xcms, the unreliable ones were removed. This was done based on a quality control sample by mixing 10µl each of all 18 samples with 108µl water and the supernatant was assayed four times scattered through the sequence of analytical runs on the LC/MS. The quality control sample runs should give the same area for each peak every time the QC sample is run (they are replicate runs of the same sample). Consequently, the variation in a peak, its relative standard deviation, will be small if it is a good peak. Quality control sample runs were used to reject any peaks that had a relative standard deviation greater

than 30%. This approach for quality control of samples was developed by Roy Goodacre and Warwick Dunn (University of Manchester). The methods used to look at the remaining peaks were firstly, the change in metabolites on a fold basis (*i.e.* if the mean in wild-type was 3.0 and in mutant was 6.0, this would constitute a 2-fold change), which was calculated expressing the fold change as bigger ÷ smaller, so all changes come out as greater than 1, and the principal component analysis (PCA) plot of the data.

For moss *PpMYB4* and *PpMYB5* genes overexpressed in Arabidopsis, a similar approach was taken. Plants already bolting (as recommended for JIC staff for phenolics extraction), grown on soil, were collected. For the phenolic analysis two independent homozygous lines (3 replicates per line) for each gene were analyzed.

5.2.8. UV-B treatment

To compare the moss response to UV-B light with the responses in Arabidopsis seedlings, the same treatment was carried out as described in Chapter 4 section 4.2.3 but with 7 day-old moss cultures grown in KNOP medium.

5.3. Results

5.3.1. Cloning, expression and characterization of *PpMYB4* and *PpMYB5*

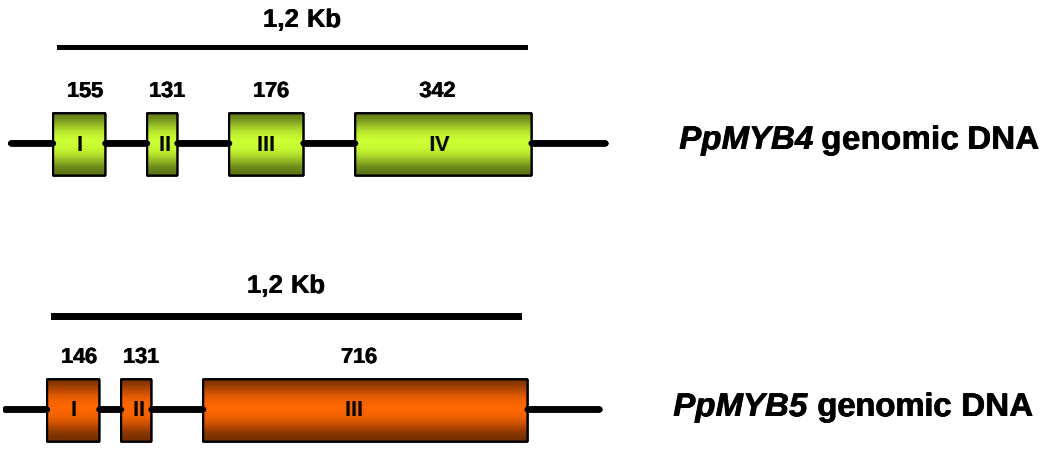
A phylogenetic analysis carried out by Dr. Bailey revealed that the two putative *P. patens* MYB proteins PpMYB4 and PpMYB5, aligned closely in the phenylpropanoid clades of MYB regulators and with AtMYB6 and AtMYB8 (Fig. 5.4). PpMYB4 and PpMYB5 EST clones (BRC resource number pdp07793 and pdp12446 respectively, http://www.brc.riken.jp/lab/epd/catalog/p_patens.html) were provided by the BRC resource centre following a TBLASTN search of the *P. patens* EST library conducted by Dr. P. Bailey with the MYB domain of MYB4 from *Arabidopsis thaliana* (GenBank™) as a query.

The cDNA clones were fully sequenced and the largest open reading frame was 804 bp for *PpMYB4* and 993 bp for *PpMYB5* respectively. The genes were predicted to encode proteins of 267 amino acids with an estimated molecular mass of 29.6 kDa in the case of *PpMYB4* and 330 amino acids with an estimated molecular mass of 36.7 kDa for *PpMYB5*. The two proteins were 36.6 % identical at the amino acid level.

PpMYB4 (Phypa_235528) and *PpMYB5* (Phypa_235529) genomic sequences were retrieved from the DOE Joint Genome Institute, JGI (http://genome.jgi-psf.org/Phypa1_1/Phypa1_1.home.html) after a BLASTN search of the data base with the *PpMYB4* and *PpMYB5* ORF sequences as queries. The EST sequences confirmed the genome annotation and Figure 5.5A, shows the exon/intron structure of the *PpMYB4* and *PpMYB5* genes from *P. patens.* In the *PpMYB4* gene there are three introns whereas in *PpMYB5* there are only two introns. Despite their differences in structure, *PpMYB4* and *PpMYB5* genes have almost identically sized coding sequences (also known as ORFs), approximately 1.2 kb from the start codon to the stop codon. Only *PpMYB4* and *PpMYB5* exon 1 and exon 2 positions are quite well conserved (relative to the aligned genomic sequences between the two *Physcomitrella* genes and those of *Arabidopsis*) with *AtMYB8* exon 1 and exon 2 positions.

Similar to MYBs transcription factors from Arabidopsis, PpMYB4 and PpMYB5 proteins from *P. patens* contained the typical conserved N-terminal MYB domain of the R2R3 MYB-like proteins consisting of two 52 amino acid repeats, each predicted to adopt a helix-turn-helix fold. This domain binds DNA (Paz-Ares *et al*., 1987; Martin and Paz-Ares, 1997) and the greatest conservation of amino acid sequence between PpMYB4, PpMYB5 and other R2R3 MYB proteins was in those regions of the MYB domains predicted to interact most closely with DNA; the C-terminal helices of each MYB repeat (Ogata *et al*., 1994). A comparison of the deduced amino acid sequences of the two Physcomitrella gene products with AtMYB6 and AtMYB8 proteins revealed a high level of conservation mostly restricted to the MYB domain whereas the

A



B

```
                                                  R2
AtMYB6  MGRSPCCEKAHT--------NKGAWTKEEDQRLVDYIRNHGEGCWRSLPKSAGLLRCGKS  52
AtMYB8  MGRSPCCEKAHM--------NKGAWTKEEDQRLIDYIRNHGEGSWRSLPKSVGLLRCGKS  52
PpMYB5  MGRAPSSSRVTTGEGL----NKGPWTAEEDSILTAFVRANGEGNWRILPKRAGLKRCRKS  56
PpMYB4  MGLFLDRCKSRRSGGT-DPVSKGPWTEAEDKILTECVEKYGDRCWRQIPKHGGLLRDGKS  59
                                                  R3
AtMYB6  CRLRWINYLRPDLKRGNFTDDEDQIIIKLHSLLGNKWSLIAGRLPGRTDNEIKNYWNTHI  112
AtMYB8  CRLRWINYLRPDLKRGNFTDGEEQIIVKLHSLFGNKWSLIAGKLPGRTDNEIKNYWNTHI  112
PpMYB5  CRLRWMNYLRPDLKRGNFSPDEDELIIKLHSLLGNKWSLIAGRIPGRTDNEIKNYWNSRL  116
PpMYB4  CRLRWKNYLKPGLKRGKFSKDEDEFIITLHALLGNRWSLIAGRIPGRTDNEIKNHWNTHL  119

AtMYB6  KRKLLSHG------ IDPQTHRQINESKTVSSQVVVPIQN----------------DAV 148
AtMYB8  KRKLLNRG------ IDPKTHGSIIEPKTTS----FHPRN----------------EDL 144
PpMYB5  KRNLQDMEGNQRQQLVATPLSVSLLDSSTEIKVEDFPCSNQLGMSGPPSPIDEKPPMEAL 176
PpMYB4  KKRLREMG------ IDHENHSVGLKLSSDSDGSLDAEES--------------KGLEF 157

AtMYB6  EYSFSNL-------AVKPKTENSSDNGASTSGTTTDEDLRQNGECYYSDNSGHIKLN--- 198
AtMYB8  KSTFPG--------SVKLKMETSCENCASTSGTTTDEDLRLSVDCDYRYDHLDKELN--- 193
PpMYB5  LGTHHHSNDLDPSIITKPQLLNSTDFHAMPRGESTLSEPTSSDGTCFGESCDHQSLEGYS 236
PpMYB4  ESTFN----------CQVDQGPYTDDPCSSASCTYDNGLSGLEQPCPKLPQSENSLDFN- 206

AtMYB6  -------------LDLTLGFGS---------------------------WSGRIVGVG 216
AtMYB8  -------------LDLTLGYS-----------------------------PTRFVGVG 209
PpMYB5  SSCNWAPFDNTQKPLSFTQASSPESVLLFQQYFPSQPQDDFLCSDFGLTDFTSTNYIPEF 296
PpMYB4  -------------LEKSLGSSFKTYPSPTDR-----------LAFDMMDYLPLSLMGLS 241

AtMYB6  SSADSKPWCDPVMEARLSLL------------- 236
AtMYB8  S-------CY------------------------ 212
PpMYB5  TSLQNLLQCGPFSPIHKNLPATWIGNYSGDGFQT 330
PpMYB4  GFAEESPHEDDVSGPLWTSIQLFDHS-------- 267
```

Figure 5.5. *(A) PpMYB4* and *PpMYB5* genes structure. Coloured regions show the exons and their size in base pairs. Non-coding DNA is shown as a solid black line. (B) Comparison of the deduced amino acid sequences of PpMYB4 and PpMYB5 proteins with the AtMYB6 and AtMYB8. The multiple alignment was generated using the computer program Clustal-W. Yellow shading indicates amino acids conserved in all entries and grey shading indicates amino acids identical in some of the entries. Above the green bar is represented the EAR motif-like LPLSLMg.

C-terminal domain was poorly conserved (Fig. 5.5B). PpMYB4 has an EAR-domain-like LxLxL motif, whereas PpMYB5 does not. Both AtMYB6 and AtMYB8 contain a putative EAR domain, LXLXLXP,(Keiichiro *et al*., 2003) in their C-terminal regions (LDLTLGf for AtMYB6 and LDLTLGy for AtMYB8) and PpMYB4 contains a similar motif towards its C-terminus; LPLSLMg) PpMYB5 does not have a motif resembling the EAR motif and seems to be the most distinct protein in the alignment. The PpMYB4 and PpMYB5 sequences are as closely related to each other as to the AtMYB6 and AtMYB8 sequences, sharing between the four proteins approximately 35-40% identity over their entirety. Over the last year, new sequencing data for *P.* patens have been released and there are two new gene sequences closely related to *PpMYB5* and *PpMYB4* (NCBI reference sequences XP_001783614.1 and XP_001752936.1 respectively) which confirms the whole genome duplication event predicted by Rensing *et al*. (2008).

In order to determine how closely related were the PpMYB4 and PpMYB5 proteins to AtMYB6 and AtMYB8, a phylogenetic tree was constructed based in the sequences of their DNA binding domains using the Neighbor-Joining method (PHYLIP package 3.67) (Fig. 5.6A). TBLASTN searches with the Genbank database at NCBI using the MYB domain of AtMYB4 were carried out to identify homologous genes in the Physcomitrella genome. Hits found were aligned against a dataset of aligned MYB domains from Arabidopsis and rice R2R3 MYB proteins (for clarity the rice MYB proteins are not shown in Figure 5.6A). Phylogenetic analysis showed that the PpMYB4 and PpMYB5 proteins align most closely to R2R3 MYB proteins from higher plants known to regulate the expression of genes involved in phenylpropanoid metabolism. Bootstrap analyses for 100 replicates were performed to provide estimates of confidence for the tree topologies. Bootstrap support was good for some branches, but not strong for those branches separating subgroup 4 from subgroup 7 and for the branch separating subgroups 5 and 6 from the other subgroups (Stracke *et al*., 2001). Depending on the sequences included in the tree dataset, the moss proteins tend to cluster with the Arabidopsis R2R3 MYB

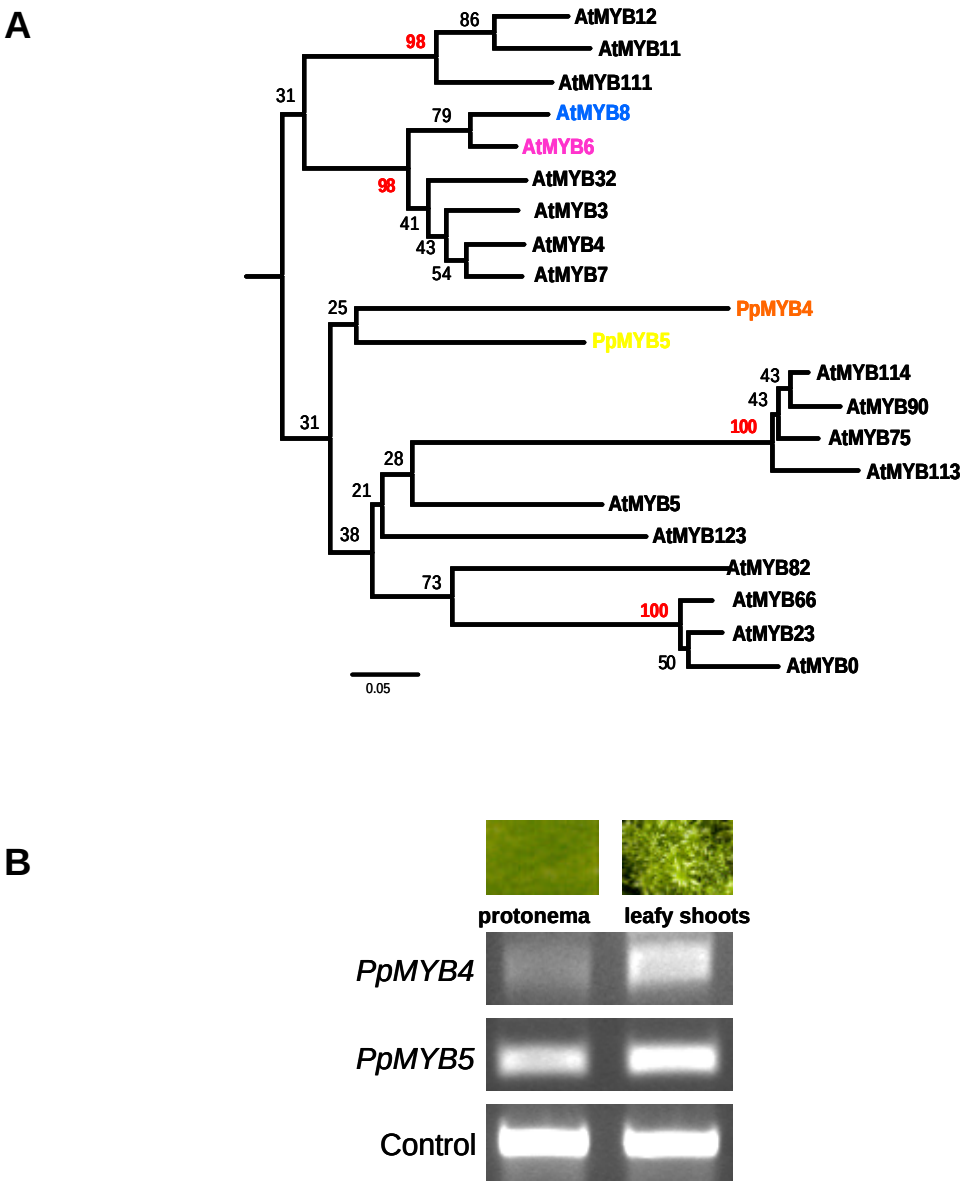


Figure 5.6. (A) PHYLIP 3.67 tree based on the full Arabidopsis thaliana R2R3MYB neighbour-joining tree which was made using the R2R3 DNA binding domain. Numbers at the branches are bootstrap values (%), significant values in red. This fragment of phylogenetic tree shows the relationships between some AtMYB transcription factors and PpMYB4 and PpMYB5. The PpMYBs (showed in orange and yellow) are located in branches corresponding to AtMYB members involved in the flavonoid pathway which suggests that it could be a conserved ancestral function. (B) RT-PCR analysis of the *PpMYB4* and PpMYB5 transcript in two developmental stages. Lower panel: internal control PCR of the constitutively expressed gene for the ribosomal protein L21.

proteins regulating anthocyanin and proanthocyanin biosynthesis (subgroups 5 and 6) but there is no strong statistical support for this. Alternatively, PpMYB4 and PpMYB5 might share a common ancestor with members of subgroup 4 and two of the moss sequences (PpMYB4 and its paralog XP_001752936.1) contain the GIDP motif that is present in most phenylpropanoid MYBs (subgroup 4) and absent in most flavonoid-related MYBs (subgroups 5 and 6).

The expression of *PpMYB4* and *PpMYB5* in protonemata and leafy shoots was analyzed using RT-PCR (SJS24-SJS69 and SJS59-SJS60 oligo pairs for *PpMYB4* and *PpMYB5* respectively) and mRNA levels were compared to the L21 ribosomal protein transcripts, used as a control (L21 primers forward and reverse, see appendix1). To describe the expression patterns of each *PpMYB* gene, specific primers were designed from central regions of the cDNAs. As illustrated in figure 5.6B, *PpMYB4* and *PpMYB5* were expressed in both protonemata and leafy shoots, although transcripts of *PpMYB4* seemed to be less abundant especially in the protonemal tissue.

5.3.2. Gene targeting for the generation of knockout mutants

To identify the functions of *PpMYB4* and *PpMYB5,* I needed to generate mutations that eliminated their activities. Genomic clones were used to construct the KO cassettes that were subsequently employed to transform *P. patens* protoplasts to generate knock-out mutants of *PpMYB4* and *PpMYB5* through PEG-mediated transformation of moss protoplasts.

To synthesize the disruption constructs, almost the whole coding sequence for *PpMYB4* and *PpMYB5* was replaced by the *nptII* gene (as a selection marker) flanked on both sides by ~1 kb of gene-specific DNA to allow the homologous recombination (Fig. 5.7A,B). Before transformation, plasmid DNA was linearised by digestion. To identify positive recombination events, approximately 60 regenerated G418-resistant colonies were prescreened by PCR with two different primer combinations to detect the disruption of the *PpMYB* genes and the correct integration of the transgene on both the 5′- and 3′-ends (Fig. 5.8A). For two *PpMYB4* transformants and six *PpMYB5*

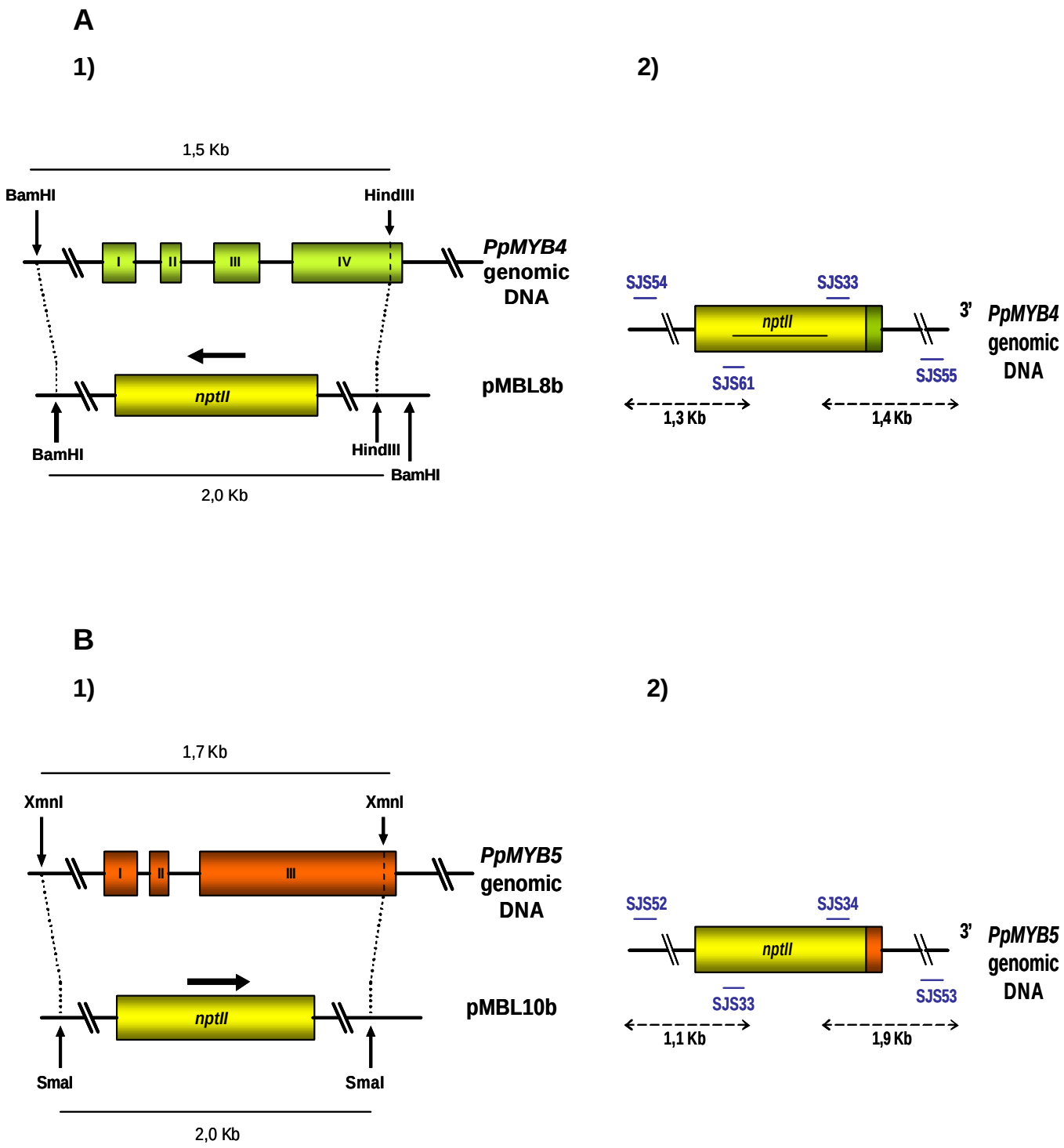


Figure 5.7. *PpMYBs* genes targeting strategy and knockout constructs

Structure of the *PpMYB4* (A1) and *PpMYB5* locus (B1) and the replacement vector containing the nptII cassette (yellow) used for gene disruption. The exons of both genes are represented by green or orange boxes respectively. Location of primers used to screen for targeted transformants is shown by arrows (A2, B2).

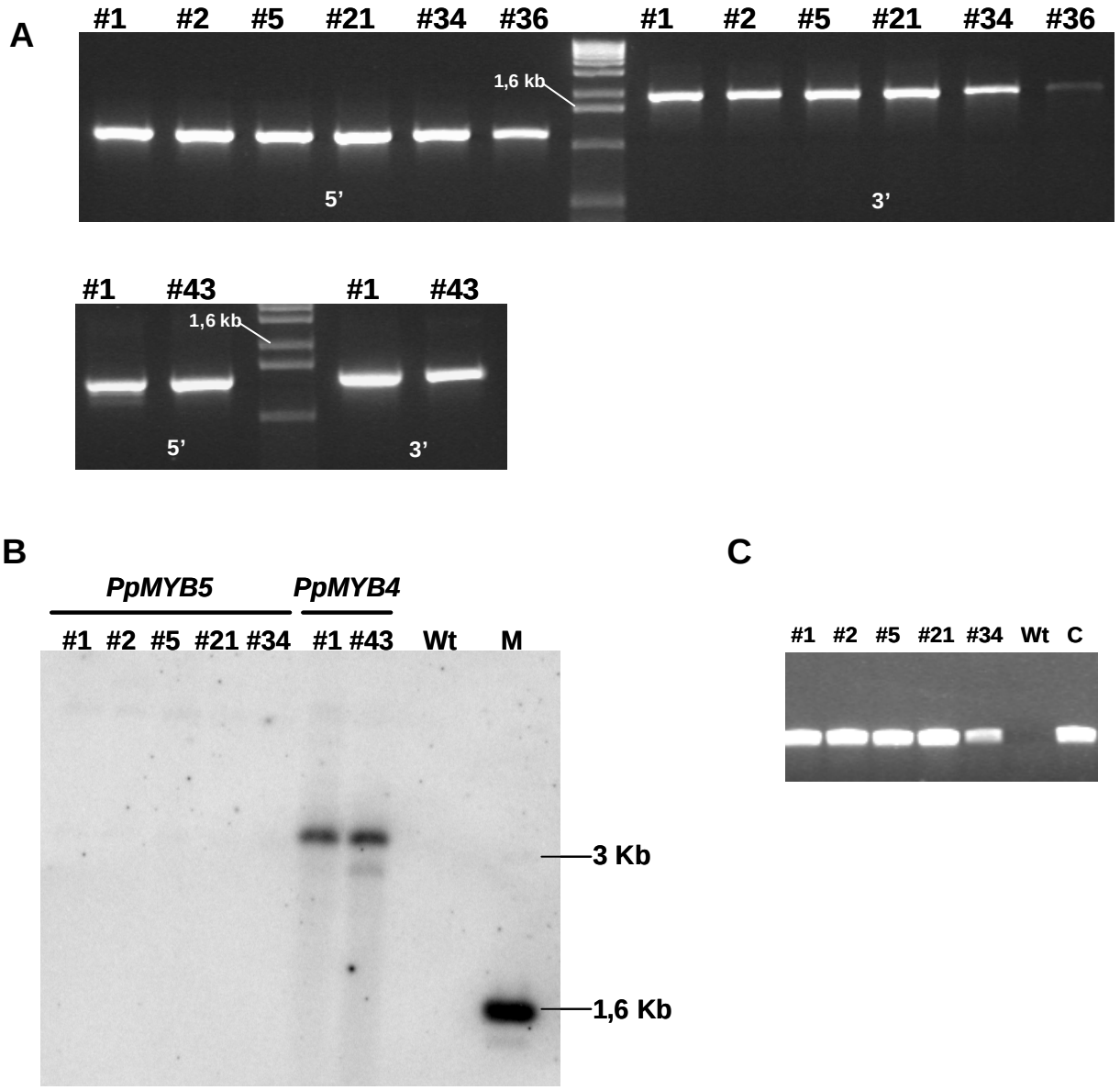


Figure 5.8. (A) PCRs of genomic DNA from *PpMYB5* (top picture) and *PpMYB4* knockouts (lower). The pictures show: product from insertion cassette to untargeted 5' region; marker and product from untargeted 3' region to insertion cassette. (B) Southern analysis of the *PpMYB5* and *PpMYB4* locus in the putative knockouts mutants. *Physcomitrella patens* DNA was digested with *Hin*dIII. Positions of the size markers are shown on the right. The *nptII-800* fragment was ^{32}P-radiolabeled and used as a probe. (C) PCR showing *PpMYB5* putative knockouts mutants (#1- #34) containing the nptII cassette (Wt, wild-type as negative control and C, as a positive control).

transformants both PCR reactions resulted in the expected products, indicating that these events represented true knockouts of the *PpMYB4* and *PpMYB5* genes. From these, two individual transformants for each gene (*PpMYB4 #1 and #43, PpMYB5* #5 and #21), were selected for further analyses.

P. patens genomic DNA mutants were analyzed by Southern blot hybridisation with the internal *nptII* gene fragment as a probe (Fig. 5.8B). The hybridization pattern indicated a unique *PpMYB4* copy in the moss genome although only a faint signal was detected for the *PpMYB5* knockouts. This signal may have been weaker than for the *PpMYB4* knockouts because the DNA fragments detected were larger and did not transfer so efficiently to the membrane prior to hybridisation. The *PpMYB5* knockouts were checked by PCR using just the external gene-specific primers and confirmed to contain the entire nptII cassette (Fig. 5.8C).

5.3.3. Phenylpropanoid profile in *PpMYB4* and *PpMYB5* knockout mutants

To address the question of whether the disruption of the *PpMYB4* and *PpMYB5* genes impacted phenylpropanoid metabolism, phenolic compounds were extracted and measured in the knockout mutants as well in the wild-type protenema.

There are no reports available of phenolics present in *P. patens* although flavonoids have been identified in other mosses (Melchert and Alston, 1965; Anhut *et al*., 1984; Stafford, 1991; Geiger and Markham, 1992; Brinkmeier *et al*., 1999).

A standard method for extracting water-soluble phenylpropanoids was adopted and extracts of wild type *P. patens* were analysed by LC/MS. Profiles were monitored by UV absorbtion at 214 nm and 325 nm, the range in which the phenolic compounds absorb (Fig. 5.9). A few strong peaks were observed in wild type protenemata, although it was not possible to identify any on the basis of their m/s data alone (ie. compounds of specified masses did not give recognisable fragmentation patterns in MS2). Lack of literature on phenylpropanoid compounds in mosses allowed us to do only a general

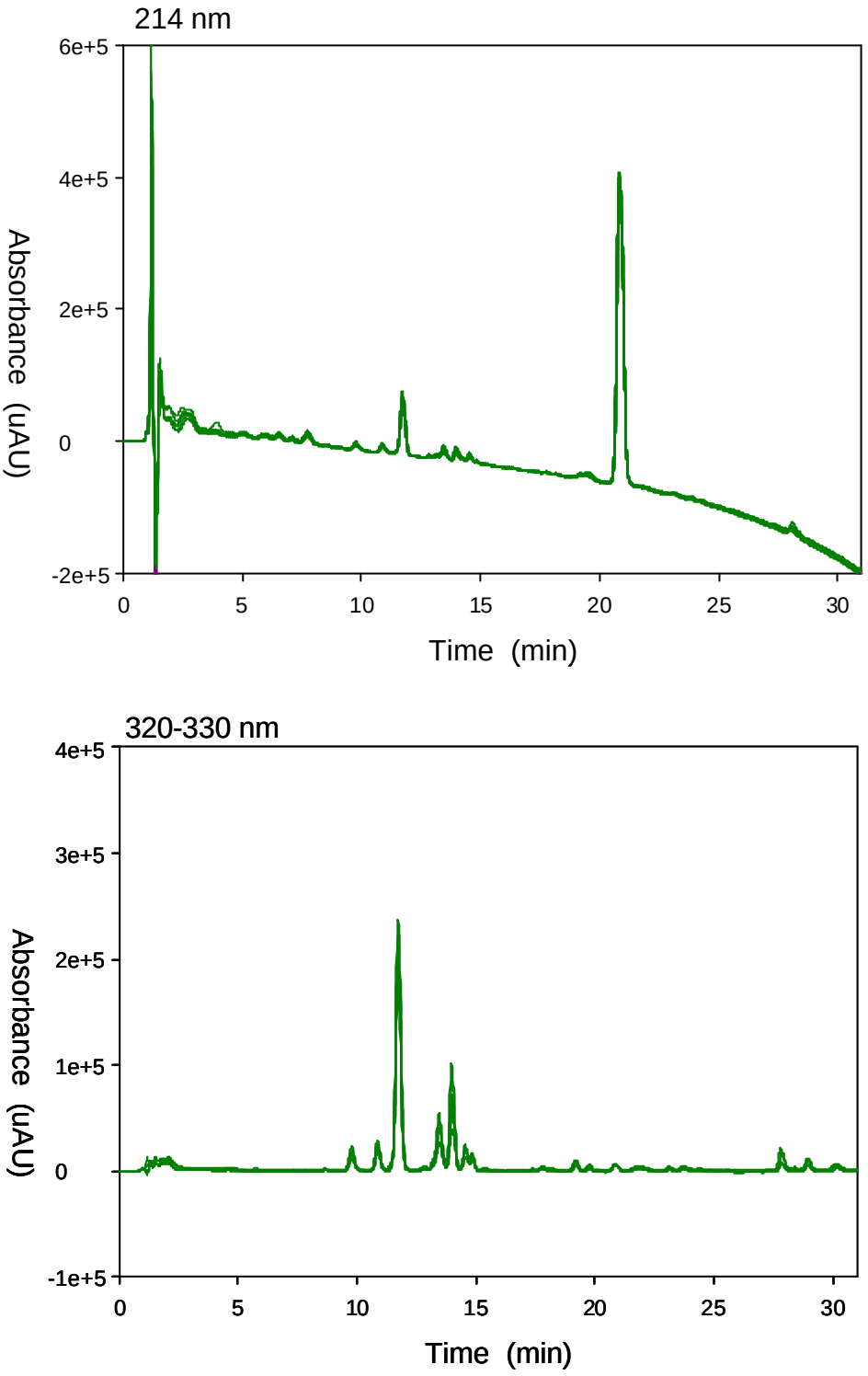


Figure 5.9. LC/MS of the phenylpropanoids extracted from *P. patens*

The figures show the UV spectra at 214 and 320-330 nm of the hypothetical phenolic compounds that can be found at different retention times in the natural Physcomitrella. Lack of literature about moss phenylpropanoids makes it difficult to identify those peaks associated with phenolic compounds.

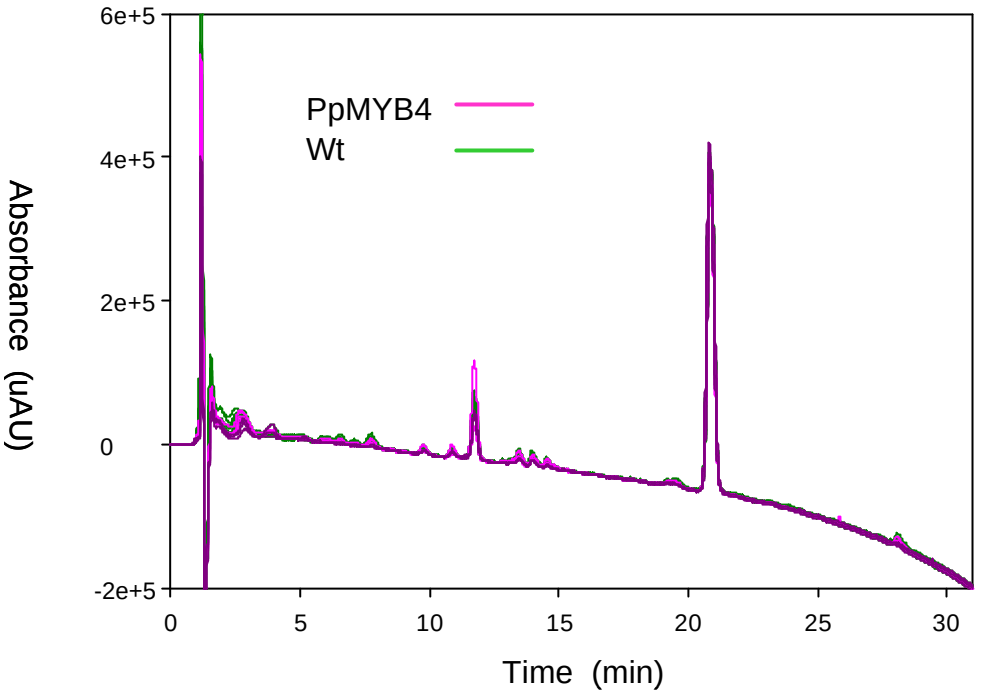
analysis of the peaks obtained. Analysing the extracted phenolics under the appropriate UV range, we discarded compounds containing nitrogen as these ones are not classified as phenylpropanoids. We identified nitrogen-containing compounds/peaks using the “even nitrogen rule” that states that compounds with an odd mass that fragment to an even mass must contain nitrogen. For example, one compound with a retention time of 12.4 minutes had a mass of 287 amu and was classified as ‘kaempferol-like’ (because kaempferol has a mass of 287, but the fragmentation pattern did not match that of kaempferol as this fragmented from an odd mass to an even mass). Another compound had a retention time of 11.4 minutes and a mass of 273 amu. This was classified as ‘naringenin-like’ (because naringenin has a mass of 273, but the fragmentation pattern did not match that of naringenin).

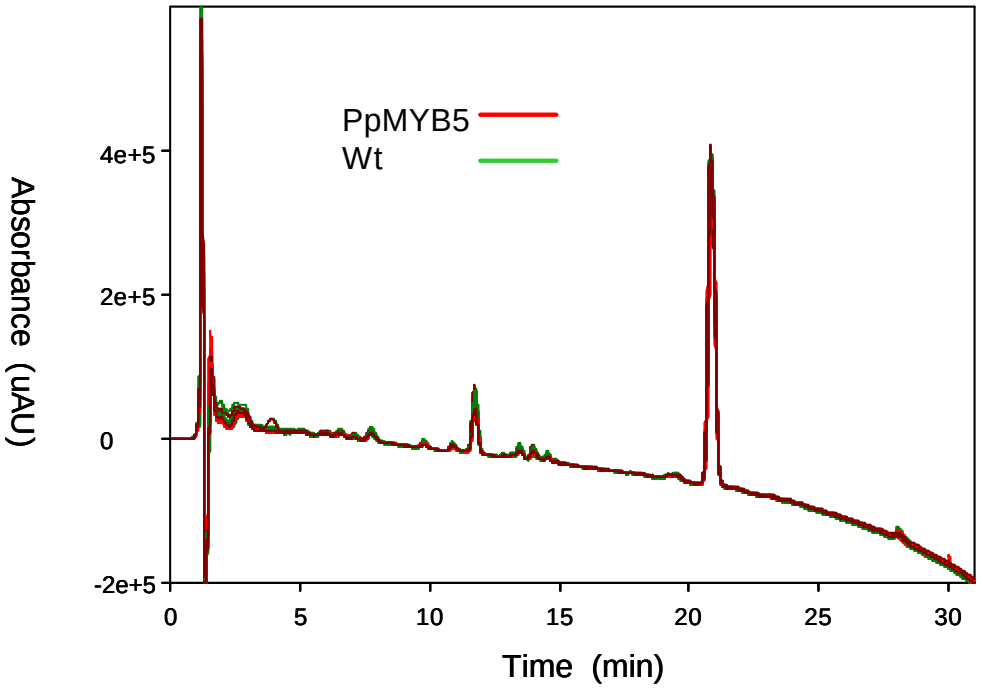
Extracts from independent *PpMYB4* and *PpMYB5* knockouts were then analysed by LC/MS. Three biological replicates from two independent transformants for each knockout line were analyzed by LC-MS. Although some differences were apparent by overlapping these profiles with the wild type profiles, the different biological replicates showed rather variable levels of compounds absorbing at 214nm (Fig. 5.10A) and 325nm (Fig. 5.10B). These differences in UV absorption were examined and an untargetted metabolomic analysis was undertaken to identify those peaks most likely to have changed in the mutants compared to the wild type protonemal tissue. Unreliable peaks were rejected (see methods) and then a principal component analysis (PCA) was undertaken (Fig. 5.11). The PCA determines whether the mutant(s) and the wild-type can be discriminated from each other on the basis of their metabolic profiles. The plot with the two principal components accounting for most of the variation between samples clusters *PpMYB4-KO* mutants to the right of wild type for principal component 1, x-axis. There was little separation between wild-type and *PpMYB5-KO* mutants with this component, but *PpMYB5-KO* mutants fell lower on the y-axis, for principal component 2 compared to wild type.

Figure 5.10. LC/MS of the phenylpropanoid extracts from *P. patens* wild-type and *PpMYB4* and *PpMYB5* knockouts

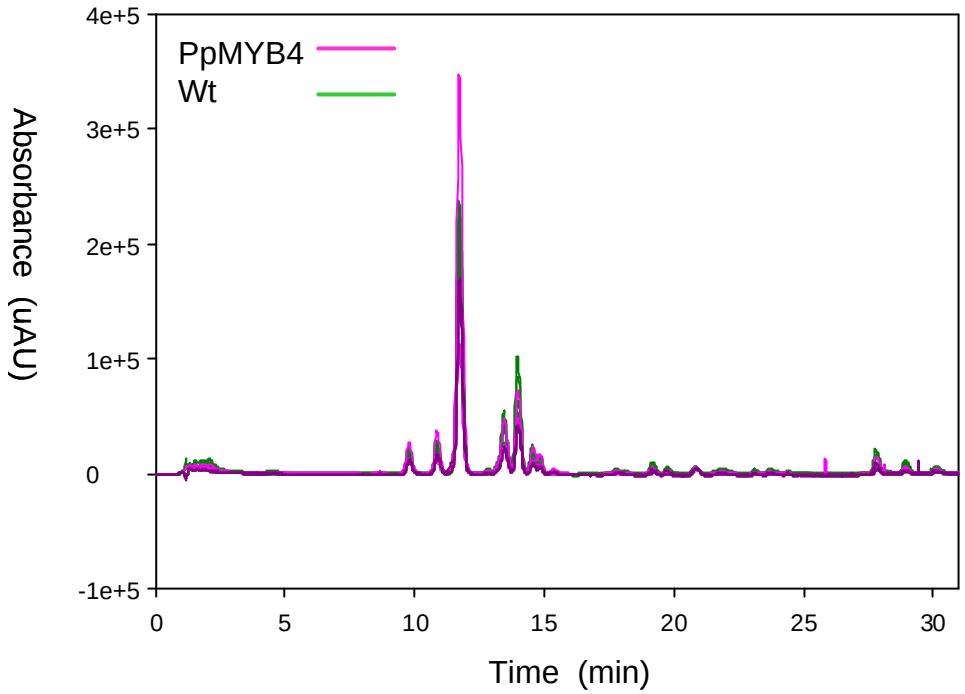
The figures show the UV spectra at 214 (A) and 325 nm (B) of the hypothetical phenolic compounds that can be found at different retention times in the wild-type (green lines) compared to the *PpMYB4* knockout (pink lines) and *PpMYB5* knockout (red lines).

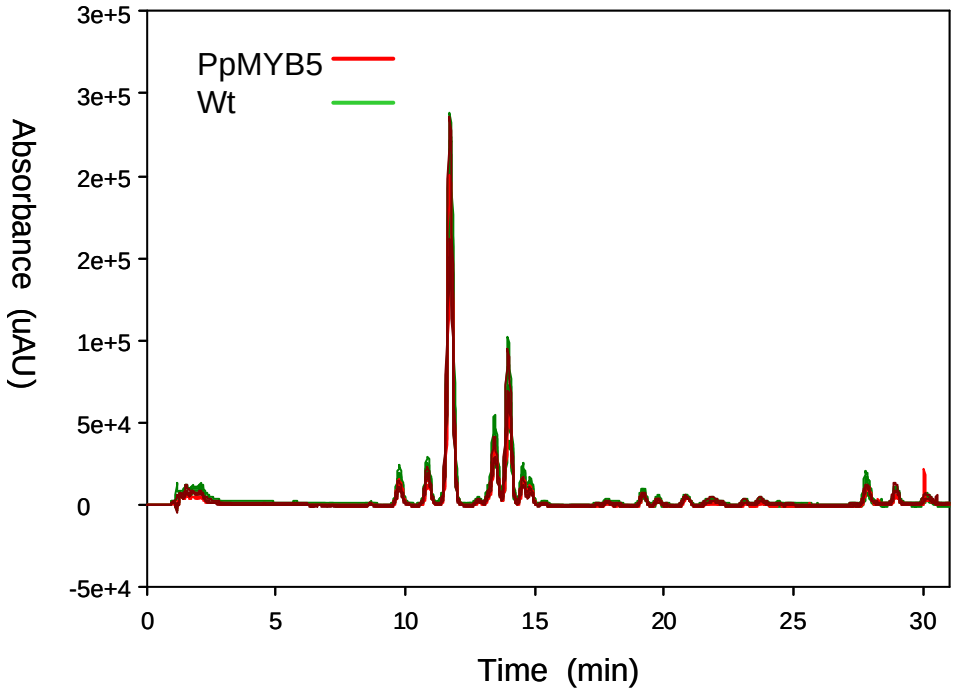
A


6e+5
4e+5
2e+5
0
-2e+5
Absorbance (uAU)
PpMYB4
Wt
0
5
10
15
20
25
30
Time (min)


4e+5
2e+5
0
-2e+5
Absorbance (uAU)
PpMYB5
Wt
0
5
10
15
20
25
30
Time (min)

B


PpMYB4
Wt
Absorbance (uAU)
4e+5
3e+5
2e+5
1e+5
0
-1e+5
0
5
10
15
20
25
30
Time (min)


PpMYB5
Wt
Absorbance (uAU)
3e+5
3e+5
2e+5
2e+5
1e+5
5e+4
0
-5e+4
0
5
10
15
20
25
30
Time (min)

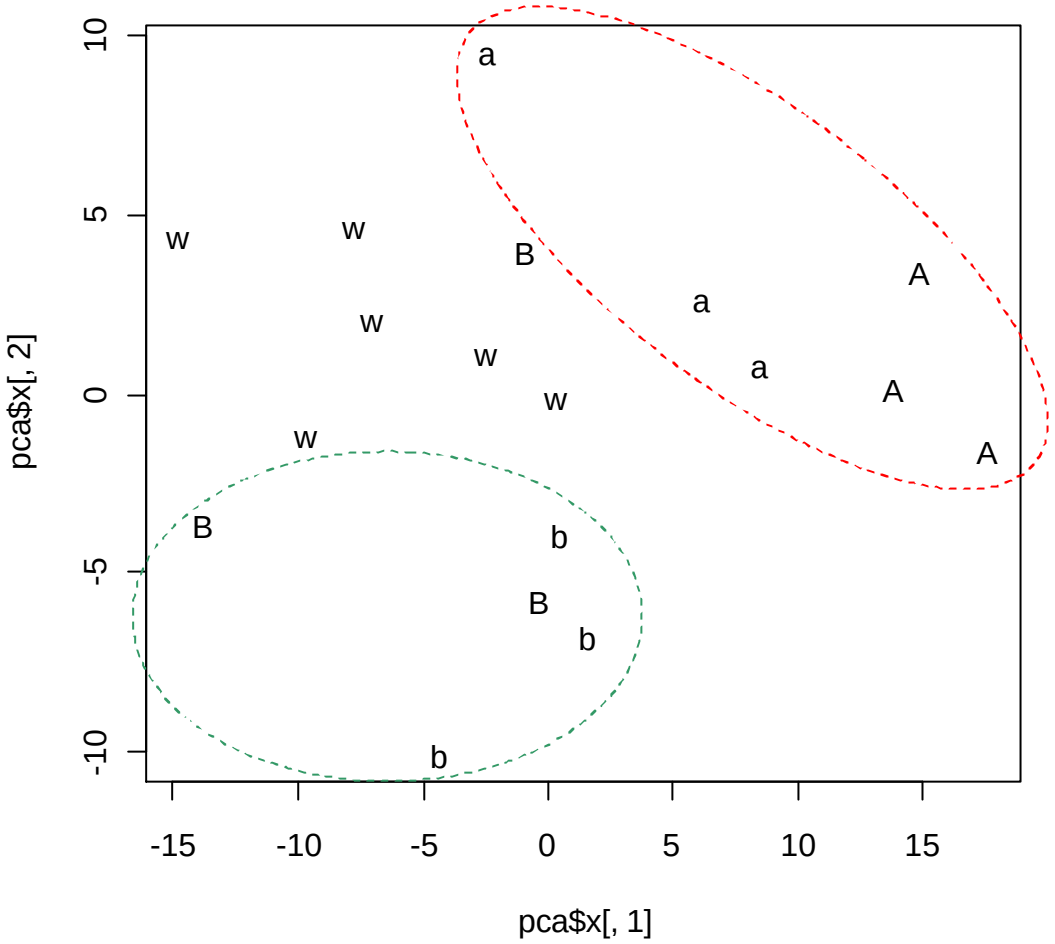


Figure 5.11. Metabolic phenotype clustering using the principal component analysis (PCA)

Data of 18 samples originating from three plant genotypes. The plot shows each individual sample represented by a letter and its position indicates its chemical content. W, wild-type; A and a, *PpMYB4-KO* mutants; B and b, *PpMYB5-KO* mutants. Mutants show metabolic phenotypes distinct from wild-type plants (Wt).

In PC1 the peak with greatest variability compared to the Wt was peak 386 which had a retention time of 357sec (6min 57sec) and a mass of 389. This compound was two fold lower in the *PpMYB4* knockouts compared to wild type. However, the MS2 data available for this compound, showed it fragmented from a compound of odd mass to a fragment of even mass, suggesting it contained nitrogen, and was therefore not a phenylpropanoid. Looking more closely at the principal components separating *PpMYB4* knockouts from wild-type, PC1 did not appear remarkably more changed than other compounds. In PC2 the most notable difference was given by peak 305 which had a retention time of 1221 sec (20min 21sec) and a mass of 453. There was less of this compound in *PpMYB5* knockouts compared to wild type, but only on average 1.1 fold less suggesting that levels of this compound were not affected substantially by loss of *PpMYB5* function. Because the PCA analysis did not identify compounds that were substantially altered in the knockout mutants compared to wild type, an alternative statistical analysis using t-tests for each genotype was adopted. This method identified a small number of peaks that gave very big differences in t-tests (Fig. 5.12).

The untargeted analysis for *PpMYB4-KOs* compared to wild type, subjected to t-tests, revealed an average 2.5-fold decrease in the majority of peak heights compared to the wild-type (Fig. 5.12A). However, one peak with a mass of 403, was increased by almost 3-fold. This analysis supported the idea that loss of *PpMYB4* activity reduced the level of a number of phenylpropanoids in moss protenema, although identification of individual peaks was not possible. When the peaks that showed the greatest differences in the *PpMYB4* knockout compared to wild type were identified amongst peaks sorted in order of change in the *PpMYB5* knockout compared to wild type, the peaks most significantly altered by loss of *PpMYB4* function were not particularly highly ranked amongst the peaks affected by loss of *PpMYB5* function (compared to wild type), see Figure 5.12B. Finally when the peaks affected by loss of *PpMYB5* function were compared to wild type, a number of peaks showed reduced levels in the mutants, but the statistical significances of these differences were lower

A

	name	mzmed	rtmed		to wt
482	M260T311	260.2121	310.5043	0.01%	2.669557
260	M561T1745	561.2218	1744.691	0.02%	2.686783
332	M383T79	383.0698	78.90242	0.02%	1.6542
1	M362T82	362.1757	81.67067	0.02%	2.707638
403	M348T127	348.0017	127.0846	0.03%	2.795219
384	M138T125	137.994	124.8874	0.03%	2.78301
348	M574T745	574.2583	744.728	0.03%	2.794578
237	M121T125	121.1184	124.5352	0.03%	2.181167
515	M272T372	272.2296	372.1095	0.04%	2.183961
370	M267T495	267.084	494.9564	0.04%	2.075155
344	M407T319	407.144	318.7294	0.04%	2.830891
464	M166T319	166.0747	318.9884	0.05%	2.20519
500	M300T372	300.1879	372.0058	0.05%	2.683233
396	M330T356	330.215	356.1513	0.06%	2.852187
321	M381T79	381.0621	78.96954	0.07%	1.436071
421	M374T470	374.1956	469.8593	0.08%	2.263457
272	M196T442	196.2525	442.3726	0.08%	3.520736
498	M761T1382	761.2651	1381.521	0.10%	2.268186
391	M373T470	373.1773	469.9358	0.10%	2.117132

B

	name	mzmed	rtmed		to wt
482	M260T311	260.2121	310.5043	11.81%	1.279941
260	M561T1745	561.2218	1744.691	63.29%	1.09211
332	M383T79	383.0698	78.90242	73.79%	1.039865
1	M362T82	362.1757	81.67067	5.41%	1.851835
403	M348T127	348.0017	127.0846	4.97%	1.533894
384	M138T125	137.994	124.8874	4.09%	1.447418
348	M574T745	574.2583	744.728	0.45%	1.728513
237	M121T125	121.1184	124.5352	2.70%	1.365142
515	M272T372	272.2296	372.1095	56.33%	1.066226
370	M267T495	267.084	494.9564	16.15%	1.230913
344	M407T319	407.144	318.7294	35.39%	1.128368
464	M166T319	166.0747	318.9884	32.34%	1.144431
500	M300T372	300.1879	372.0058	30.15%	1.148921
396	M330T356	330.215	356.1513	16.59%	1.198511
321	M381T79	381.0621	78.96954	51.53%	1.034696
421	M374T470	374.1956	469.8593	39.80%	1.085955

	name	mzmed	rtmed		to wt
286	M860T746	860.3661	746.0644	0.00%	2.658768
442	M476T1222	476.4536	1221.523	0.08%	1.302661
348	M574T745	574.2583	744.728	0.45%	1.728513
279	M287T745	287.2158	744.728	1.54%	1.320806
210	M288T745	288.2265	744.728	1.75%	1.338514
513	M213T161	212.76	160.9843	2.09%	1.461652
506	M182T1221	182.2089	1221.209	2.31%	1.333045
389	M467T594	467.1066	593.7703	2.49%	1.671956
237	M121T125	121.1184	124.5352	2.70%	1.365142
353	M554T881	554.1025	881.0401	2.81%	1.424482
216	M361T862	361.2444	861.6955	2.91%	1.316999
200	M378T848	378.1745	848.0364	2.92%	1.356558
423	M273T631	273.2031	630.5597	3.44%	1.288553
415	M475T1221	475.4504	1221.232	3.50%	1.104346
384	M138T125	137.994	124.8874	4.09%	1.447418
269	M573T745	573.2412	744.9487	4.55%	1.48676
403	M348T127	348.0017	127.0846	4.97%	1.533894
57	M881T746	881.4137	745.9902	5.24%	2.01105
198	M859T745	859.3297	745.1438	5.27%	1.915865

Figure 5.12. Summary of the untargeted analysis of *PpMYB4-KO* and *PpMYB5-KO*

The tables are sorted by probability P-value from the ttest in *PpMYB4-KO* (table A) and *PpMYB5-KO* (table B).The retention time (seconds) and mass are in columns rtmed and mzmed. For each genotype, a two-tailed t-test probability has been calculated meaning that each peak is the same as the corresponding peak in the wild-type, assuming different variances. Fold-change compared to wild-type, and an indication of whether the analyte went up or down is also given. Yellow highlight represents best P-values, green indicates P-values < 5%.

than for the differences between *PpMYB4* knockouts compared to wild type (Fig. 5.12B). On average the levels of these compounds were reduced by about 1.3 to 1.4 fold. Overall, these data suggested that both *PpMYB4-KO* and *PpMYB5-KO* have effects on phenylpropanoids in moss, but that they have different effects. *PpMYB5-KO* mutants seemed to have less pronounced effects, which also fits with it being harder to separate from Wt in the PCA analysis.

Mass spectra of peaks apparently differing in t-test results were analysed further by comparing extracted ion chromatograms. One unidentified compound in the *PpMYB4-KO* compared to wild type, which had a mass of 260 with a retention time (RT) of 5.17 min, had slightly increased level in the control when compared to *PpMYB4-KO* (Fig. 5.13A). Another chemical whose spectrum contained masses 561 and 539 (RT 29 min) did not show an obvious difference between the *PpMYB5-KO* mutant and the control but was identified as a possible flavonoid, biapigenin.

In contrast, *PpMYB5-KO*, showed two interesting peaks with masses of 287 and 476 (RT of 12.4 min and 20.4 min respectively) which were decreased in the mutant compared to the wild-type (Fig. 5.13B, C). Curiously, the peak with a mass of 287 had a mass appropriate for kaempferol but this was not kaempferol because it fragmented differently to the library standard. Consistently, in a previous analysis with other samples, it was possible also to identify the 287 amu compound in the *PpMYB5-KO* mutant and it was found that it was quite similar in its spectrum to the 273 amu chemical detected in the same samples. Interestingly, 273 amu is an appropriate mass for a hydrogen ion adduct of either naringenin or any isomer such as naringenin chalcone (Fig. 5.13D). This particular peak probably was not naringenin, as it fragmented differently to the library standard. However, in a library search, it showed an extremely weak similarity to chrysin, suggesting that it may be a molecule in the same general family of compounds as naringenin.

In Physcomitrella, only two genes encoding enzymes of phenylpropanoid metabolism have been identified, cloned and further characterized.

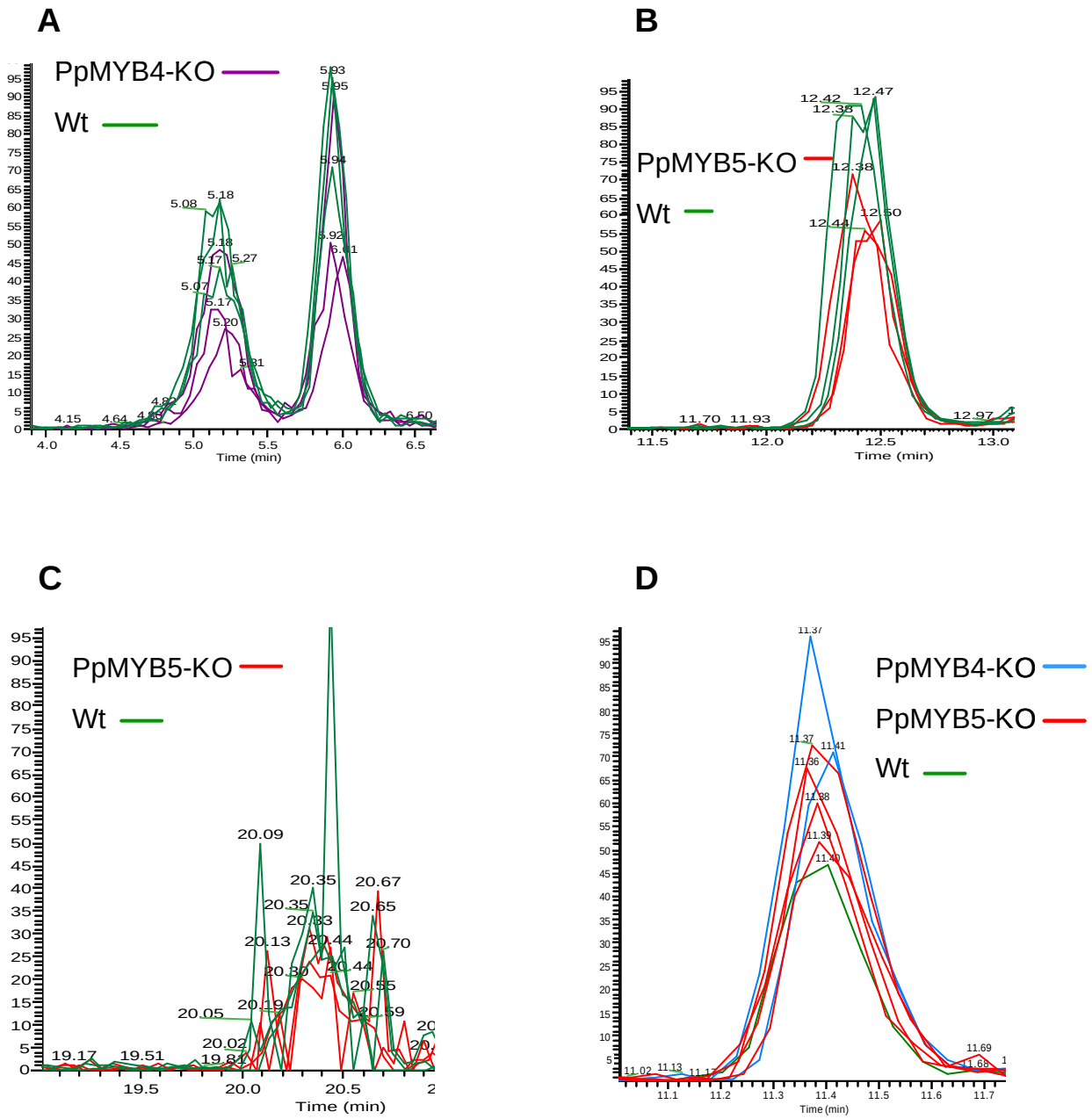


Figure 13. Relevant MS chromatograms peaks in Physcomitrella knockouts

Colour code of each line (genotype) represented in each graph. (A) Peak 260 amu (5.17 min) of *PpMYB4-KO* replicates compared to the control. (B) Compound of 287 amu (12.4 min) similar to kaempferol and (C) Peak 476 amu (20.4 min) of *PpMYB5-KO* replicates compared to the wild-type. (D) 273 amu molecule resembling the phenylpropanoid naringenin in the different knockouts and the wt.

They are the chalcone synthase (Jiang *et al*., 2006) that catalyzes the first step of flavonoid biosynthesis and the 4-coumarate:CoA ligase (4CL), controlling the divergence point from general phenylpropanoid metabolism to several branch pathways (Silber *et al*., 2008). As changes in the phenylpropanoids were suggested when *PpMYB4-KO* and *PpMYB5-KO* mutants were compared to wild-type, I decided to test whether either CHS or 4CL were affected in their expression in the moss knockout mutants (although both enzymes are encoded by multigene families and analysing one member may not be very informative). Oligonucleotides for each single gene encoding putative enzymes already described (CHS13a and 4CL2), were used for the expression analysis (Fig. 5.14A). In general, the expression of the CHS13a gene was lower in all the samples compared to the expression of 4CL2. No differences were found in the transcription of either gene in the wild-type compared to the *PpMYB4* and *PpMYB5* knockouts.

To identify other possible genes of the phenylpropanoid biosynthetic pathway in Physcomitrella, a BLASTP search of the *P. patens* genome sequence in the JGI genome annotation database was performed. Protein sequences of some genes (encoding PAL1, C4H, 4CL3, C3H, CHS, CHI and F3H) involved in phenylpropanoid metabolism in *Arabidopsis thaliana* (TAIR) were used as queries. Hits for every enzyme were found in the moss genome except for CHI (At5g05270.1) and although transcript analysis was not carried out, these genes could be also targets of PpMYB4 and PpMYB5 (To avoid a large list of genes, Appendix 2 only shows the best hit in Physcomitrella for every enzyme).

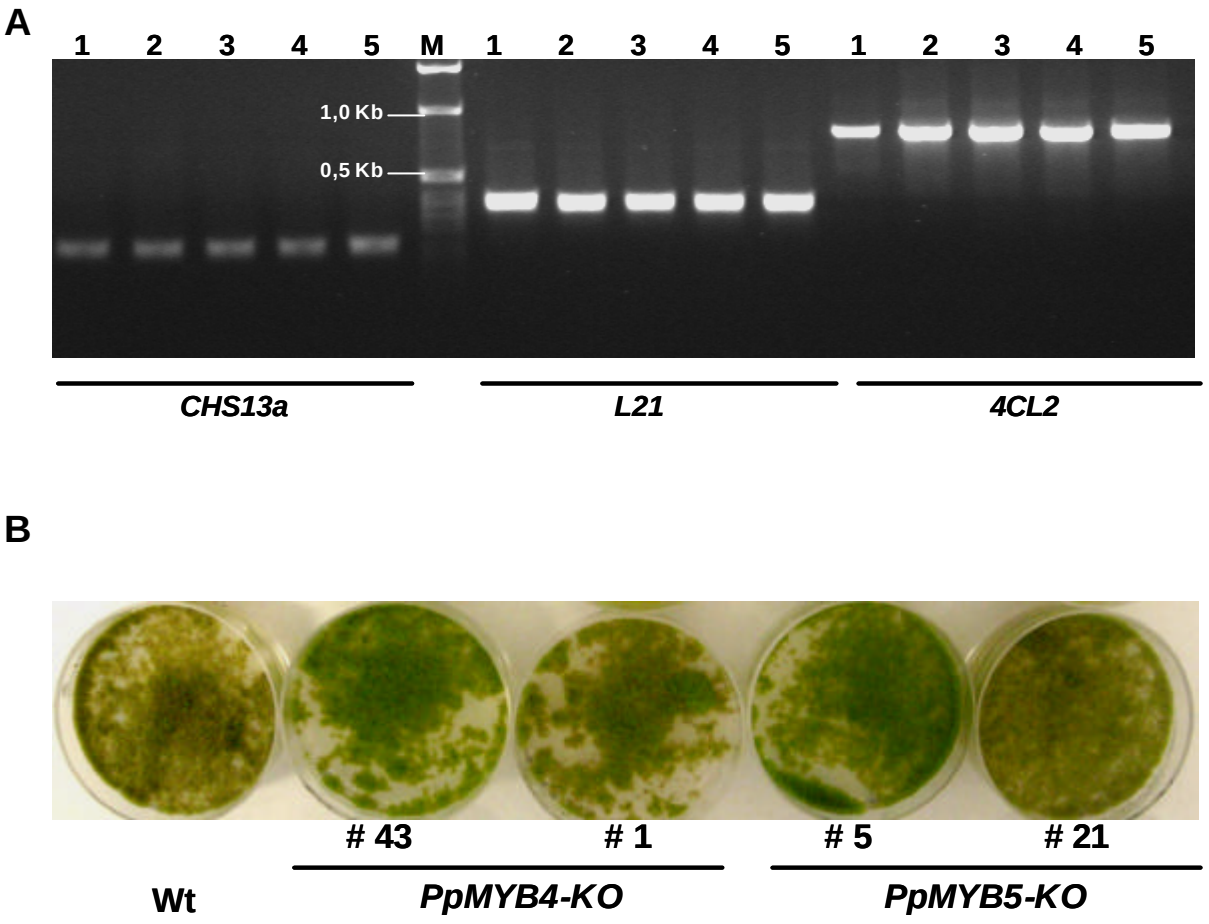


Figure 5.14. (A) RT-PCR of putative CHS and 4CL2 genes in 7-days old tissue from *PpMYBs* mutants and wt. Constitutive ribosomal protein L21 was used as control. 1, wt; lanes 2 and 3, *PpMYB4* independent KO lines #1, #43 respectively; lanes 4 and 5, *PpMYB5* independent KO lines #5 and #21. (B) Wt, *PpMYB4* and *PpMYB5* mutant lines 15 days after UV-B treatment. (C) Replicate of the UV-B treatment showing more clearly the resistance of the knockouts to the radiation.

5.3.4. UV-B treatment of *PpMYB4* and *PpMYB5* knockouts

Ultraviolet light is known to have effects on the levels of compounds of the phenylpropanoid pathway of plants, and flavonoids and hydroxycinnamates act as protectants from the damaging effects of UV-B radiation (Arvind and Jitendra, 1997).

The metabolic profile discussed above indicates variations in levels of compounds that could be phenolics different from those found in Arabidopsis. To test if these compounds could have a role as UV-B protectants, I tested the response of *PpMYB4 and PpMYB5* knockout isolates to UV-B light.

One week old moss cultures of two independent isolates of each knockout mutant and a wild-type were treated with UV-B radiation (as described above in Materials and Methods). Following exposure, moss cultures were grown on for two weeks more, prior to scoring and photography. Surprisingly, the *PpMYB4 and PpMYB5* knockouts showed less damage than the wild-type (Fig. 5.14B). Whereas the mutants were greener and started to develop green leafy shoots, the wild-type was almost dead with protonema of pale green-brown colour. The experiment was repeated using just one isolate of each knockout mutant and one wild type culture and the increased resistance of both mutants to UV-B irradiation was confirmed (Fig.5.14C). These data suggested that the *PpMYB4* and *PpMYB5* knockouts may produce more phenylpropanoid derivatives that could serve as sunscreens by absorbing UV-B irradiation and so have enhanced resistance to short, intense UV-B irradiation.

5.3.5. Overexpression of *PpMYB4* and *PpMYB5* in Arabidopsis

I examined the effects of expression of the moss genes *PpMYB4* and *PpMYB5* on the content of phenolics in Arabidopsis, by generating transgenic plant lines carrying *PpMYB4* and *PpMYB5* cDNAs (pJAM1502) under the control of the strong CaMV 35S promoter. Two homozygous independent lines for each transgene were retrieved by kanamycin selection and characterized further.

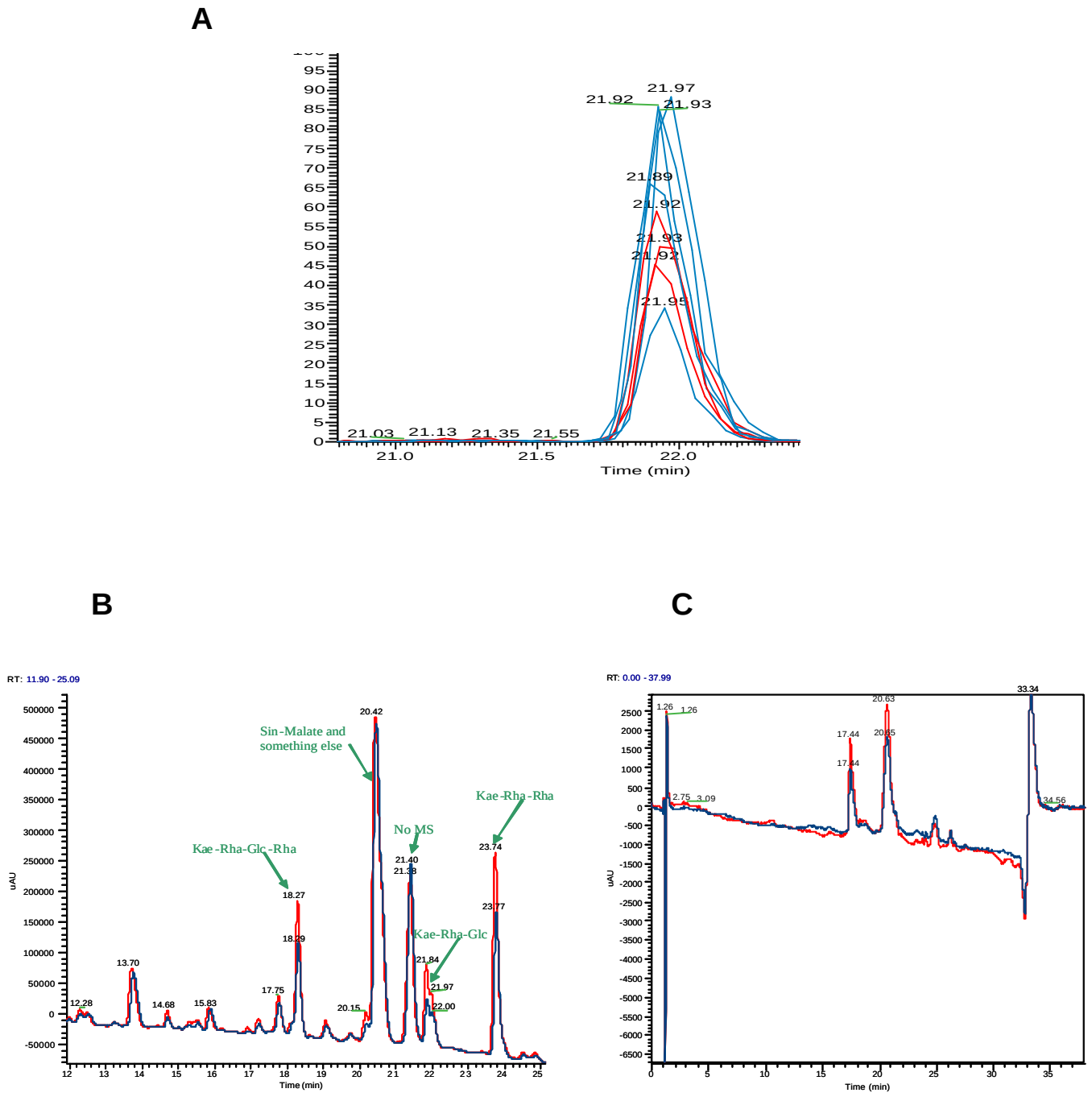


Figure 5.15. Phenolic profile in Arabidopsis Col-0 overexpressing *PpMYB4* and *PpMYB5*

(A) Peak 434 amu corresponding to kaempferol-rhamnosyl-glucoside in Arabidopsis overexpressing *PpMYB4* (transgenic line in red). (B-C) Arabidopsis phenylpropanoid profile overexpressing *PpMYB5* gene (transgenic line in blue). Mutant line has lower kaempferol derivatives (B) and anthocyanins (C) compared to the wild-type.

Soluble phenolics of whole Arabidopsis plantlets (including roots) at the bolting stage were analyzed by liquid chromatography-mass spectrometry (LC-MS). Only the statically significant differences are discussed in this section.

For Arabidopsis *PpMYB4* overexpressing lines, the analysis of two independent lines revealed that there were two significant common peaks that appeared in both lines and therefore seemed most relevant. These are mass 240 at 534 sec (8.9 min) and mass 434 at 1317 sec (22 min). The extracted ion chromatograms for 240 amu did not show very large differences, but levels appeared slightly higher in transgenic plants (data not shown). It was not possible to identify this compound due to the lack of a similar fragmentation pattern in the database. The mass 434 corresponds to kaempferol-rhamnosyl-glucoside and it looked as though the transgenic plants contained less of this compound (Fig.5.15A).

For *PpMYB5* overexpressing lines there were also peaks which changed significantly in their levels compared to wild type plants. Overlay of UV chromatograms at 214 nm, in the region of the chromatogram where the phenolics tend to elute (Fig.5.15B), there were peaks that were different and peaks that were the same in PpMYB5 overexpressing plants compared to wild type controls. It was noticeable that all the kaempferol glycosides were lower in *PpMYB5* KO lines although the other peaks (unidentified peaks and sinapoyl malate), were similar to the wild type. In an overlay of chromatograms at 520 nm, the two peaks at 17.44 and 20.63 min were weakly associated with an ion of 1343 amu, which is an anthocyanin (Fig.5.15C). This compound decreased in the mutant.

UV-B light exposure was applied to Arabidopsis seedlings overexpressing *PpMYB4* and *PpMYB5,* but visible changes in growth or development were not observed compared to wild type (pictures not shown).

5.4. Discussion

The PpMYB4 and PpMYB5 proteins are related structurally to AtMYB6 and AtMYB8 in Arabidopsis. Phylogenetic analysis shows that these MYB proteins are the only ones from *P. patens* that align within the clade of R2R3 MYB proteins that regulate phenylpropanoid metabolism (Arabidopsis R2R3 MYB subgroups 4, 5, 6 and 7; Stracke *et al*., 2001). I report them here as likely structural and putative functional homologues of the Arabidopsis MYB transcription factors characterized in previous chapters of this work. Since starting my PhD, the moss genome sequence has been published and the alignment of MYB genes in moss against those in Arabidopsis looks somewhat different (new close paralogues to PpMYB4 and PpMYB5, XP_001783614.1 and XP_001752936.1, have been identified). It is very hard to say which moss genes are orthologous to AtMYB6 and AtMYB8. However, since PpMYB4 and PpMYB5 are the only moss MYBs that align within the phenylpropanoid clades of MYB proteins in Arabidopsis, it seems fair to claim that *PpMYB4* and *PpMYB5* are descended from a common ancestor with *AtMYB6* and *AtMYB8*. This alignment also means that the functions of PpMYB4 and PpMYB5 are most likely (of all the moss MYB proteins) to be associated with the regulation of phenylpropanoid metabolism.

Bryophytes and other lower plants are not well studied, and up until the publication of the genome sequence of *P. patens* (Rensing *et al*., 2008), only two MYB-like proteins, Pp1 and Pp2, had been characterized from the moss *Physcomitrella patens* (Leech *et al*., 1993).

Sequence comparison showed that *PpMYB4* and *PpMYB5* contain the conserved MYB domain indicating that *R2R3 MYB* genes arose early during the evolution of land plants. In fact, it is thought that the MYB domain first originated over 1 billion years ago, shortly after the divergence of eubacteria and eukaryotes (Lipsick, 1996). All MYB-type genes are probably homologous genes which originated from a single common ancestral gene; they were probably generated by gene duplications and sequence diversification (Rosinski and Atchley, 1998; Braun and Grotewold, 1999; Harald *et al*., 2000;

Rensing *et al*., 2008). That implies that *PpMYB* genes reported here could share a common ancestor with some of the well-characterized *MYB* genes from flowering plants and be related to them functionally as well. In our phylogenetic analysis comparing the Arabidopsis MYB family (Kranz *et al*., 1998; Stracke *et al*., 2001) with PpMYB4 and PpMYB5, it was possible to group them close to the subgroups containing flavonoid-related and phenylpropanoid-related regulatory proteins. In fact, *PpMYB4* and its later-found paralog (XP_001783614.1), encode the GIDP motif present in most phenylpropanoid MYBs suggesting that higher plant R2R3 MYB subgroup 4 was derived from an ancestor shared with these two moss genes. Despite the homology amongst all the *MYB* proteins described in this study, within their DNA binding domains, they differed highly in their C-terminal domains and in the number and length of exons and introns. My data indicate that these proteins not only share structural similarities but share some aspects of their functionality as well.

PpMYB4 and *PpMYB5* genes are expressed in protonema and in gametophores, the two main tissue types of moss, suggesting that the encoded MYB proteins could be required in both tissues.

The pioneering ancestor of land plants that conquered terrestrial habitats around 500 million years evolved from the Charophyceae that faced a harsh environment characterised by stresses including desiccation, UV radiation and potential microbial attack. The evolution of the phenylpropanoid pathway is thought to have been critical to the colonization of land as this must have been the point at which many compounds essential for the survival of plants on land, such as pigments, lignins and phytoalexins evolved (Emiliani *et al*., 2009). As major phenylpropanoid-derived compounds, flavonoids are present in higher plants as well as in mosses (Stafford, 1991). In earlier chapters, I presented evidence that *AtMYB6* and *AtMYB8* might have roles as transcriptional regulators of phenylpropanoid metabolism, specifically flavonoid metabolism. To analyze the role of PpMYB4 and PpMYB5 in regulating this metabolic pathway and to compare these data to those previously obtained for Arabidopsis, mutants of the *PpMYB4* and *PpMYB5* genes were prepared

(Didier and Jean-Pierre, 1997). By disruption of these *PpMYBs* genes it was hoped to remove their activity, with the expectation that PpMYB knockouts would be altered in the phenylpropanoid metabolism. This expectation was tempered upon release of the genome sequence of *P. patens* which revealed the existence of very similar, paralogous genes to *PpMYB4* and *PpMYB5* (XP_001752936.1 and XP_001783614.1, respectively) in the *P. patens* genome. However, I continued to study the *PpMYB4* and *PpMYB5* knockouts (since there were no EST sequences of XP_001783614.1 or XP_001752936.1 suggesting that they might not be expressed genes) in the hope that they would exhibit phenotypes that would provide clues to their functionalities, despite the possibility of this being masked by functional redundancy with the paralogues. The metabolic profiles for phenylpropanoid compounds of the knockout mutants compared to the wild type revealed differences in phenolic levels between the two mutant types and compared to the wild-type. However, the differences were small and difficult to detect probably because the disruption of a single MYB gene may not necessarily cause a drastic change of the phenotype as a result of functional redundancy with the paralogues, as suggested for MYB transcription factors in Arabidopsis (Jin and Martin, 1999; Stracke *et al*., 2001). The MYB family has not been extensively studied in Physcomitrella but, although it is not as large a family as in Arabidopsis, it has a total of 48 genes encoding R2R3 MYB proteins according to the Plant Transcription Factor database (http://plntfdb.bio.uni-potsdam.de/v3.0/) and data-mining by Paul Bailey.

Lower plants seem to differ from higher plants by producing a relatively simple range of UV-B absorbing (phenolic) compounds. Unfortunately, phenylpropanoid compounds in lower plants are almost completely uncharacterised. Few have been identified and most of the secondary metabolites produced are not yet characterized. Identification of phenolic peaks from LC/MS chromatograms of tissue extracts proved very difficult due to the poor literature reporting the fragmentation patterns of phenylpropanoid compounds. Several flavonoids have been reported in bryophytes including

flavone C- and O-glycosides, biflavones, isoflavones, aurones and anthocyanins (Melchert and Alston, 1965; Anhut *et al*., 1984; Geiger and Markham, 1992; Brinkmeier *et al*., 1999). Despite difficulties in identifying individual compounds, it was possible to detect general changes in the putative phenylpropanoid profile in the *PpMYB* knockouts compared to wild type. It was possible to identify few peaks corresponding to phenolics that changed in the mutants. The *PpMYB4* knockout showed a general decrease in the majority of these compounds although one or two were increased significantly in the mutant. For the *PpMYB5* knockouts there were small decreases in most peaks compared to wild type. For this mutant it was possible to identify changes in peaks structurally similar to flavonoids. Those peaks were slightly decreased in the *PpMYB5* knockouts and from their mass characteristics, resembled the flavonoids, kaempeferol and chrysin. Kaempferol (together with quecertin) are the major flavonoids found in *Arabidopsis thaliana*. Kaempferol is a flavonol (Veit and Pauli, 1999; Kerhoas *et al*., 2006) whereas chrysin is a flavone.

It is now generally accepted that the development of phenylpropanoid metabolism has played a major role in the evolution of land plants. Phenolics in lower plants may have acted as UV filters in the progenitors of vascular plants and supported the transition of plant life from aquatic environments, (where UV-B irradiance is low) to terrestrial environments (with high solar UV-B irradiances). It has been assumed that filtering UV-B irradiation by phenolics provided land plants with internal screens against damaging solar UV-B irradiation and facilitated the evolution of terrestrial plant life (Stafford, 1991; Waters, 2003). To assess the effects of altered phenylpropanoid levels in *PpMYB* knockout mutants compared to the control, a UV-B light treatment was applied. It was not possible to correlate the content of phenolics overall with resistance to UV-B irradiation because I could not identify which of all the phenylpropanoids present in moss were most closely associated with tolerance of UV-B and not all phenolic compounds act as UV light protectants. Results suggested that *PpMYB4* and *PpMYB5* may be acting as negative regulators of UV-B protection mechanisms, since knockouts of these genes were more

resistant to the UV-B irradiation. Possibly, both genes are repressing a key regulatory point affecting one or several branches of phenylpropanoid metabolism that controls the accumulation of protective compounds in moss. The similar effects of the *PpMYB4* and *PpMYB5* genes compared to those of *AtMYB6* and *AtMYB8* of Arabidopsis on UV-B protection led me to think about a shared role of both MYB transcription factors which could work together to coordinate the production of protective compounds, possibly phenylpropanoids.

To determine whether *PpMYB4* and *PpMYB5* could act similarly to the MYB proteins of higher plants, their function was tested by overexpression in Arabidopsis. Evaluation of the phenolic content in transformed Arabidopsis plants showed clear alterations especially in the Arabidopsis plants overexpressing *PpMYB5*. In this case, mutant plants contained less kaempferol derivatives as well as anthocyanins, whereas the phenolic profile for *PpMYB4* overexpressers was not very different to the control. Both metabolic profiles resembled those obtained in the Arabidopsis overexpression lines with *AtMYB6* and *AtMYB8* described in Chapter 4, especially, the *AtMYB8* overexpressing line which also had reduced phenylpropanoid levels. This suggested that the PpMYB TFs could act as repressors of phenylpropanoid biosynthesis and that their functions have been conserved during land plant evolution. However, irradiation with UV-B light showed no significant differences in growth or survival of the lines overexpressing *PpMYB4* or *PpMYB5* compared to the Arabidopsis wild-type plants. This result was consistent with plants expressing *PpMYB4* since phenylpropanoid levels were almost identical in these and wild type plants, but it does not explain the absence of UV-B sensitivity of the Arabidopsis *PpMYB5* transgenic plants whose low flavonoid levels were predicted to result in an increased sensitivity to UV-B irradiation. One explanation could be that sinapoyl malate levels (the major UV light sunscreen in Arabidopsis) were the same in *PpMYB5* overexpressing lines as in the wild-type. In addition, the analysis of phenolics was performed on older plants where the content in these chemicals is possibly different to the phenylpropanoid content in young seedlings, the stage at which the UV-B irradiance treatments

were carried out. A future approach in young seedlings could give more consistent results. Taken together my data suggest that the common ancestor to *PpMYB4*, *PpMYB5* and the R2R3MYBs of subgroups 4, 5, 6 and 7 in Angiosperms, may originally have played an important role in UV protection through regulating phenylpropanoid metabolism as plants emerged onto the land.

The lack of information available about phenylpropanoid metabolism in bryophytes makes it difficult to track the biosynthetic steps (genes) that PpMYB4 and PpMYB5 might regulate. A BLASTP search using the publicly accessible *P. patens* genome sequence data (JGI database), revealed the presence of the main biosynthetic enzymes of the phenylpropanoid pathway in moss (Appendix 2). I identified genes encoding proteins with significant protein homology to the enzymes of general phenylpropanoid metabolism (PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; and 4CL, 4-coumarate: CoA ligase), flavonoid metabolism (CHS, chalcone synthase; and F3H, flavonoid 3 hydroxylase; and hydroxycinnamate metabolism (C3'H, 4-coumarate 3-hydroxylase) in *P. patens*. The presence of these genes (except CHI), suggests that *P. patens* is able to synthesise simple flavonoids and hydroxycinnamates. In Physcomitrella, each enzyme from the main branch of the phenylpropanoid pathway (PAL1, C4H, 4CL3, C3H, CHS, and F3H) seems to be encoded by a multigene family. I attempted to decipher if the two characterized genes encoding enzymes in this metabolic pathway in Physcomitrella, PpCHS and Pp4CL (Jiang *et al*., 2006; Silber *et al*., 2008) were regulated by PpMYB4 or PpMYB5, by assaying their expression in wild type and the knock out mutants. I found no evidence for either mutant affecting the levels of CHS transcripts, but the transcripts of 4CL2 were more abundant in both the *PpMYB4* and *PpMYB5* knockouts compared to the wild type indicating that both MYBs may be potential negative regulators of 4CL2 expression. However, these data should be viewed as preliminary because one can not really draw quantitative conclusions about transcript levels from the RT-PCR

due to problems affecting the quantitative basis of such data like saturating levels of DNA *etc*.

Most of the moss MYBs are represented by two very closely related genes/proteins as with PpMYB4 and PpMYB5. A large-scale recent duplication of the *Physcomitrella patens* genome, supports these observations (Rensing *et al*., 2008). From a practical point of view, a recent genome duplication could imply that the knockouts I generated might have no strong phenotype because of redundancy with their paralogous genes, in both cases. However, some phenotypic effects on putative phenylpropanoid levels were observed for the knockouts, so complete functional redundancy between the gene pairs for *PpMYB4* and *PpMYB5* therefore seems unlikely. However, the phenotypes might be much stronger if double knockouts of the duplicate gene pairs could be achieved.

The information presented in this chapter suggests the involvement of *PpMYB4* and *PpMYB5* in the same or similar functions as the regulatory genes *AtMYB6* and *AtMYB8* from Arabidopsis showing that, despite the 200-250 Mya that separate mosses from flowering plants, these MYB transcription factors may have been functionally conserved and may therefore have functions essential for long-term plant survival.

CONCLUSIONS

In this study, I have characterized AtMYB6 and AtMYB8 as two MYB transcription factors probably involved in regulating phenylpropanoid metabolism.

I have demonstrated that both genes are nuclear-localized supporting their identity as transcription factors. I have shown that they are widely and similarly expressed in Arabidopsis suggesting that they are playing overlapping roles in the plant and therefore may be functionally redundant. This interpretation is supported by experimental data obtained as well as by the structural-functional similarity of these genes, between themselves and with other members of their same R2R3MYB subfamily. Absence of visual phenotype in the mutant plants analyzed for both genes is in agreement with the hypothesis of functional redundancy between them and/or other members of the R2R3MYB subgroup IV family.

Metabolic analyses revealed that transgenic Arabidopsis plants overxpressing/misexpressing *AtMYB6* or *AtMYB8* had altered phenylpropanoid contents. Overexpression of these genes in a heterologous system, tobacco, showed similar results, supporting the idea of these proteins as transcriptional regulators of the phenylpropanoid pathway and therefore highlighting again their possible functional redundancy, since both genes seems to participate in the controlling the same metabolism.

In addition, I have characterised *PpMYB4* and *PpMYB5* genes from the lower plant *Physcomitrella patens* as putatively derived from a common ancestor with *AtMYB6* and *AtMYB8*. The phylogenetic positions of PpMYB4 and PpMYB5, the analysis of phenylpropanoid compounds in the *PpMYB4* and *PpMYB5* knockout mutants as well as their overexpression in Arabidopsis, suggested that these moss proteins have similar functions to their homologues in Arabidopsis AtMYB6 and AtMYB8. These results indicate that their roles in the phenylpropanoid pathway could have been conserved during plant evolution since mosses are descended from the first plants to colonize the land.

My work has assigned functions, for the first time, to genes that had not previously been correctly characterized and provided new insights into the evolution of phenylpropanoid metabolism, a topic which remains largely undescribed. Very little information about phenylpropanoids is available for mosses and characterization of these compounds could shed much more light on the evolution of one of the most important metabolic pathways in plants. Further explanations are necessary to decipher the evolution of this pathway and lower plants could provide important clues as to how this happened.

BIBLIOGRAPHY

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signaling. Plant Cell **15:** 63-78

Aeschbacher RA, Hauser MT, Feldmann KA, Benfey PN (1995) The SABRE gene is required for normal cell expansion in Arabidopsis. Genes & Development **9:** 330-340

Agarwal M, Hao Y, Kapoor A, Dong C-H, Fujii H, Zheng X, Zhu J-K (2006) A R2R3 Type MYB Transcription Factor Involved in the Cold Regulation of CBF Genes and in Acquired Freezing Tolerance. Journal of Biological Chemistry **281:** 37636-37645

Albert S, Delseny M, Devic M (1997) BANYULS, a novel negative regulator of flavonoid biosynthesis in the Arabidopsis seed coat. The Plant Journal **11:** 289-299

Ana C-D, Steven P, Caroline S, Merryn C, Michael B (2003) Reduced cellulose synthesis invokes lignification and defense responses in Arabidopsis thaliana. The Plant Journal **34:** 351-362

Anhut S, Dietmar Zinsmeister H, Mues R, Barz W, Mackenbrock K, Köster J, Markham KR (1984) The first identification of isoflavones from a bryophyte. Phytochemistry **23:** 1073-1075

Arnon DI (1949) Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. Plant Physiol. **24:** 1-15

Arvind KB, Jitendra PK (1997) Mutants of Arabidopsis as Tools to Understand the Regulation of Phenylpropanoid Pathway and UVB Protection Mechanisms. Photochemistry and Photobiology **65:** 765-776

Asami T, Nakano T, Fujioka S (2005) Plant brassinosteroid hormones. Vitam Horm **72:** 479-504

Avila J, Nieto C, Canas L, Benito MJ, Paz-Ares J (1993) Petunia hybrida genes related to the maize regulatory C1 gene and to animal myb proto-oncogenes. Plant J **3:** 553-562

Bai F, Matsui T, Ohtani-Fujita N, Matsukawa Y, Ding Y, Sakai T (1998) Promoter activation and following induction of the p21/WAF1 gene by flavone is involved in G1 phase arrest in A549 lung adenocarcinoma cells. FEBS Letters **437:** 61-64

Bang WY, Kim SW, Jeong IS, Koiwa H, Bahk JD (2008) The C-terminal region (640-967) of Arabidopsis CPL1 interacts with the abiotic stress- and ABA-responsive transcription factors. Biochemical and Biophysical Research Communications **372:** 907-912

Baranowskij N, Frohberg C, Prat S, Willmitzer L (1994) A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. Embo J **13:** 5383-5392

Baucher M, Bernard-vailhé MA, Chabbert B, Besle J-M, Opsomer C, Van Montagu M, Botterman J (1999) Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (*Medicago sativa* L.) and the effect on lignin composition and digestibility. Plant Molecular Biology **39:** 437-447

Baumann K, Perez-Rodriguez M, Bradley D, Venail J, Bailey P, Jin H, Koes R, Roberts K, Martin C (2007) Control of cell and petal morphogenesis by R2R3 MYB transcription factors. Development **134:** 1691-1701

Ben Zvi MM, Florence NZ, Masci T, Ovadis M, Shklarman E, Ben-Meir H, Tzfira T, Dudareva N, Vainstein A (2008) Interlinking showy traits: co-engineering of scent and colour biosynthesis in flowers. Plant Biotechnology Journal **6:** 403-415
Bender J, Fink GR (1998) A Myb homologue, ATR1, activates tryptophan gene expression in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **95:** 5655-5660
Bentley R, Haslam E (1990) The Shikimate Pathway â€" A Metabolic Tree with Many Branche. Critical Reviews in Biochemistry and Molecular Biology **25:** 307-384
Berger MF, Bulyk ML (2009) Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. Nat. Protocols **4:** 393-411
Berger MF, Philippakis AA, Qureshi AM, He FS, Estep PW, Bulyk ML (2006) Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. Nat Biotech **24:** 1429-1435
Bishop GJ, Koncz C (2002) Brassinosteroids and Plant Steroid Hormone Signaling. Plant Cell **14:** S97-110
Booij-James IS, Dube SK, Jansen MAK, Edelman M, Mattoo AK (2000) Ultraviolet-B Radiation Impacts Light-Mediated Turnover of the Photosystem II Reaction Center Heterodimer in Arabidopsis Mutants Altered in Phenolic Metabolism. Plant Physiol. **124:** 1275-1284
Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C (2000) Activation Tagging Identifies a Conserved MYB Regulator of Phenylpropanoid Biosynthesis. Plant Cell **12:** 2383-2394
Braun EL, Grotewold E (1999) Newly discovered plant c-myb-like genes rewrite the evolution of the plant myb gene family. Plant Physiol **121:** 21-24
Briggs GC, Osmont KS, Shindo C, Sibout R, Hardtke CS (2006) Unequal genetic redundancies in Arabidopsis - a neglected phenomenon? Trends in Plant Science **11:** 492-498
Brinkmeier E, Geiger H, Zinsmeister HD (1999) Biflavonoids and 4,2'-epoxy-3-phenylcoumarins from the moss Mnium hornum. Phytochemistry **52:** 297-302
Broun P (2005) Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in Arabidopsis. Current Opinion in Plant Biology **8:** 272-279
Brown DE, Rashotte AM, Murphy AS, Normanly J, Tague BW, Peer WA, Taiz L, Muday GK (2001) Flavonoids Act as Negative Regulators of Auxin Transport in Vivo in Arabidopsis. Plant Physiol. **126:** 524-535
Buer CS, Djordjevic MA (2009) Architectural phenotypes in the transparent testa mutants of Arabidopsis thaliana. J. Exp. Bot. **60:** 751-763
Buer CS, Muday GK, Djordjevic MA (2007) Flavonoids Are Differentially Taken Up and Transported Long Distances in Arabidopsis. Plant Physiol. **145:** 478-490
Burke GL, Vitolins MZ, Bland D (2000) Soybean Isoflavones as an Alternative to Traditional Hormone Replacement Therapy: Are We There Yet? J. Nutr. **130:** 664-
Butelli E, Titta L, Giorgio M, Mock H-P, Matros A, Peterek S, Schijlen EGWM, Hall RD, Bovy AG, Luo J, Martin C (2008) Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nat Biotech **26:** 1301-1308

Cano-Delgado AI, Metzlaff K, Bevan MW (2000) The eli1 mutation reveals a link between cell expansion and secondary cell wall formation in Arabidopsis thaliana. Development **127:** 3395-3405
Castle LA, Meinke DW (1994) A FUSCA Gene of Arabidopsis Encodes a Novel Protein Essential for Plant Development. Plant Cell **6:** 25-41
Cedroni ML, Cronn RC, Adams KL, Wilkins TA, Wendel JF (2003) Evolution and expression of MYB genes in diploid and polyploid cotton. Plant Mol Biol **51:** 313-325
Chapple CCS, Vogt T, Ellis BE, Somerville CR (1992) An Arabidopsis Mutant Defective in the General Phenylpropanoid Pathway. Plant Cell **4:** 1413-1424
Chen B-J, Wang Y, Hu Y-L, Wu Q, Lin Z-P (2005) Cloning and characterization of a drought-inducible MYB gene from Boea crassifolia. Plant Science **168:** 493-500
Cheng H, Song S, Xiao L, Soo HM, Cheng Z, Xie D, Peng J (2009) Gibberellin Acts through Jasmonate to Control the Expression of MYB21, MYB24, and MYB57 to Promote Stamen Filament Growth in Arabidopsis. PLoS Genet **5:** e1000440
Cheong YH, Chang H-S, Gupta R, Wang X, Zhu T, Luan S (2002) Transcriptional Profiling Reveals Novel Interactions between Wounding, Pathogen, Abiotic Stress, and Hormonal Responses in Arabidopsis. Plant Physiol. **129:** 661-677
Choi B, Roh K (2003) UV-B radiation affects chlorophyll and activation of rubisco by rubisco activase inCanavalia ensiformis L. leaves. Journal of Plant Biology **46:** 117-121
Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16:** 735-743
Clouse SD, Sasse JM (1998) Brassinosteroids: Essential Regulators of Plant Growth and Development. Annual Review of Plant Physiology and Plant Molecular Biology **49:** 427-451
Collard BCY, Das A, Virk PS, Mackill DJ (2007) Evaluation of 'quick and dirty' DNA extraction methods for marker-assisted selection in rice (*Oryza sativa* L.). Plant Breeding **126:** 47-50
Cominelli E, Galbiati M, Vavasseur A, Conti L, Sala T, Vuylsteke M, Leonhardt N, Dellaporta SL, Tonelli C (2005) A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. Curr Biol **15:** 1196-1200
Cominelli E, Tonelli C A new role for plant R2R3-MYB transcription factors in cell cycle regulation. Cell Res **19:** 1231-1232
Cove D (2005) The moss *Physcomitrella patens*. Annual Review of Genetics **39:** 339-358
Dai X, Xu Y, Ma Q, Xu W, Wang T, Xue Y, Chong K (2007) Overexpression of an R1R2R3 MYB Gene, OsMYB3R-2, Increases Tolerance to Freezing, Drought, and Salt Stress in Transgenic Arabidopsis. Plant Physiol. **143:** 1739-1751
Dai Y, Rahmani M, Grant S (2003) Proteasome inhibitors potentiate leukemic cell apoptosis induced by the cyclin-dependent kinase inhibitor flavopiridol through a SAPK//JNK- and NF-[kappa]B-dependent process. Oncogene **22:** 7108-7122
de Vetten N, Quattrocchio F, Mol J, Koes R (1997) The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes & Development **11:** 1422-1434

Debeaujon I, Koornneef M (2000) Gibberellin Requirement for Arabidopsis Seed Germination Is Determined Both by Testa Characteristics and Embryonic Abscisic Acid. Plant Physiol. **122:** 415-424
Debeaujon I, Leon-Kloosterziel KM, Koornneef M (2000) Influence of the Testa on Seed Dormancy, Germination, and Longevity in Arabidopsis. Plant Physiol. **122:** 403-414
Denekamp M, Smeekens SC (2003) Integration of Wounding and Osmotic Stress Signals Determines the Expression of the AtMYB102 Transcription Factor Gene. Plant Physiol. **132:** 1415-1423
Dias AP, Braun EL, McMullen MD, Grotewold E (2003) Recently duplicated maize R2R3 Myb genes provide evidence for distinct mechanisms of evolutionary divergence after duplication. Plant Physiol **131:** 610-620
Didier GS, Jean-Pierre Z (1997) Efficient gene targeting in the moss *Physcomitrella patens*. The Plant Journal **11:** 1195-1206
Dixon RA, Lamb CJ, Masoud S, Sewalt VJH, Paiva NL (1996) Metabolic engineering: prospects for crop improvement through the genetic manipulation of phenylpropanoid biosynthesis and defense responses -- a review. Gene **179:** 61-71
Dixon RA, Paiva NL (1995) Stress-Induced Phenylpropanoid Metabolism. Plant Cell **7:** 1085-1097
Du H, Zhang L, Liu L, Tang X-F, Yang W-J, Wu Y-M, Huang Y-B, Tang Y-X (2009) Biochemical and molecular characterization of plant MYB transcription factor family. Biochemistry (Moscow) **74:** 1-11
Duprey SP, Boettiger D (1985) Developmental regulation of c-myb in normal myeloid progenitor cells. Proc Natl Acad Sci U S A **82:** 6937-6941
Elkind Y, Edwards R, Mavandad M, Hedrick SA, Ribak O, Dixon RA, Lamb CJ (1990) Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene. Proceedings of the National Academy of Sciences of the United States of America **87:** 9057-9061
Emiliani G, Fondi M, Fani R, Gribaldo S (2009) A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. Biology Direct **4:** 7
Ferrer JL, Austin MB, Stewart Jr C, Noel JP (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiology and Biochemistry **46:** 356-370
Filippos V, Emmanouil T, Carl D, Guenter V, Georg K, Nickolas P (2007) Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. Biotechnology Journal **2:** 1214-1234
Forkmann G, Martens S (2001) Metabolic engineering and applications of flavonoids. Current Opinion in Biotechnology **12:** 155-160
Fornalé S, Sonbol F-M, Maes T, Capellades M, Puigdomènech P, Rigau J, Caparrós-Ruiz D (2006) Down-regulation of the maize and Arabidopsis thaliana caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors. Plant Molecular Biology **62:** 809-823

Friml J (2003) Auxin transport -- shaping the plant. Current Opinion in Plant Biology **6:** 7-12
Fu X, Harberd NP (2003) Auxin promotes Arabidopsis root growth by modulating gibberellin response. Nature **421:** 740-743
Geiger H, Markham KR (1992) Campylopusaurone, an auronoflavanone biflavonoid from the mosses *Campylopus clavatus* and *Campylopus holomitrium*. Phytochemistry **31:** 4325-4328
Gigolashvili T, Berger B, Flügge U-I (2009) Specific and coordinated control of indolic and aliphatic glucosinolate biosynthesis by R2R3-MYB transcription factors in Arabidopsis thaliana. Phytochemistry Reviews **8:** 3-13
Gómez-Gómez L, Boller T (2000) FLS2: An LRR Receptor-like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis. Molecular Cell **5:** 1003-1011
Gubler F, Kalla R, Roberts JK, Jacobsen JV (1995) Gibberellin-Regulated Expression of a myb Gene in Barley Aleurone Cells: Evidence for Myb Transactivation of a High-pl [alpha]-Amylase Gene Promoter. Plant Cell **7:** 1879-1891
Guo A-Y, Chen X, Gao G, Zhang H, Zhu Q-H, Liu X-C, Zhong Y-F, Gu X, He K, Luo J (2008) PlantTFDB: a comprehensive plant transcription factor database. Nucl. Acids Res. **36:** D966-969
Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J (2005) DATF: a database of Arabidopsis transcription factors. Bioinformatics **21:** 2568-2569
Halls C, Yu O (2008) Potential for metabolic engineering of resveratrol biosynthesis. Trends in Biotechnology **26:** 77-81
Hammarstrom M, Hellgren N, Van den Berg S, Berglund H, Hard T (2002) Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli. Protein Science **11:** 313-321
Harald K, Kai S, Bernd W (2000) c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. The Plant Journal **21:** 231-235
Harashima S, Takano H, Ono K, Takio S (2004) Chalcone synthase-like gene in the liverwort, Marchantia paleacea var. diptera. Plant Cell Reports **23:** 167-173
Hemm MR, Herrmann KM, Chapple C (2001) AtMYB4: a transcription factor general in the battle against UV. Trends in Plant Science **6:** 135-136
Hirano SS, Upper CD (2000) Bacteria in the Leaf Ecosystem with Emphasis on Pseudomonas syringae-a Pathogen, Ice Nucleus, and Epiphyte. Microbiol. Mol. Biol. Rev. **64:** 624-653
Hohe A, Egener T, Lucht JM, Holtorf H, Reinhard C, Schween G, Reski R (2004) An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, Physcomitrella patens. Curr Genet **44:** 339-347
Hulskamp M (2004) Plant trichomes: a model for cell differentiation. Nat Rev Mol Cell Biol **5:** 471-480
Ishitani M, Xiong L, Stevenson B, Zhu JK (1997) Genetic analysis of osmotic and cold stress signal transduction in Arabidopsis: interactions and convergence of abscisic acid-dependent and abscisic acid-independent pathways. Plant Cell **9:** 1935-1949

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung GH, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, SchuSzler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lucking R, Budel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JS, Vilgalys R (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature **443:** 818 - 822

Jansen MAK, Gaba V, Greenberg BM (1998) Higher plants and UV-B radiation: balancing damage, repair and acclimation. Trends in Plant Science **3:** 131-135

Jiang C, Gu J, Chopra S, Gu X, Peterson T (2004) Ordered origin of the typical two- and three-repeat Myb genes. Gene **326:** 13-22

Jiang C, Gu X, Peterson T (2004) Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of Arabidopsis and Oryza sativa L. ssp. indica. Genome Biol **5:** R46

Jiang C, Schommer CK, Kim SY, Suh DY (2006) Cloning and characterization of chalcone synthase from the moss, Physcomitrella patens. Phytochemistry **67:** 2531-2540

Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. EMBO J **19:** 6150-6161

Jin H, Martin C (1999) Multifunctionality and diversity within the plant MYB-gene family. Plant Molecular Biology **41:** 577-585

Jorgensen R (1993) The origin of land plants: a union of alga and fungus advanced by flavonoids? Biosystems **31:** 193-207

Jorgensen RA (2000) Directed Cell-to-Cell Movement of Functional Proteins: Do Transcription Factors Double as Signal Molecules in Plants? Sci. STKE **2000:** pe2-

Jung C, Seo JS, Han SW, Koo YJ, Kim CH, Song SI, Nahm BH, Choi YD, Cheong J-J (2008) Overexpression of AtMYB44 Enhances Stomatal Closure to Confer Abiotic Stress Tolerance in Transgenic Arabidopsis. Plant Physiol. **146:** 623-635

Karimi M, Inzé D, Depicker A (2002) GATEWAY(TM) vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science **7:** 193-195

Keiichiro H, Kyoko M, Tomotsugu K, Masaru O-T (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. The Plant Journal **34:** 733-739

Kerhoas L, Aouak D, Cingoz A, Routaboul J-M, Lepiniec L, Einhorn J, Birlirakis N (2006) Structural Characterization of the Major Flavonoid Glycosides from Arabidopsis thaliana Seeds. Journal of Agricultural and Food Chemistry **54:** 6603-6612

Kim DW, Lee SH, Choi S-B, Won S-K, Heo Y-K, Cho M, Park Y-I, Cho H-T (2006) Functional Conservation of a Root Hair Cell-Specific cis-Element in Angiosperms with Different Root Hair Distribution Patterns. Plant Cell**:** tpc.106.045229
Klempnauer KH, Gonda TJ, Bishop JM (1982) Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. Cell **31:** 453-463
Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends in Plant Science **10:** 236-242
Koornneef M (1990) Mutations Affecting the Testa Color in Arabidopsis. Arabidopsis Information Service**:** 1-4
Koornneef M, Bentsink L, Hilhorst H (2002) Seed dormancy and germination. Current Opinion in Plant Biology **5:** 33-36
Kranz H, Scholz K, Weisshaar B (2000) c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. Plant J **21:** 231-235
Kranz HD, Denekamp M, Greco R, Jin H, Leyva A, Meissner RC, Petroni K, Urzainqui A, Bevan M, Martin C, Smeekens S, Tonelli C, Paz-Ares J, Weisshaar B (1998) Towards functional characterisation of the members of the R2R3-MYB gene family from Arabidopsis thaliana. Plant J **16:** 263-276
Kucera B, Cohn MA, Leubner-Metzger G (2005) Plant hormone interactions during seed dormancy release and germination. Seed Science Research **15:** 281-307
Kurata T, Okada K, Wada T (2005) Intercellular movement of transcription factors. Current Opinion in Plant Biology **8:** 600-605
Lamb C, Dixon RA (1997) The oxidative burst in plant disease resistance. Annual Review of Plant Physiology and Plant Molecular Biology **48:** 251-275
Landry LG, Chapple CCS, Last RL (1995) Arabidopsis Mutants Lacking Phenolic Sunscreens Exhibit Enhanced Ultraviolet-B Injury and Oxidative Damage. Plant Physiol. **109:** 1159-1166
Lee M-W, Qi M, Yang Y (2007) A Novel Jasmonic Acid-Inducible Rice myb Gene Associates with Fungal Infection and Host Cell Death. Molecular Plant-Microbe Interactions **14:** 527-535
Lee MM, Schiefelbein J (2002) Cell Pattern in the Arabidopsis Root Epidermis Determined by Lateral Inhibition with Feedback. Plant Cell **14:** 611-618
Lee TI, Young RA (2000) Transcription of eukaryotic protein-coding genes. Annual Review of Genetics **34:** 77-137
Leech MJ, Kammerer W, Cove DJ, Martin C, Wang TL (1993) Expression of myb-related genes in the moss, Physcomitrella patens. Plant J **3:** 51-61
Lepiniec Lc, Debeaujon I, Routaboul J-M, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annual Review of Plant Biology **57:** 405-430
Li F, Cao Q-E, Ding Z (2007) Separation and Determination of Three Phenylpropanoids in the Traditional Chinese Medicine and Its Preparations by Capillary Electrophoresis. Journal of Chromatographic Science **45:** 354-359
Li J, Ou-Lee TM, Raba R, Amundson RG, Last RL (1993) Arabidopsis Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. Plant Cell **5:** 171-179

Li SF, Parish RW (1995) Isolation of two novel myb-like genes from Arabidopsis and studies on the DNA-binding properties of their products. Plant J **8:** 963-972
Liang XW, Dron M, Schmid J, Dixon RA, Lamb CJ (1989) Developmental and environmental regulation of a phenylalanine ammonia-lyase-beta-glucuronidase gene fusion in transgenic tobacco plants. Proceedings of the National Academy of Sciences of the United States of America **86:** 9284-9288
Liang YK, Dubos C, Dodd IC, Holroyd GH, Hetherington AM, Campbell MM (2005) AtMYB61, an R2R3-MYB transcription factor controlling stomatal aperture in Arabidopsis thaliana. Curr Biol **15:** 1201-1206
Lippold F, Sanchez DH, Musialak M, Schlereth A, Scheible W-R, Hincha DK, Udvardi MK (2009) AtMyb41 Regulates Transcriptional and Metabolic Responses to Osmotic Stress in Arabidopsis. Plant Physiol. **149:** 1761-1772
Lipsick JS (1996) One billion years of Myb. Oncogene **13:** 223-235
Lourdes G-G, Georg F, Thomas B (1999) A single locus determines sensitivity to bacterial flagellin in Arabidopsis thaliana. The Plant Journal **18:** 277-284
Martin C, Paz-Ares J (1997) MYB transcription factors in plants. Trends in Genetics **13:** 67-73
McCabe DE, Swain WF, Martinell BJ, Christou P (1988) Stable Transformation of Soybean (Glycine Max) by Particle Acceleration. Nat Biotech **6:** 923-926
McKay DL, Blumberg JB (2007) Roles for Epigallocatechin Gallate in Cardiovascular Disease and Obesity: An Introduction. J Am Coll Nutr **26:** 362S-365
Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M (1998) Arabidopsis thaliana: a model plant for genome analysis. Science **282:** 662, 679-682
Melchert TE, Alston RE (1965) Flavonoids from the moss Mnium affine Bland. Science **150:** 1170-1171
Mellway RD, Tran LT, Prouse MB, Campbell MM, Constabel CP (2009) The Wound-, Pathogen-, and Ultraviolet B-Responsive MYB134 Gene Encodes an R2R3 MYB Transcription Factor That Regulates Proanthocyanidin Synthesis in Poplar. Plant Physiol. **150:** 924-941
Mengiste T, Chen X, Salmeron J, Dietrich R (2003) The BOTRYTIS SUSCEPTIBLE1 Gene Encodes an R2R3MYB Transcription Factor Protein That Is Required for Biotic and Abiotic Stress Responses in Arabidopsis. Plant Cell **15:** 2551-2565
Meyer P, Heidmann I, Forkmann G, Saedler H (1987) A new petunia flower colour generated by transformation of a mutant with a maize gene. Nature **330:** 677-678
Michael JH, Leónie B, Wim JJS (2008) Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. New Phytologist **179:** 33-54
Mo Y, Nagel C, Taylor LP (1992) Biochemical complementation of chalcone synthase mutants defines a role for flavonols in functional pollen. Proceedings of the National Academy of Sciences of the United States of America **89:** 7213-7217
Mockaitis K, Estelle M (2008) Auxin Receptors and Plant Development: A New Signaling Paradigm. Annual Review of Cell and Developmental Biology **24:** 55-80
Mol J, Grotewold E, Koes R (1998) How genes paint flowers and seeds. Trends in Plant Science **3:** 212-217

Moore K, Subba Rao PV, Towers GHN (1968) Degradation of Phenylalanine and Tyrosine by Sporobolomyces roseus. Biochem J **106:** 507 - 514
Moyano E, Martinez-Garcia JF, Martin C (1996) Apparent Redundancy in myb Gene Function Provides Gearing for the Control of Flavonoid Biosynthesis in Antirrhinum Flowers. Plant Cell **8:** 1519-1532
Mu R-L, Cao Y-R, Liu Y-F, Lei G, Zou H-F, Liao Y, Wang H-W, Zhang W-K, Ma B, Du J-Z, Yuan M, Zhang J-S, Chen S-Y (2009) An R2R3-type transcription factor gene AtMYB59 regulates root growth and cell cycle progression in Arabidopsis. Cell Res **19:** 1291-1304
Muller S, Fuchs E, Ovecka M, Wysocka-Diller J, Benfey PN, Hauser M-T (2002) Two New Loci, PLEIADE and HYADE, Implicate Organ-Specific Regulation of Cytokinesis in Arabidopsis. Plant Physiol. **130:** 312-324
Murphy A, Peer WA, Taiz L (2000) Regulation of auxin transport by aminopeptidases and endogenous flavonoids. Planta **211:** 315-324
Nair HK, Rao KVK, Aalinkeel R, Mahajan S, Chawda R, Schwartz SA (2004) Inhibition of Prostate Cancer Cell Colony Formation by the Flavonoid Quercetin Correlates with Modulation of Specific Regulatory Genes. Clin. Diagn. Lab. Immunol. **11:** 63-69
Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell **2:** 279-289
Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L (2000) The TT8 Gene Encodes a Basic Helix-Loop-Helix Domain Protein Required for Expression of DFR and BAN Genes in Arabidopsis Siliques. Plant Cell **12:** 1863-1878
Newman LJ, Perazza DE, Juda L, Campbell MM (2004) Involvement of the R2R3-MYB, AtMYB61, in the ectopic lignification and dark-photomorphogenic components of the det3 mutant phenotype. Plant Journal **37:** 239-250
Nishiyama T, Fujita T, Shin IT, Seki M, Nishide H, Uchiyama I, Kamiya A, Carninci P, Hayashizaki Y, Shinozaki K, Kohara Y, Hasebe M (2003) Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution. Proc Natl Acad Sci U S A **100:** 8007-8012
Noel JP, Austin MB, Bomati EK (2005) Structure-function relationships in plant phenylpropanoid biosynthesis. Current Opinion in Plant Biology **8:** 249-253
Nowak MA, Boerlijst MC, Cooke J, Smith JM (1997) Evolution of genetic redundancy. Nature **388:** 167-171
Ogata K, Hojo H, Aimoto S, Nakai T, Nakamura H, Sarai A, Ishii S, Nishimura Y (1992) Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. Proc Natl Acad Sci U S A **89:** 6428-6432
Ogata K, Morikawa S, Nakamura H, Sekikawa A, Inoue T, Kanai H, Sarai A, Ishii S, Nishimura Y (1994) Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. Cell **79:** 639-648
Ohta M, Matsui K, Hiratsu K, Shinshi H, Ohme-Takagi M (2001) Repression Domains of Class II ERF Transcriptional Repressors Share an Essential Motif for Active Repression. Plant Cell **13:** 1959-1968

Otsuka H, Van Haastert PJ (1998) A novel Myb homolog initiates Dictyostelium development by induction of adenylyl cyclase expression. Genes Dev **12:** 1738-1748
Paul EV, Manu A, Surekha K-A, Jianhua Z, Jian-Kang Z (2006) Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. The Plant Journal **45:** 523-539
Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H (1987) The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. Embo J **6:** 3553-3558
Peer WA, Brown DE, Tague BW, Muday GK, Taiz L, Murphy AS (2001) Flavonoid Accumulation Patterns of Transparent Testa Mutants of Arabidopsis. Plant Physiol. **126:** 536-548
Peer WA, Murphy AS (2007) Flavonoids and auxin transport: modulators or regulators? Trends in Plant Science **12:** 556-563
Preston J, Wheeler J, Heazlewood J, Li SF, Parish RW (2004) AtMYB32 is required for normal pollen development in Arabidopsis thaliana. Plant Journal **40:** 979-995
Pu L, Li Q, Fan X, Yang W, Xue Y (2008) The R2R3 MYB Transcription Factor GhMYB109 Is Required for Cotton Fiber Development. Genetics **180:** 811-820
Rabinowicz PD, Braun EL, Wolfe AD, Bowen B, Grotewold E (1999) Maize R2R3 Myb genes: Sequence analysis reveals amplification in the higher plants. Genetics **153:** 427-444
Raven JA (2000) Land plant biochemistry. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **355:** 833-846
Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud P-F, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin-I T, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S-i, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu S-H, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The Physcomitrella Genome Reveals Evolutionary Insights into the Conquest of Land by Plants. Science **319:** 64-69
Richard AD, Lahoucine A, Parvathi K, Chang-Jun L, Reddy MSS, Liangjiang W (2002) The phenylpropanoid pathway and plant defence—a genomics perspective. Molecular Plant Pathology **3:** 371-390
Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science **290:** 2105-2110
Ringli C, Bigler L, Kuhn BM, Leiber R-M, Diet A, Santelia D, Frey B, Pollmann S, Klein M (2008) The Modified Flavonol Glycosylation Profile in the Arabidopsis rol1 Mutants Results in Alterations in Plant Growth and Cell Shape Formation. Plant Cell**:** tpc.107.053249

Romero, Fuertes, Benito, Malpica, Leyva, Paz2010Ares (1998) More than 80 R2R3 2010;MYB regulatory genes in the genome of Arabidopsis thaliana. The Plant Journal **14:** 273-284
Romero, Fuertes, Benito, Malpica, Leyva, Paz A (1998) More than 80R2R3 MYB regulatory genes in the genome of Arabidopsis thaliana. The Plant Journal **14:** 273-284
Rosinski JA, Atchley WR (1998) Molecular Evolution of the Myb Family of Transcription Factors: Evidence for Polyphyletic Origin. Journal of Molecular Evolution **46:** 74-83
Rowan DD, Cao M, Lin-Wang K, Cooney JM, Jensen DJ, Austin PT, Hunt MB, Norling C, Hellens RP, Schaffer RJ, Allan AC (2009) Environmental regulation of leaf colour in red 35S:PAP1 Arabidopsis thaliana. New Phytologist **182:** 102-115
Ruzicka K, Ljung K, Vanneste S, Podhorska R, Beeckman T, Friml J, Benkova E (2007) Ethylene Regulates Root Growth through Effects on Auxin Biosynthesis and Transport-Dependent Auxin Distribution. Plant Cell **19:** 2197-2212
Ryu KH, Kang YH, Park Y-h, Hwang I, Schiefelbein J, Lee MM (2005) The WEREWOLF MYB protein directly regulates CAPRICE transcription during cell fate specification in the Arabidopsis root epidermis. Development **132:** 4765-4775
Saibo NJM, Lourenco T, Oliveira MM (2009) Transcription factors and regulation of photosynthetic and related metabolism under environmental stresses. Ann Bot **103:** 609-623
Saikumar P, Murali R, Reddy EP (1990) Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions. Proceedings of the National Academy of Sciences of the United States of America **87:** 8452-8456
Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. Nat Rev Genet **8:** 424-436
Schaefer DG, Zryd JP (2001) The moss Physcomitrella patens, now and then. Plant Physiol **127:** 1430-1438
Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C (2006) A Small Family of MYB-Regulatory Genes Controls Floral Pigmentation Intensity and Patterning in the Genus Antirrhinum. Plant Cell **18:** 831-851
Sharrocks AD (2000) Introduction: the regulation of eukaryotic transcription factor function. Cellular and Molecular Life Sciences **57:** 1147-1148
Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM (1995) Analysis of Arabidopsis Mutants Deficient in Flavonoid Biosynthesis. Plant Journal **8:** 659-671
Silber MV, Meimberg H, Ebel J (2008) Identification of a 4-coumarate:CoA ligase gene family in the moss, Physcomitrella patens. Phytochemistry **69:** 2449-2456
Song Feng L, Roger WP (1995) Isolation of two novel myb-like genes from Arabidopsis and studies on the DNA-binding properties of their products. The Plant Journal **8:** 963-972
Spelt C, Quattrocchio F, Mol JNM, Koes R (2000) anthocyanin1 of Petunia Encodes a Basic Helix-Loop-Helix Protein That Directly Activates Transcription of Structural Anthocyanin Genes. Plant Cell **12:** 1619-1632

Stafford HA (1991) Flavonoid Evolution: An Enzymic Approach. Plant Physiol. **96:** 680-685
Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in Arabidopsis thaliana. Curr Opin Plant Biol **4:** 447-456
Sylvie A, Michel D, Martine D (1997) BANYULS, a novel negative regulator of flavonoid biosynthesis in the Arabidopsis seed coat. The Plant Journal **11:** 289-299
Taito T, Yuko S, Makoto S, Hidemi K, Miyako U-T, Motoyuki A, Makoto M, Chiharu U (2003) The OsTB1 gene negatively regulates lateral branching in rice. The Plant Journal **33:** 513-520
Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Macia FA, Roberts K, Martin C (1998) The AmMYB308 and AmMYB330 Transcription Factors from Antirrhinum Regulate Phenylpropanoid and Lignin Biosynthesis in Transgenic Tobacco. Plant Cell **10:** 135-154
Todd CD, Zeng P, Huete AMR, Hoyos ME, Polacco JC (2004) Transcripts of MYB-like genes respond to phosphorous and nitrogen deprivation in Arabidopsis. Planta **219:** 1003-1009
Tominaga R, Iwata M, Okada K, Wada T (2007) Functional Analysis of the Epidermal-Specific MYB Genes CAPRICE and WEREWOLF in Arabidopsis. Plant Cell **19:** 2264-2277
Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K (1993) An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. Plant Cell **5:** 1529-1539
Vailleau F, Daniel X, Tronchet M, Montillet J-L, TriantaphylidÃ¨s C, Roby D (2002) A R2R3-MYB gene, AtMYB30, acts as a positive regulator of the hypersensitive cell death program in plants in response to pathogen attack. Proceedings of the National Academy of Sciences of the United States of America **99:** 10179-10184
Van Doorsselaere J, Baucher M, Chognot E, Chabbert B, Tollier MT, Petit-Conil M, Lepl, eacute, J C, Pilate G, Cornu D, Monties B, Van Montagu M, Inz, eacute, D, Boerjan W, Jouanin L (1995) A novel lignin in poplar trees with a reduced caffeic acid/5-hydroxyferulic acid O-methyltransferase activity. The Plant Journal **8:** 855-864
van Nocker S, Ludwig P (2003) The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function. BMC Genomics **4:** 50
Veit M, Pauli GF (1999) Major Flavonoids from Arabidopsis thaliana Leaves. Journal of Natural Products **62:** 1301-1303
Verdonk JC, Haring MA, van Tunen AJ, Schuurink RC (2005) ODORANT1 Regulates Fragrance Biosynthesis in Petunia Flowers. Plant Cell **17:** 1612-1624
Verslues PE, Agarwal M, Katiyar-Agarwal S, Zhu J, Zhu JK (2006) Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. Plant J **45:** 523-539
Walhout AJM (2006) Unraveling transcription regulatory networks by protein-DNA and protein-protein interaction mapping. Genome Research **16:** 1445-1454
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

Waters ER (2003) Molecular adaptation and the origin of land plants. Molecular Phylogenetics and Evolution **29:** 456-463
Weijers D, Jürgens G (2004) Funneling auxin action: specificity in signal transduction. Current Opinion in Plant Biology **7:** 687-693
Weisshaar B, Jenkins GI (1998) Phenylpropanoid biosynthesis and its regulation. Current Opinion in Plant Biology **1:** 251-257
Weston K (1998) Myb proteins in life, death and differentiation. Curr Opin Genet Dev **8:** 76-81
Wiedemeier AMD, Judy-March JE, Hocart CH, Wasteneys GO, Williamson RE, Baskin TI (2002) Mutant alleles of Arabidopsis RADIALLY SWOLLEN 4 and 7 reduce growth anisotropy without altering the transverse orientation of cortical microtubules or cellulose microfibrils. Development **129:** 4821-4830
Winkel-Shirley B (1999) Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. Physiologia Plantarum **107:** 142-149
Winkel-Shirley B (2001) Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology. Plant Physiol. **126:** 485-493
Wolberger C (1999) Multiprotein-dna complexes in transcriptional regulation. Annual Review of Biophysics and Biomolecular Structure **28:** 29-56
Woo H-H, Faull KF, Hirsch AM, Hawes MC (2003) Altered Life Cycle in Arabidopsis Plants Expressing PsUGT1, a UDP-Glucuronosyltransferase-Encoding Gene from Pea. Plant Physiol. **133:** 538-548
Wu X, Dinneny JR, Crawford KM, Rhee Y, Citovsky V, Zambryski PC, Weigel D (2003) Modes of intercellular transcription factor movement in the Arabidopsis apex. Development **130:** 3735-3745
Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu Rev Plant Biol **57:** 781-803
Yamazaki Y, Suh D-Y, Sitthithaworn W, Ishiguro K, Kobayashi Y, Shibuya M, Ebizuka Y, Sankawa U (2001) Diverse chalcone synthase superfamily enzymes from the most primitive vascular plant, Psilotum nudum. Planta **214:** 75-84
Yanagisawa S (1998) Transcription factors in plants: Physiological functions and regulation of expression. Journal of Plant Research **111:** 363-371
Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, Yunping S, Li Z, Xiaohui D, Jingchu L, Xing-Wang D, Zhangliang C, Hongya G, Li-Jia Q (2006) The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol Biol **60:** 107-124
Zhou J, Lee C, Zhong R, Ye Z-H (2009) MYB58 and MYB63 Are Transcriptional Activators of the Lignin Biosynthetic Pathway during Secondary Cell Wall Formation in Arabidopsis. Plant Cell **21:** 248-266
Zhu J, Shi H, Lee BH, Damsz B, Cheng S, Stirm V, Zhu JK, Hasegawa PM, Bressan RA (2004) An Arabidopsis homeodomain transcription factor gene, HOS9, mediates cold tolerance through a CBF-independent pathway. Proc Natl Acad Sci U S A **101:** 9873-9878
Zhu J, Verslues PE, Zheng X, Lee BH, Zhan X, Manabe Y, Sokolchik I, Zhu Y, Dong CH, Zhu JK, Hasegawa PM, Bressan RA (2005) HOS10 encodes an R2R3-

type MYB transcription factor essential for cold acclimation in plants. *In* Proc Natl Acad Sci U S A, Vol 102, pp 9966-9971
Zhu J, Verslues PE, Zheng X, Lee BH, Zhan X, Manabe Y, Sokolchik I, Zhu Y, Dong CH, Zhu JK, Hasegawa PM, Bressan RA (2010) Retraction for Zhu et al., HOS10 encodes an R2R3-type MYB transcription factor essential for cold acclimation in plants. Proc Natl Acad Sci U S A **107:** 13972
Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis Microarray Database and Analysis Toolbox. Plant Physiol. **136:** 2621-2632

APPENDIX

Appendix 1. List of primers used

Primer name	Sequence 5'- 3'
SJS1	CCGTAATCACGGTGAAGGTT
SJS2	AAATTATTTTGGAAACCATGTATCG
SJS3III	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTTGACTGAACGACAGCTT
SJS4II	GGGGACCACTTTGTACAAGAAAGCTGGGTCGGCCAGATTTAATAGTCTGGCT
SJS5II	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCACCATGGGACGGGCTCCC
SJS6	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATGTCTGGAAGCCGTCC
SJS7	ATCATGGTGAAGGCTCTTGG
SJS8	TTTAACAGAACCGGGAAACG
SJS9	CCGGGTGTTGTGAGTCTTTT
SJS10	TTAGTTGTCGCCATTCACCA
SJS13	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCAGTGTCATATAGAATACTTATATAG GTGTT
SJS14II	GGGGACCACTTTGTACAAGAAAGCTGGGTCGTAGTAATTGCTCACGAATTTTAG
SJS15	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCAAACCCTAAT TTCTCCCAAAACTAA
SJS16	GGGGACCACTTTGTACAAGAAAGCTGGGTCCACAGAACATT GTGTGTGTGTCT
SJS18	CAACAGTGACAAACGCGCCT
SJS24	TCCTGAGAGGCCCATTAAG
SJS26	ATAAGATTCTGGACAGAGA
SJS33	GTGCAATCCATCTTGTTCAATCAT
SJS34	CCTTCTTGACGAGTTCTTCTGA
SJS52	CCTCCACCAGAAACATCTCG
SJS53	TGCCTCAATTTGGTTATGAATG
SJS54	GCGATTAACAAAAGCATTCTAAA
SJS55	GGCGACAAAAGAGTTCTTCG
SJS56	AAGGTGAAATCATGCCGAAC
SJS59	TGAACTATCTGCGCCCTGACTTGA
SJS60	TGCCTCTGGTTTCCTTCCATGTCT
SJS61	AGGCCCTGTGCAAGGTAAGAAGAT
SJS62	AGGCTCTTGGCGTTCTCTTCCTAA
SJS63	TAACAGAACCGGGAAACGTGGACT
SJS64	GGGGACA AGTT GTACAAAAAAGCAGGCTCCTTCACAATAAAT AGCTATTGATTAAAGATAGCATTGTC
SJS65	GGGGACCACTTTGTACAAGAAAGCTGGGTCGGTGAATGAAGAAGATGGTGTAAGA G
SJS69	ATTAATCGCTGGCCGCATT
SJS71	TAACAGAGGCCCATTAATGCA
SJS73	TGCCGGAAAAAGAAGATCAG

SJS75II	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTCGCTTTTGGCCTTGTTTAT
SJS76II	GGGGACCACTTTGTACAAGAAAGCTGGGTCGATGATGGTTACATGTGTGGGTG
SJS78	TCTGATGATTTTATCTTTTCTTTAGGG
SJS79	AGCACAACCTAATTCCACGAGCCCG
SJS80	ATATGTTTGCAAACAGCCTGA
SJS81	TCCTTAATGGACATGCTCTAAG
SJS87	ACAGAGTCGTTCACTTTTTTGT
LB-Salk	GACCGCTTGCTGCAACTCTC
LB-GabiKat	CATATTGACCATCATACTCATTGC
Att-Fw	ACAAGTTTGTACAAAAAAGCAGGCT
Att-Rv	ACCACTTTGTACAAGAAAGCTGGGT
NptII-Fw	CGCGACGTCAATGATTGAACAAGATGGATTGCA
NptII-Rv	CAATCTTAAGAAACTTTTCAGAAGAACTCGTCAAGAAGG
Cyclophillin-Fw	TCTTCCTCTTCGGAGCCATA
Cyclophillin-Rv	AAGCTGGGAATGATTCGATG
L21-Fw	GGTTGGTCATGGGTTGCG
L21-Rv	GAGGTCAACTGTCTCGCC
PpCHS-Fw	CAAGCGCATCTGTGACAAGT
PpCHS-Rv	GCGAAGACGATGTGGGTAAT
Pp4CL2-Fw	AGCATCAGGATGGTGATGTCCG
Pp4CL2-Rv	ACATTGTGGGAGGATACCGC
AtPAL-Fw	CCAAAAACGGTGTCGCACT
AtPAL-Rv	GCTTCCGAATATTCCGGCGTTAA
AtC4H-Fw	GGAGAAATCAACGAGGACAATGTTC
AtC4H-Rv	CCACTCGATAGACCACAATGTTGT
At4CL3-Fw	GGATACAACCAACCTGGTGAGA
At4CL3-Rv	ACCTTCTTCGTCTATTGTTGCTGAA
AtCHS-Fw	CCGACCTCAAGGAGAAGTTCAAG
AtCHS-Rv	GCATGTGACGTTTCCGAATTGT
AtCHI-Fw	CTCTCTTACGGTTGCGTTTTC
AtCHI-Rv	GTTCTTCCCGATGATAGATTC
AtF3H-Fw	ACCACGGCCATTTTTTGAGC
AtF3H-Rv	CGTGGCTATGGATAATCTGC
AtC3H-Fw	GACAGCGATAACAGCGGAAT

Appendix 2. Best BLASTP hits for Arabidopsis genes of the phenylpropanoid pathway in the Physcomitrella genome

PAL1

Sequence: AT2G37040.

MEINGAHKSNGGGVDAMLCGGDIKTKNMVINAEDPLNWGAAAEQMKGSHLDEVKRMVAEFRKPVVNLGGE
TLTIGQVAAISTIGNSVKVELSETARAGVNASSDWVMESMNKGTDSYGVTTGFGATSHRRTKNGVALQKE
LIRFLNAGIFGSTKETSHTLPHSATRAAMLVRINTLLQGFSGIRFEILEAITSFLNNNITPSLPLRGTIT
ASGDLVPLSYIAGLLTGRPNSKATGPNGEALTAEEAFKLAGISSGFFDLQPKEGLALVNGTAVGSGMASM
VLFETNVLSVLAEILSAVFAEVMSGKPEFTDHLTHRLKHHPGQIEAAAIMEHILDGSSYMKLAQKLHEMD
PLQKPKQDRYALRTSPQWLGPQIEVIRYATKSIEREINSVNDNPLIDVSRNKAIHGGNFQGTPIGVSMDN
TRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLTASRNPSLDYGFKGAEIAMASYCSELQYLANPVTSHVQ
SAEQHNQDVNSLGLISSRKTSEAVDILKLMSTTFLVAICQAVDLRHLEENLRQTVKNTVSQVAKKVLTTG
VNGELHPSRFCEKDLLKVVDREQVYTYADDPCSATYPLIQKLRQVIVDHALINGESEKNAVTSIFHKIGA
FEEELKAVLPKEVEAARAAYDNGTSAIPNRIKECRSYPLYRFVREELGTELLTGEKVTSPGEEFDKVFTA
ICEGKIIDPMMECLNEWNGAPIPIC-

Sequence Pp: 121522

>e_gw1.36.41.1 [Phypa1_1:121522]
MVGFRDVNAQSEDPLNWGKVAAGLAGSHLQEVRQMVTAFFECSVVVLEGASLTIAQVAAIARRPEVKVVL
DADTAKERVDESSDWVLSCALKGTDTYGVTTGFGATSHRRTNQVAELQFELIRFLNAGIVGKGDGNTLPS
PTTRAAMLVRTNTLMQGYSGIRWEILVAISKLLNANVTPKLPLRGTITASGDLVPLSYIAGCLTGRQNSR
AYTAEGKEVTSEEALRIAGVGKPFQLEPKEGLALVNGTAVGSALASTVCFDANVLVVMAEVLSALFCEVM
QGKPEFADPLTHKLKHHPGQMEAAAMMEWVLDGSSYMKAAAEAHATDPLSKPKQDRYALRTSPQWLGPHI
ETIRAATHSIQREINSVNDNPLIDVSRDMALHGGNFQGTPIGVSMDNMRLSIAGIGKLMFAQFSELVNDF
YNNGLPSNLSGGPNPSLDYGMKGGEIAMASYTSELQYLANPVTTHVQSAEQHNQDVNSLGLISARKTAEA
IEILKLMTSTYLVALCQAIDLRHLEETMQATVKTVVAQAAKKTLFLGTNGALQVARFCEKELLKVVERQP
VFTYVDDPANTKYPLMEKLRQVLVEYSLKYITEKDDGLSIFNKIPAFEEEVKTQLRLEVAAARASFDNGV
SPIPNRVMDCRSYPLYQFVRSELGTALLTGQSSQSPGTDFEKVYDAICEGRHVAHLMKVLEGWSCVPGSL
CV*

C4H

Sequence: AT2G30490.1

MDLLLLEKSLIAVFVAVILATVISKLRGKKLKLPPGPIPIPIFGNWLQVGDDLNHRNLVDYAKKFGDLFL
LRMGQRNLVVVSSPDLTKEVLLTQGVEFGSRTRNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTN
KVVQQNREGWEFEAASVVEDVKKNPDSATKGIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFLRLKALN
GERSRLAQSFEYNYGDFIPILRPFLRGYLKICQDVKDRRIALFKKYFVDERKQIASSKPTGSEGLKCAID
HILEAEQKGEINEDNVLYIVENINVAAIETTLWSIEWGIAELVNHPEIQSKLRNELDTVLGPGVQVTEPD
LHKLPYLQAVVKETLRLRMAIPLLVPHMNLHDAKLAGYDIPAESKILVNAWWLANNPNSWKKPEEFRPER
FFEEESHVEANGNDFRYVPFGVGRRSCPGIILALPILGITIGRMVQNFELLPPPGQSKVDTSEKGGQFSL
HILNHSIIVMKPRNC-

Sequence Pp: 140533

>e_gw1.168.38.1 [Phypa1_1:140533]
MNLEGAVQALLVAVLLGLLIAKLRAPKLNLPPGPVALPIVGNWLQVGDDLNHRNLAEMSQKYGDVFLLKM
GQRNLVVVSSPEVAKEVLHTQSVEFGSRTRNVVFDIFTGNGQDMVFTVYGDHWRRMRRIMTVPFFTNKVV
QQSRGAWEDEALRVIQDLKAKPEASTTGVVIRKRLQLMMYNIMYRLMFDSRFESEEDPLFLKLKALNGER
SRLAQSFEYNYGDFIPVLRPLLRGYLKVCQEIKDRRLALFKEHFLDERKKLLSTLGPRPDGEKAAIDLIL
EAQKRGEINEENVLYIVENINVAAIETTLWSIEWGVAELVNNPEMQTRIREELDSTLGKGNLITEPDTYN

NKLPYLSAFVKEVMRLHMAIPLLVPHMNLHQAKLAGYDIPAESKILVNAWWIANNPNHWDQPEKFIPERF
LDGKIEAKGDDFRFLPFGSGRRSCPGIIIAMPLLSIVLGRLVQSLELLPPPGTKKVDVSEKGGQFSLHIA
THSTVVCKPIA*

4CL3

Sequence: AT1G65060.1

MITAALHEPQIHKPTDTSVVSDDVLPHSPPTPRIFRSKLPDIDIPNHLPLHTYCFEKLSSVSDKPCLIVG
STGKSYTYGETHLICRRVASGLYKLGIRKGDVIMILLQNSAEFVFSFMGASMIGAVSTTANPFYTSQELY
KQLKSSGAKLIITHSQYVDKLKNLGENLTLITTDEPTPENCLPFSTLITDDETNPFQETVDIGGDDAAAL
PFSSGTTGLPKGVVLTHKSLITSVAQQVDGDNPNLYLKSNDVILCVLPLFHIYSLNSVLLNSLRSGATVL
LMHKFEIGALLDLIQRHRVTIAALVPPLVIALAKNPTVNSYDLSSVRFVLSGAAPLGKELQDSLRRRLPQ
AILGQGYGMTEAGPVLSMSLGFAKEPIPTKSGSCGTVVRNAELKVVHLETRLSLGYNQPGEICIRGQQIM
KEYLNDPEATSATIDEEGWLHTGDIGYVDEDDEIFIVDRLKEVIKFKGFQVPPAELESLLINHHSIADAA
VVPQNDEVAGEVPVAFVVRSNGNDITEEDVKEYVAKQVVFYKRLHKVFFVASIPKSPSGKILRKDLKAKL
C-

Sequence Pp: 140413

>e_gw1.167.67.1 [Phypa1_1:140413]
MESEVCDFIYRSKLPDIDIPNHMPLSDYCLEKAAQWPDKVCLIDGVTGREHTYGEIELSTRRVAAGLFKI
GVKQGDVIALLLPNCAEFVQVFLGAAKMGAIVTTANPFYTSAELEKQTIASGAGIVVTHSSYIEKLAGLN
VQVPTTSHPVSIITVDQHVDKCMHISMLLEPNEAECPQVEIHPDDVVCLPYSSGTTGLPKGVMLTHKSLV
SSVSQQVDGDSPNFNITVEDTLMCVLPMFHIYSLNSILLCGLRVGATLVIMPKFELSKMLELIQKHKVTM
GPFVPPIVLAIAKNPIVENYDLSSIKMVMSGAAPLGKELEDAFRARLPNAVLGQGYGMTEAGPVLAMCLA
FAKSPFPVKPGSCGTVVRNAEVKIVDTETGMSLPYNQPGEICIRGPQIMKGYLNNPEATANTIDKDGFLH
TGDVAFIDEDEEMFIVDRVKEIIKFKGFQVPPAELEALLLSNEEIQHAAVVSRKDDVAGEVPVAFVVRQA
GSTISEEEVKDYVAKQVVFYKKIHNVYFVDSIPESPSGKILRKDLRNKV*

C3H

Sequence: AT2G40890.1

MSWFLIAVATIAAVVSYKLIQRLRYKFPPGPSPKPIVGNLYDIKPVRFRCYYEWAQSYGPIISVWIGSIL
NVVVSSAELAKEVLKEHDQKLADRHRNRSTEAFSRNGQDLIWADYGPHYVKVRKVCTLELFTPKRLESLR
PIREDEVTAMVESVFRDCNLPENRAKGLQLRKYLGAVAFNNITRLAFGKRFMNAEGVVDEQGLEFKAIVS
NGLKLGASLSIAEHIPWLRWMFPADEKAFAEHGARRDRLTRAIMEEHTLARQKSSGAKQHFVDALLTLKD
QYDLSEDTIIGLLWDMITAGMDTTAITAEWAMAEMIKNPRVQQKVQEEFDRVVGLDRILTEADFSRLPYL
QCVVKESFRLHPPTPLMLPHRSNADVKIGGYDIPKGSNVHVNVWAVARDPAVWKNPFEFRPERFLEEDVD
MKGHDFRLLPFGAGRRVCPGAQLGINLVTSMMSHLLHHFVWTPPQGTKPEEIDMSENPGLVTYMRTPVQA
VATPRLPSDLYKRVPYDM-

Sequence Pp: 98753

>fgenesh1_pg.scaffold_366000008 [Phypa1_1:98753]
MYKSLRSSHKLPPGPRPLPVVGNLTHITPVRFKCFMEWAQTYGSVLSVWMGPTLNVVVSSADAAKEMLKE
RDHALSSRPLTRAAARFSRNGQDLIWADYGPHYVKVRKVCTLELFTFKRLESLKPVREDEVGAMVAALFK
DCADSRPLNLKKYVSAMAFNNITRIVFGKRFVDDKGNIDNQGVEFKEIVSQGMKLGASLKMSEHIPYLRW
MFPLQEEEFAKHGARRDNLTKAIMQEHRLQSQKNGPGHHFVDALLSMQKQYDLSETTIIGLLWDMITAGM
DTTAISVEWAIAELVRNPDVQVKAQQELDQVVGQDRVVTEADFSQLPYLQAVAKEALRLHPPTPLMLPHK
ATETVKIGGYDVPKGTVVHCNVYAISRDPTVWEEPLRFRPERFLEEDIDIKGHDYRLLPFGAGRRVCPGA
QLGLNMVQLMLARLLHHFSWAPPPGVTPAAIDMTERPGVVTFMAAPLQVLATPRLRAALYKNGSSPS*

CHS

Sequence: AT5G13930.1

MVMAGASSLDEIRQAQRADGPAGILAIGTANPENHVLQAEYPDYYFRITNSEHMTDLKEKFKRMCDKSTI
RKRHMHLTEEFLKENPHMCAYMAPSLDTRQDIVVVEVPKLGKEAAVKAIKEWGQPKSKITHVVFCTTSGV
DMPGADYQLTKLLGLRPSVKRLMMYQQGCFAGGTVLRIAKDLAENNRGARVLVVCSEITAVTFRGPSDTH
LDSLVGQALFSDGAAALIVGSDPDTSVGEKPIFEMVSAAQTILPDSDGAIDGHLREVGLTFHLLKDVPGL
ISKNIVKSLDEAFKPLGISDWNSLFWIAHPGGPAILDQVEIKLGLKEEKMRATRHVLSEYGNMSSACVLF
ILDEMRRKSAKDGVATTGEGLEWGVLFGFGPGLTVETVVLHSVPL-

Sequence Pp: 104998

>estExt_fgenesh1_pm.C_220002 [Phypa1_1:104998]
MASAGDVTRVALPRGQPRAEGPACVLGIGTAVPPAEFLQSEYPDFFFNITNCGEKEALKAKFKRICDKSG
IRKRHMFLTEEVLKANPGICTYMEPSLNVRHDIVVVQVPKLAAEAAQKAIKEWGGRKSDITHIVFATTSG
VNMPGADHALAKLLGLKPTVKRVMMYQTGCFGGASVLRVAKDLAENNKGARVLAVASEVTAVTYRAPSEN
HLDGLVGSALFGDGAGVYVVGSDPKPEVEKPLFEVHWAGETILPESDGAIDGHLTEAGLIFHLMKDVPGL
ISKNIEKFLNEARKPVGSPAWNEMFWAVHPGGPAILDQVEAKLKLTKDKMQGSRDILSEFGNMSSASVLF
VLDQIRHRSVKMGASTLGEGSEFGFFIGFGPGLTLEVLVLRAAPNSA

CHI

Sequence: AT5G05270.1

MGTEMVMVHEVPFPPQIITSKPLSLLGQGITDIEIHFLQVKFTAIGVYLDPSDVKTHLDNWKGKTGKELA
GDDDFFDALASAEMEKVIRVVVIKEIKGAQYGVQLENTVRDRLAEEDKYEEEEETELEKVVGFFQSKYFK
ANSVITYHFSAKDGICEIGFETEGKEEEKLKVENANVVGMMQRWYLSGSRGVSPSTIVSIADSISAVLT-

No hits found in Physcomitrella

F3H

Sequence: AT3G51240.1

MAPGTLTELAGESKLNSKFVRDEDERPKVAYNVFSDEIPVISLAGIDDVDGKRGEICRQIVEACENWGIF
QVVDHGVDTNLVADMTRLARDFFALPPEDKLRFDMSGGKKGGFIVSSHLQGEAVQDWREIVTYFSYPVRN
RDYSRWPDKPEGWVKVTEEYSERLMSLACKLLEVLSEAMGLEKESLTNACVDMDQKIVVNYYPKCPQPDL
TLGLKRHTDPGTITLLLQDQVGGLQATRDNGKTWITVQPVEGAFVVNLGDHGHFLSNGRFKNADHQAVVN
SNSSRLSIATFQNPAPDATVYPLKVREGEKAILEEPITFAEMYKRKMGRDLELARLKKLAKEERDHKEVD
KPVDQIFA-

Sequence Pp：108581

>estExt_fgenesh1_pm.C_2870008 [Phypa1_1:108581]
MLGPSSFERVQSLSEQGLLEVPSSYIRPAEERPSISELVGEIPVIDLADGSLDVTAQIGQACREWGFFQV
VNHGVPKELLNRMLELGAHFYAKPMEEKLAYACKDPGTAPEGYGSRMLVKEEQVMDWRDYIDHHTLPLSR
RNPSRWPSDPPHYRSSMEEFSDETCKLARRILGHISESLGLPTQFLEDAVGEPAQNIVINYYPTCPQPQL
TLGLQAHSDMGAITLLLQDDVAGLQVKKNNEWSTIQPIRDTFVVNLGDMLQILSNDKYRSVEHRTVVNGE
RARKSVAVFYDPAKNRLISPAAPLVDKDHPALFPSILYGDHVLNWYSKGPDGKRTIDSLIIE