

The role of class 1 *KNOX* genes in sporophyte evolution

Eftychios Frangedakis

The Queen's College
University of Oxford

Thesis submitted for the degree of Doctor of Philosophy
Michaelmas term 2014

Abstract

Land plants are characterized by the alternation of two generations: the haploid gametophyte and the diploid sporophyte. As land plants evolved from bryophytes to vascular plants, the sporophyte became the dominant generation in the life cycle. The sporophytes of bryophytes are developmentally simple structures characterized by determinate growth. In contrast, the sporophytes of vascular plants diverged to become highly variable and often complex structures with indeterminate growth.

KNOTTED1-LIKE HOMEODOMAIN (KNOX) genes encode homeodomain containing transcription factors that are key regulators of sporophyte development. *KNOX* genes are divided into two subclasses, class 1 and class 2. The critical role of class 1 *KNOX* genes in the apical growth of all extant land plants studied to date, suggests that modifications to class 1 *KNOX* gene function may have played an important role in sporophyte evolution. However, the nature of any such modifications is largely unknown. In this study, a number of cross-species complementation experiments were carried out to determine the extent to which class 1 *KNOX* gene function is conserved between different land plant groups. The role of *KNOX* genes in hornworts was also investigated because hornworts are believed to be the sister group to vascular plants.

The work presented in this thesis demonstrates that the function of class 1 *KNOX* genes is conserved between lycophytes and angiosperms. In contrast, the function of class 1 *KNOX* genes appears to have diversified between bryophytes and vascular plants. Collectively, these results indicate increased complexity of class 1 *KNOX* gene function during the evolution of land plants.

Acknowledgements

Firstly, I would like to thank my super-supervisor Jane Langdale for her generosity, guidance and support throughout the course of my DPhil. She has been a great supervisor. Ευχαριστώ πολύ Jane!

Furthermore, I am very grateful to the fantastic Laura Moody for ‘one million’ reasons (e.g. always having the best answer to my endless questions, showing me how to play with moss protoplasts, and revising my thesis) and to the awesome Yuki Yasumura for helping me set foot in the moss world. I especially want to thank Ester Rabbinoiwtsch for collecting a quantity of *Arabidopsis* seeds that is double the greek debt; Andy Plackett for always being very helpful, revising my thesis and for the best cakes ever made by a plant scientist; Sayuri Ando for sharing with me her japanese secrets about moss; Denis Saint-Marcoux for bryo-evo-devo help and Heather Sanders for help during the beginning of my DPhil. I would also like to thank my second supervisor Liam Dolan for his helpful comments and advice.

I must also thank all the wonderful people in the Langdale lab for creating a unique working environment that I am sure will be very difficult to find again. In particular, I want to thank Mara Schuler for helping to establish the “Plant Sciences Sigmund Freud Centre”; Julie Bull who is the best lab manager in the Galaxy; Jim Fouracre who was the best student of my Turkish history lessons; and Peng Wang for being a rare example of scientific perfection.

Many thanks also go to Keiko Sakakibara (University of Tokyo) for the moss mutant lines and for sharing her expertise of moss sporophytes; Juan Carlos Villarreal (University of Munich) for introducing me to the world of hornworts; Steve Kelly for annotating the genome of *Anthoceros punctatus*; David Emms for helping with the phylogenetic trees; David Long and Laura Forrest (RBGE) for helping to obtain *A. punctatus* and *A. agrestis* plants; Dietmar Quandt (University of Bonn) for advice on hornwort sporophyte induction media; and John Baker for taking the best photos.

Finally, I have to thank (alphabetically) my sister Anna, my friend Demetra and my friend Pavlos for immense support along the way.

Nomenclature

For consistency the following gene nomenclature is used throughout this thesis:

Gene names are capitalized and italicized. Protein names are capitalized.

With the exception of *Arabidopsis thaliana*, names of genes and proteins include a two letter abbreviation of the species name (e.g. *Selaginella kraussiana* KNOX1 is *SkKNOX1*).

To designate multiple mutants, gene names are written in italicized lower case letters and separated by semi-colons (e.g, *mkn2;mkn4;mkn5*).

Promoter/gene fusions are designated by the italicized names of the gene promoter and coding sequence separated by two consecutive colons (e.g. *pSTM::SkKNOX1*).

Table of Contents

List of figures.....	8
List of abbreviations.....	10
1 Introduction	
1.1 Evolution of land plants	14
1.2 Evolution of the land plant life cycle.....	16
1.2.1 The sporophyte: a key innovation during land plant evolution	18
1.2.1.1 Major steps in sporophyte evolution.....	18
1.3 Indeterminate shoot growth in the sporophyte.....	22
1.3.1 Establishment and maintenance of an indeterminate stem cell population	24
1.3.1.1 The role of <i>KNOX</i> in the establishment and maintenance of the SAM	25
1.3.1.2 The role of <i>WUS</i> in maintenance of the stem cell population in the SAM.....	25
1.3.1.3 The role of hormones in promoting SAM activity	26
1.3.2 Processes involved in the acquisition of determinate fate by stem cell derivatives.....	26
1.3.2.1 Establishment of boundaries between indeterminate and determinate regions of the SAM	27
1.3.2.1.1 SAM and cotyledon boundary	27
1.3.2.1.2 Central zone and peripheral zone boundary	28
1.3.3 Transition from an indeterminate to a determinate cell fate	28
1.3.3.1 The role of auxin in leaf primordium initiation.....	28
1.3.3.2 Importance of <i>KNOX</i> repression in leaf development	29
1.3.3.3 <i>KNOX</i> repression during valve development	29
1.4 The role of <i>KNOX</i> genes in the evolution of the sporophyte	30
1.4.1 The <i>KNOX</i> gene family	30
1.4.2 Significance of gene duplications in the evolution of <i>KNOX</i> gene function.....	33
1.4.3 Two distinct expression patterns of class 1 <i>KNOX</i> genes.....	37
1.5 Evolution of <i>KNOX</i> gene function	40
1.5.1 Function of <i>KNOX</i> genes in green algae.....	40
1.5.2 Function of <i>KNOX</i> genes in mosses	40
1.5.3 <i>KNOX</i> genes in lycophytes	43
1.5.4 <i>KNOX</i> genes in monilophytes	43
1.5.5 <i>KNOX</i> genes in gymnosperms.....	46
1.6 Aims of the project.....	46
2 Material and Methods	48
2.1 General Materials	49
2.2 General equipment.....	49
2.2.1 Centrifuges and rotors.....	49
2.2.2 Microscopy	49
2.2.3 Spectrophotometry	50
2.3 Plant material and growth conditions	50
2.3.1 <i>Arabidopsis</i>	50
2.3.1.1 Growth of <i>Arabidopsis</i> plants on Murashige-Skoog medium	50
2.3.1.2 Growth of <i>Arabidopsis</i> plants in soil	51
2.3.2 <i>P. patens</i>	51
2.3.2.1 <i>P. patens</i> growth medium	51
2.3.2.2 <i>P. patens</i> culture conditions	52
2.3.2.3 <i>P. patens</i> gametophore induction	52
2.3.2.4 <i>P. patens</i> sporophyte induction.....	52
2.3.3 Hornwort strains	53
2.3.3.1 Hornwort growth conditions.....	53
2.3.3.2 Hornwort spore sterilization.....	53
2.3.3.3 Hornwort sporophyte induction	54
2.4 Bacterial material, growth conditions and transformation	54

2.4.1	Bacteria strains	54
2.4.2	Bacterial growth conditions	54
2.4.3	Bacterial methods	55
2.4.3.1	Preparation <i>E. coli</i> competent cells	55
2.4.3.2	Preparation of electrocompetent <i>A. tumefaciens</i> GV3101 cells	56
2.4.3.3	Bacterial transformations	56
2.4.3.3.1	Transformation of <i>E. coli</i> by heat-shock	56
2.4.3.3.2	Transformation of <i>A. tumefaciens</i> by electroporation	56
2.5	Plant transformation	57
2.5.1	Transformation of <i>Arabidopsis</i> using floral dipping	57
2.5.2	Transformation of <i>P. patens</i>	57
2.5.2.1	Preparation of protoplasts	57
2.5.2.2	PEG-based transformation	58
2.5.2.3	Regeneration of protoplasts	59
2.5.2.4	Screening of <i>P. patens</i> transformants	59
2.6	Nucleic acid extraction methods	60
2.6.1	Plasmid DNA	60
2.6.2	<i>P. patens</i> and <i>Arabidopsis</i> genomic DNA extraction	60
2.6.3	Hornwort genomic DNA extraction	61
2.6.4	<i>Arabidopsis</i> , <i>S. kraussiana</i> and <i>P. patens</i> RNA extraction	62
2.6.5	Hornwort RNA extraction	63
2.7	Nucleic acid manipulation	64
2.7.1	Molecular Cloning	64
2.7.2	Ligation of PCR products	64
2.7.3	Site directed mutagenesis	64
2.8	Nucleic acid analysis	65
2.8.1	Polymerase Chain Reaction (PCR)	65
2.8.1.1	Standard PCR	65
2.8.1.2	Colony PCR	66
2.8.1.3	Nested PCR	66
2.8.1.4	Touchdown PCR	67
2.8.1.5	Reverse-transcription PCR (RT-PCR)	67
2.8.2	DNA sequencing	68
2.8.3	Southern blotting	68
2.8.3.1	Genomic DNA restriction digest	68
2.8.3.2	Gel electrophoresis	69
2.8.3.3	Capillary transfer of DNA onto nylon membrane	69
2.8.3.4	DNA Radiolabelling	70
2.8.3.5	Hybridization	70
2.9	Bioinformatics	71
2.9.1	Sequence alignment	71
2.9.2	Phylogenetic analysis	71
3	Comparative analysis of class 1 KNOX gene function in lycophytes, ferns and angiosperms	72
3.1	Introduction	73
3.2	Aims.	73
3.3	Generation of constructs for transformation	74
3.3.1	Expression under the <i>Arabidopsis BP</i> promoter	74
3.3.2	Expression under the <i>Arabidopsis STM</i> promoter	74
3.3.3	Constitutive expression	75
3.4	Generation of transformed lines	79
3.5	Genetic analysis of transgenic lines	81
3.5.1	Analysis of transgene copy number	81
3.5.2	Transgene expression analysis	81
3.6	Phenotypic characterization of plants	84

3.6.1	Cross-species complementation of <i>bp</i> mutants	84
3.6.2	Analysis of homozygous <i>pBP::KNOX</i> lines	84
3.6.3	Analysis of T ₂ segregating <i>pBP::KNOX</i> lines	88
3.7	Discussion	93
4	Comparative analysis of class 1 <i>KNOX</i> genes in bryophytes and vascular plants..	95
4.1	Introduction	96
4.2	Aims.	96
4.3	Generation of constructs for transformation	97
4.4	Generation of transgenic lines	97
4.5	Analysis of transformed lines.....	101
4.5.1	Genetic analysis.....	101
4.5.2	Expression analysis	109
4.6	Phenotypic characterization of transgenic plants.....	111
4.7	Discussion	116
5	Investigation of <i>KNOX</i> genes in hornworts	118
5.1	Introduction	119
5.2	Aims	120
5.3	Establishment of a hornwort experimental system	120
5.3.1	Selection of hornwort species and establishment of axenic cultures	120
5.3.2	Optimisation of growth conditions	121
5.3.3	Establishment of sporophyte induction conditions	121
5.3.4	Development of nucleic acid extraction methods.....	127
5.3.5	Identification of a single <i>KNOX</i> gene in <i>A. punctatus</i>	128
5.3.6	Phylogenetic analysis of <i>ApKNOX</i> gene in land plants.....	128
5.4	Discussion	133
6	General discussion	135
6.1	Gene family origins and expansions during land plant evolution	136
6.2	The role of <i>KNOX</i> genes in the evolution of the SAM.....	136
6.2.1	<i>KNOX</i> gene function in bryophytes	136
6.2.2	Evolution of the SAM in vascular plants.....	137
6.3	Divergent class 1 <i>KNOX</i> gene function in bryophytes and vascular plants	138
6.4	Conserved class 1 <i>KNOX</i> gene function in lycophytes and angiosperms...	139
6.5	Mechanisms for the specialization of <i>KNOX</i> gene function during land plant evolution	140
6.5.1	<i>KNOX</i> gene duplication.....	140
6.5.2	<i>KNOX</i> regulatory element changes.....	140
6.5.3	Evolution of novel <i>KNOX</i> / <i>BELL</i> interactions	141
6.5.4	Sequence changes in open reading frames.....	142
6.6	Future perspectives	143
	References.....	145
	Appendix 1: List of <i>KNOX</i> genes used for <i>KNOX</i> phylogeny.....	157
	Appendix 2: List of primers.....	161

List of figures

Chapter 1

Figure 1.1	Phylogenetic relationships of the major lineages of land plants	15
Figure 1.2	Increase in sporophyte size during land plant evolution	17
Figure 1.3	Key innovations during sporophyte evolution	21
Figure 1.4	<i>Arabidopsis</i> shoot apical meristem with key regulators of SAM activity	23
Figure 1.5	KNOX protein structure and KNOX/BELL interactions.	32
Figure 1.6	KNOX gene phylogeny	34
Figure 1.7	Importance of gene duplications in the evolution of KNOX gene functions	36
Figure 1.8	Expression patterns of class 1 KNOX genes in angiosperms	38
Figure 1.9	Loss of function class 1 KNOX mutant phenotypes	39
Figure 1.10	Life cycle of the moss <i>Physcomitrella patens</i> and the phenotype of <i>P. patens</i> class 1 KNOX mutants	42
Figure 1.11	Comparison of class 1 KNOX gene expression patterns in <i>Selaginella kraussiana</i> and <i>Arabidopsis thaliana</i>	45

Chapter 3

Figure 3.1	Schematic representation of <i>pBP::KNOX</i> constructs	76
Figure 3.2	Schematic representation of <i>pSTM::KNOX</i> constructs	77
Figure 3.3	Schematic representation of <i>35S::KNOX</i> constructs	78
Figure 3.4	Southern blot analysis of <i>pBP::KNOX</i> constructs	82
Figure 3.5	Transcript accumulation in <i>pBP::KNOX</i> lines	83
Figure 3.6	Phenotypes of homozygous <i>pBP::KNOX</i> lines	85
Figure 3.7	Average silique angle of homozygous <i>pBP::KNOX</i> lines	86
Figure 3.8	Average height of homozygous <i>pBP::KNOX</i> lines	87
Figure 3.9	Phenotypes of T ₂ segregating <i>pBP::KNOX</i> lines	90
Figure 3.10	Average silique angle of T ₂ segregating <i>pBP::KNOX</i> lines	91
Figure 3.11	Average height of T ₂ segregating <i>pBP::KNOX</i> lines	91

Chapter 4

Figure 4.1	Schematic representation of <i>pMKN2::KNOX</i> constructs	99
Figure 4.2	Schematic representation of the generation of <i>pMKN2::KNOX</i> lines	100
Figure 4.3	Schematic representation of <i>pMKN2::KNOX</i> constructs targeted into the <i>MKN2</i> locus	103
Figure 4.4	Southern blot analysis of <i>pMKN2::KNOX</i> lines	104
Figure 4.5	Southern blot analysis of <i>pMKN2::KNOX</i> lines	105
Figure 4.6	Southern blot analysis of <i>pMKN2::KNOX</i> lines	106
Figure 4.7	Southern blot analysis of <i>pMKN2::KNOX</i> lines	107
Figure 4.8	Southern blot analysis of <i>pMKN2::KNOX</i> lines	107
Figure 4.9	Southern blot analysis of <i>pMKN2::SkKNOX2</i> lines	108
Figure 4.10	Transcript accumulation in <i>pMKN2::KNOX</i> lines	110
Figure 4.11	Phenotypes of <i>pMKN2::KNOX</i> lines	113
Figure 4.12	Average ratio of total sporophyte surface area to the area occupied by spores, for <i>pMKN2::KNOX</i> lines	115

Chapter 5

Figure 5.1	Life cycle of the hornwort <i>Anthoceros punctatus</i>	123
Figure 5.2	Life cycle of the hornwort <i>Anthoceros agrestis</i>	124
Figure 5.3	Sequence alignment of <i>ApKNOX</i>	129
Figure 5.4	<i>ApKNOX</i> gene phylogeny	132

List of abbreviations

ABA	Absciscic acid
<i>ARP</i>	<i>ASYMMETRIC LEAF1</i> <i>ROUGH SHEATH2</i> <i>PHANTASTICA</i>
ATP	Adenine tri-phosphate
BELL	BEL1-like
<i>BP</i>	<i>BREVIPEDICELLUS</i>
<i>C. reinhardtii</i>	<i>Chlamydomonas reinhardtii</i>
<i>C. richardii</i>	<i>Ceratopteris richardii</i>
CaMV	Cauliflower mosaic virus
cDNA	Complementary DNA
CK	Cytokinin
<i>CrKNOX1</i>	<i>C. richardii</i> <i>KNOTTED1-like</i> <i>HOMEODOMAIN 1</i>
<i>CrKNOX2</i>	<i>C. richardii</i> <i>KNOTTED1-like</i> <i>HOMEODOMAIN 2</i>
<i>CrKNOX3</i>	<i>C. richardii</i> <i>KNOTTED1-like</i> <i>HOMEODOMAIN 3</i>
CTAB	Cetyl trimethylammonium bromide
<i>CUC1</i> , 2 and 3	<i>CUPSHAPED COTYLEDON 1</i> , 2 and 3
dATP	Deoxyadenosine tri-phosphate
dCTP	Deoxycytidine tri-phosphate
DEPC	Diethyl pyrocarbonate
dGTP	Deoxyguanine tri-phosphate
dNTP	Deoxynucleic tri-phosphate
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylene diaminetetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
FM	Floral meristem
GA	Gibberellic acid
<i>GA20ox1</i>	<i>GA 20-OXIDASE1</i>
<i>GA2ox1</i>	<i>GA 2-OXIDASE1</i>
<i>Glk</i>	<i>Golden 2-like</i>
GSM1	GAMETE SPECIFIC MINUS 1
GSP1	GAMETE SPECIFIC PLUS 1
GUS	β-glucuronidase

h	Hour
Hyg	Hygromycin
IM	Inflorescence meristem
<i>IPT7</i>	<i>ISOPENTENYL TRANSFERASE7</i>
Kan	Kanamycin
kan ^R	Kanamycin resistant
Kan ^S	Kanamycin sensitive
Kb	Kilobase
<i>KN1</i>	<i>KNOTTED1</i>
<i>KNAT1</i>	<i>Kn1-like Arabidopsis thaliana 1</i>
<i>KNAT2</i>	<i>Kn1-like Arabidopsis thaliana 2</i>
<i>KNAT6</i>	<i>Kn1-like Arabidopsis thaliana 6</i>
<i>KNOX</i>	<i>KNOTTED1-like HOMEOBOX</i>
kV	Kilovolt
L1	Layer 1
L2	Layer 2
L3	Layer 3
LB	Luria Bertani
Lux	Luminous flux
M	Molar
MES	2-(N-morpholino)ethanesulfonic acid
Min	Minute
<i>MKN1-3</i>	<i>Moss KNOTTED1-like 1-3</i>
<i>MKN2</i>	<i>Moss KNOTTED1-like 2</i>
<i>MKN4</i>	<i>Moss KNOTTED1-like 4</i>
<i>MKN5</i>	<i>Moss KNOTTED1-like 5</i>
<i>MKN6</i>	<i>Moss KNOTTED1-like 6</i>
mM	Millimolar
<i>nptII</i>	<i>Neomycin phosphotransferase II</i>
OD	Optical density
<i>OSH1</i>	<i>ORYZA SATIVA HOMEOBOX 1</i>
<i>OSH15</i>	<i>ORYZA SATIVA HOMEOBOX 15</i>
<i>P. patens</i>	<i>Physcomitrella patens</i>
P0	Plastochron 0
P1	Plastochron 1
PCR	Polymerase chain reaction
PEG	Polyethylen glycol

PVP-40	Polyvinylpyrrolidone (average molecular weight 40,000)
RT-PCR	Reverse transcription polymerase chain reaction
<i>S. kraussiana</i>	<i>Selaginella kraussiana</i>
SAM	Shoot apical meristem
<i>SkARP</i>	<i>S. kraussiana ARP</i>
<i>SkKNOX1</i>	<i>S. kraussiana KNOTTED1-like HOMEODOMAIN 1</i>
<i>SkKNOX2</i>	<i>S. kraussiana KNOTTED1-like HOMEODOMAIN 2</i>
<i>STM</i>	<i>SHOOTMERISTEMLESS</i>
TALE	Three amino acid loop extension
TE	Tris-EDTA
Tet	Tetracycline
TF	Transcription factor
UV	Ultra violet
V	Volt
Wt	Wild type
<i>WUS</i>	<i>WUSCHEL</i>
μM	Micromolar

1 Introduction

1.1 Evolution of land plants

The plant lineage (*Viridiplantae*) first emerged 1.3 billion years ago, and followed two evolutionary trajectories: the *Chlorophyta*, comprising the majority of green algae, and the *Streptophyta* comprising the charophytes (*Charophyta*) and the land plants (*Embryophyta*) (Leliaert et al., 2012). As indicated by biochemical, morphological and molecular data, land plants evolved from charophycean green algae (Karol et al., 2001; Lewis and McCourt, 2004; Becker and Marin, 2009). For example, cell cleavage by phragmoplasts and plasmodesmatal connections between cells are found only in charophytes and land plants, and the biopolymer sporopollenin is found only in charophyte zygotes and land plant spores (Graham et al., 2000; Becker and Marin, 2009).

The first evidence of land colonization by plants is the presence of fossil spores found in middle Ordovician strata (Gensel, 2008; Rubinstein et al., 2010). This suggests that land plants evolved at least 480 million years ago. Colonization of land by plants led to their diversification into the following taxa: the bryophytes and the vascular plants, with vascular plants comprising lycophytes, monilophytes, gymnosperms and angiosperms (Figure 1.1). All of the vascular plant groups are monophyletic whereas the bryophytes are a poorly-resolved paraphyletic group in which liverworts are the most basal and hornworts are the group closest to vascular plants; the moss lineage falls in between the two (Qiu et al., 2006).

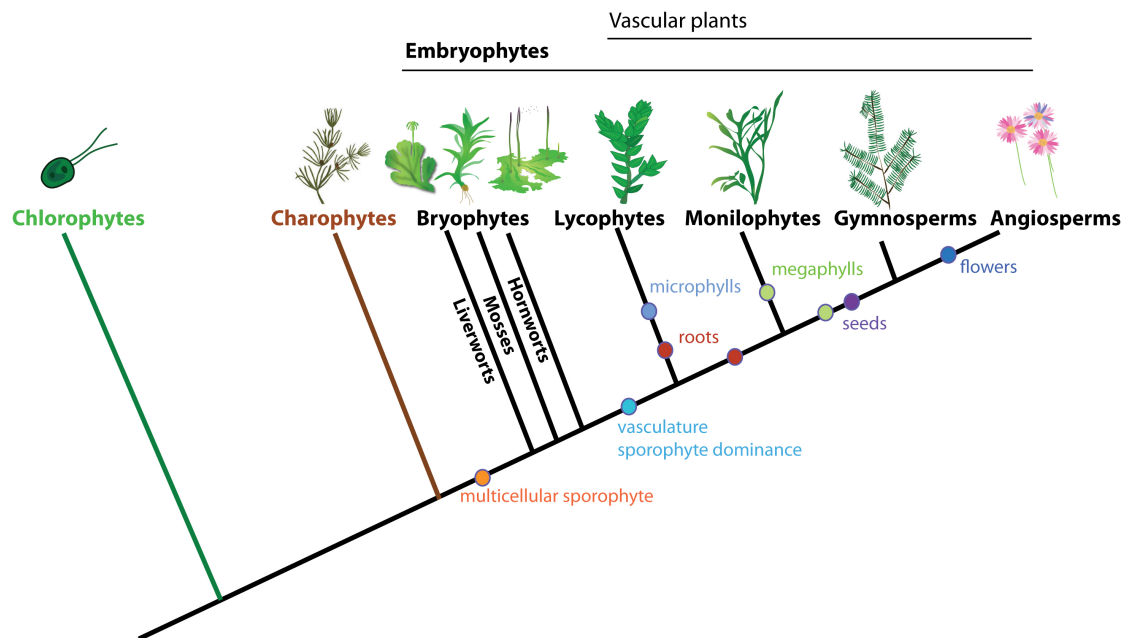


Figure 1.1

Phylogenetic relationships of the major lineages of land plants (embryophytes) illustrating the paraphyly of bryophytes and the monophyly of lycophytes, monilophytes, gymnosperms and angiosperms. Key innovations during land plant evolution are indicated with filled circles. Two types of leaves evolved in vascular plants: the microphylls and the megaphylls. Microphylls are morphologically simpler than megaphylls and evolved within the lycophytes. Megaphyll evolution occurred at least two times, once within monilophytes and a second time before the diversification of seed plants (gymnosperms and angiosperms) (Langdale, 2008). Root evolution also occurred at least twice, once within lycophytes and once before the diversification of monilophytes and seed plants (Langdale, 2008).

1.2 Evolution of the land plant life cycle

In charophytes, the sister group to land plants, the majority of the life cycle is represented by the haploid gametophyte and only the unicellular zygote is diploid (Karol et al., 2001; Niklas and Kutschera, 2010). In contrast, all extant land plant lineages are characterized by the alternation of two multicellular generations: the haploid gametophyte and the diploid sporophyte. Gametes are produced during the gametophyte generation and the fertilization of eggs by sperm results in a diploid zygote, which subsequently develops into a multicellular sporophyte. Spores are produced within the sporophyte generation and upon germination, spores develop into gametophytes.

By comparing the life cycles of early diverging land plants to those that have diverged more recently it is obvious that there is an evolutionary trend towards increasing dominance of the sporophyte generation and reduction of the gametophyte (Figure 1.2). In bryophytes the gametophyte is the dominant phase of the life cycle and the sporophyte is dependent on the gametophyte for nourishment. In contrast, the gametophytes of angiosperms are dependent on the dominant sporophyte (Niklas and Kutschera, 2009).

In all cases, the sporophyte is diploid and the gametophyte is haploid. Thus the sporophyte can mask deleterious mutations through the dominant-recessive interaction of alleles, whereas the gametophyte cannot.

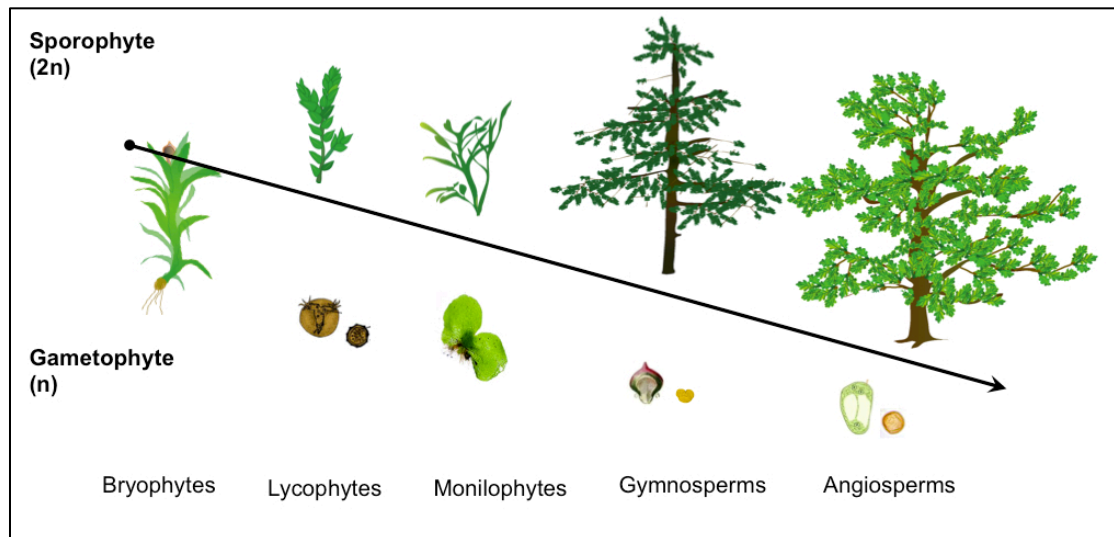


Figure 1.2

Schematic representation of the increase in sporophyte size and complexity during land plant evolution. All bryophytes have a sporophyte that is small, structurally simple and dependent on the gametophyte for nutrition. Vascular plants evolved a free-living sporophyte that progressively increased in complexity, whereas the gametophyte reduced to just a few cells in seed plants. In flowering plants the male and female gametophytes are the pollen and the embryo sac respectively.

1.2.1 The sporophyte: a key innovation during land plant evolution

The terrestrial environment is dramatically different from aquatic environments. For example, reduced water availability, increased temperature variation and greater exposure to radiation are major issues that must be addressed. For organisms that evolved in water the transition to terrestrial habitats was thus really challenging (Bowman et al., 2007; Bennici, 2008).

1.2.1.1 Major steps in sporophyte evolution

The land plant sporophyte has undergone major morphological and physiological changes during evolution. It was transformed from a developmentally simple structure to a highly variable and often elaborate structure with an extraordinary array of organs and tissue systems (Figure 1.3A).

Evidence of the morphology of early land plants comes from the discovery of sporangia fragments that resemble the sporangia of extant liverworts (Wellman et al., 2003). These fossils suggest that ancestral land plants were morphologically similar to extant bryophytes in which sporophytes have a single axis of growth, forming an unbranched structure that terminates in a sporangium containing haploid spores.

As indicated by protracheophyte (extinct precursors of vascular plants) fossil sporophytes such as *Aglaophyton major* (Crane et al., 2004), the next step in the evolution of sporophyte morphology was branching. Another key step was apical shoot growth from one or more initial cells. This pattern of growth is different from that found in bryophytes where diverse strategies for sporophyte growth exist. For example, liverwort sporophytes develop from cell divisions throughout the structure and moss sporophytes develop from both a transient apical cell and basal meristem (Bowman, 2013).

Acquisition of branching and apical growth opened the way for the evolution of more

complex body plans which, along with competition for light, led to the evolution of larger plants (Langdale, 2008). As a consequence, the need for efficient transport of water and nutrients, and for mechanical support intensified. These drivers led to the evolution of conducting cells (tracheids) and subsequently to a vascular system. The precise morphology of early vascular plants is not clear but fossils of the first known vascular plant, *Cooksonia*, provide some insight (Edwards and Feehan, 1980). *Cooksonia* had leafless branching axes bearing multiple sporangia at the tips.

During the evolution of vascular plants, sporophyte complexity further increased. Roots evolved to improve both transport capability and mechanical support, and leaves evolved to improve photosynthetic efficiency (Figure 1.3B). It has been suggested that the evolution of a sporophyte with photosynthetic and reproductive potential intensified the conflict of interest between the gametophyte and the sporophyte in terms of water and nutrient allocation (Haig and Wilczek, 2006). This could be one of the main factors that led to the evolution of a free-living autonomous sporophyte in vascular plants.

The next important step in sporophyte evolution was the appearance of seeds. The first land plants were homosporous, meaning that the spores they produced all had the same size and the same developmental potential. The transition to plants producing two types of spores (heterosporous) paved the way to seed evolution. It is believed that larger spores (megaspores) were the precursors of ovules whereas smaller spores (microspores) were the precursors of pollen. Initially, there was a reduction in the number of megaspores to one that was retained in the sporangium (megasporangium). Later, an outer protective coat (seed coat) evolved that conferred protection from desiccation (Stewart and Rothwell, 1993) (Figure 1.3C). The evolution of seeds allowed plants to proliferate and spread to drier areas of Earth. Within

angiosperms, sepals and petals evolved and allowed for specialization to attract animal pollinators (Figure 1.3D).

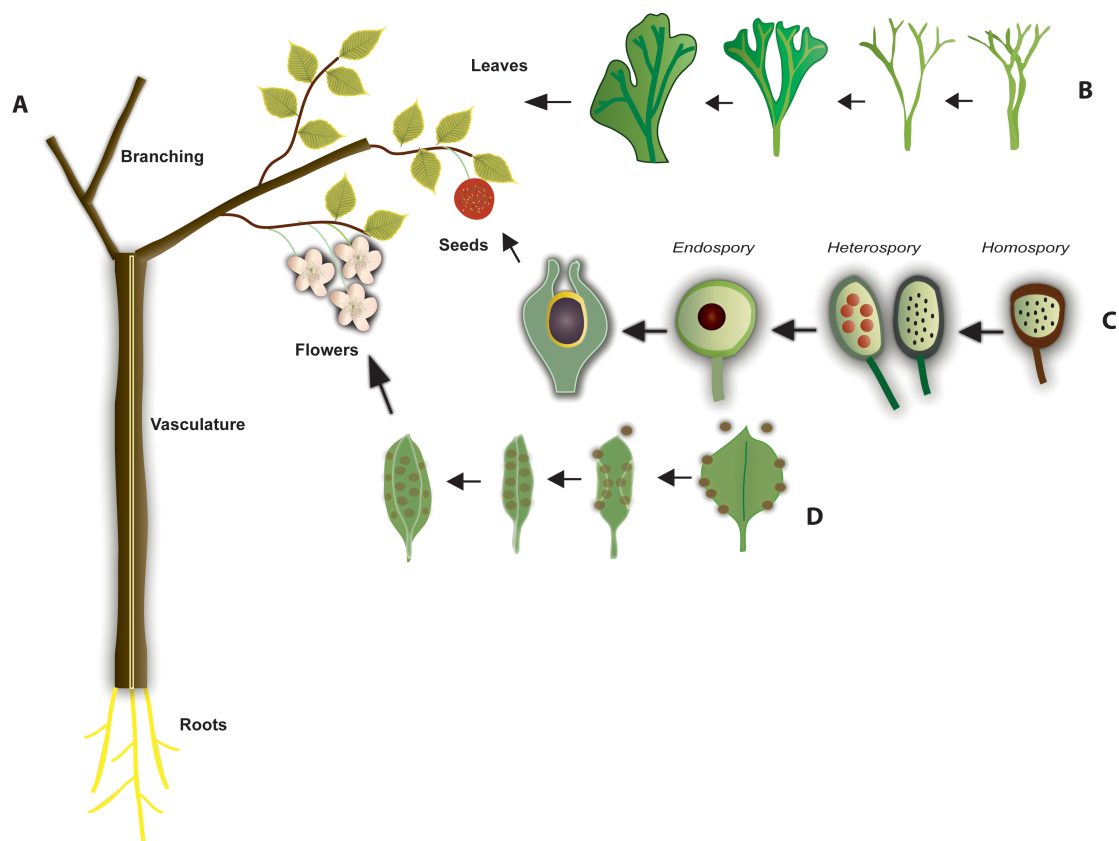


Figure 1.3

A) Schematic representation of key innovations during sporophyte evolution: vasculature, roots, leaves, seeds and flowers. B) Possible scenario for leaf evolution. Fossil evidence suggests that megaphylls (leaves of monilophytes and seed plants) evolved from photosynthetic branching systems with the development of a lamina between different branches (adapted from Kenrick, 2001). A similar modification probably led to the evolution of microphylls. C) Possible scenario for seed evolution. The first step towards the evolution of seeds was the production of two types of spores (microspores and megaspores) called heterospory. The reduction to a single functional megaspore per sporangium and its retention within the sporangium (endospory) was the next step. Finally, the development of a seed coat (integument) facilitated the evolution of seeds (Linkies et al., 2010). D) Possible scenario for carpel and flower evolution. A modified leaf bearing sporangia folds and then the margins fuse to form a carpel or simple pistil found in some species. Compound pistils found in other species are thus the product of the fusion of several carpels (adapted from Purves et al., 1995).

1.3 Indeterminate shoot growth in the sporophyte

The capacity for continuing organogenesis and growth throughout the life of a plant is called indeterminate growth, whereas determinate growth is temporally and/or spatially restricted. Indeterminate growth confers morphological flexibility and improved survival potential in response to adverse environmental conditions (Rowe and Speck, 2005). Shoot growth of vascular plants is indeterminate whereas shoot growth in the sporophyte of bryophytes is determinate. Shoot growth in vascular plants results from the action of a single stem cell (most lycophytes, monilophytes and gymnosperms) or a population of multiple stem cells (angiosperms) at the apex of the shoot (Langdale and Harrison, 2008). In angiosperms this pluripotent population of stem cells is called the shoot apical meristem (SAM).

In the model angiosperm *Arabidopsis thaliana* the SAM is established in early embryogenesis. Embryogenesis starts with the zygote dividing asymmetrically to form an apical and a basal cell which each have different developmental fates. The apical cell undergoes several transverse and longitudinal cell divisions and gives rise to the globular embryo. Further rounds of cell division result in the formation of the heart stage embryo where the cotyledon primordia and the SAM become distinct (Barton and Poethig, 1993; Ueda and Laux, 2012).

The SAM is organized into three layers based on cell division planes and contribution to the formation of different shoot tissues (Figure 1.4). Layer 1 (L1) is the outermost layer that divides anticlinally (perpendicular to the nearest layer surface) and forms the epidermis. Layer 2 (L2) is the sub-epidermal layer that predominantly divides anticlinally and layer 3 (L3) is the innermost layer that divides in all planes. Both L2 and L3 layers contribute different numbers of cells to different plant organs (Poethig, 1987). The SAM is also divided into functional zones: the central zone which contains

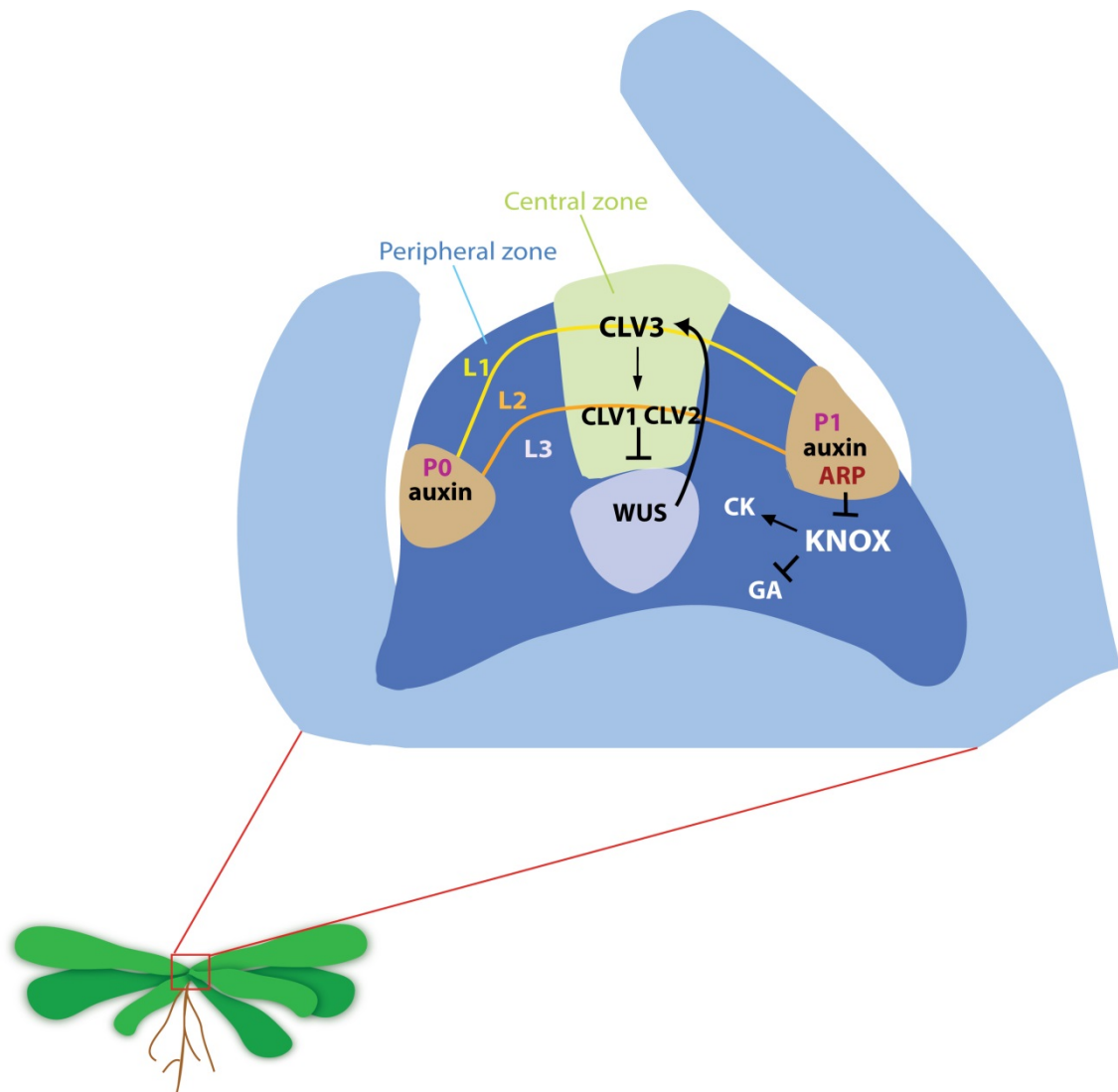


Figure 1.4

Schematic representation of a longitudinal section of the *Arabidopsis* shoot apical meristem (SAM) with key regulators of SAM activity. L1: layer 1, L2: layer 2 and L3: layer 3. Central zone: green. Peripheral zone: dark blue. *WUS* expression zone: light blue. P0: plastochron 0, P1: plastochron 1. *KNOX* expression that marks the entire SAM except leaf primordia is not indicated for simplicity. Modified from Barton, 2010.

the stem cell population, the peripheral zone where organ initiation occurs and the rib zone that contributes to the stem.

During the transition from vegetative to reproductive development, the SAM changes its identity. It converts into the inflorescence meristem (IM) which instead of leaf primordia initiates floral meristems (FM) in the peripheral zone. In contrast to the SAM and IM, the FM is determinate and meristematic activity ceases after the initiation of carpels (Ha et al., 2010; McKim and Hay, 2010). In *Arabidopsis* two fused carpels form the female reproductive organ (gynoecium) which is composed of the ovary, the style, and the stigma. The ovary comprises of three tissue types: valves, replum and valve margins. A region of the internal surface of the replum called the placenta reinitiates meristematic activity to form ovules (Robles and Pelaz, 2005).

Stem cells in the SAM are continuously self-renewing and at the same time they produce derivatives that differentiate into plant organs such as leaves and flowers. Mechanisms controlling SAM function must thus balance two opposing activities:

1. Establishment and maintenance of an indeterminate stem cell population.
2. Production of determinate organs.

1.3.1 Establishment and maintenance of an indeterminate stem cell population

In *Arabidopsis*, it has been shown that *KNOTTED1-like HOMEODOMAIN* (*KNOX*) and *WUSCHEL* (*WUS*) genes play a critical role in the establishment and maintenance of SAM indeterminacy, in combination with the plant hormones cytokinin (CK) and gibberellin (GA).

1.3.1.1 The role of KNOX in the establishment and maintenance of the SAM

The *KNOX* gene *SHOOTMERISTEMLESS* (*STM*) is first expressed during the globular stage of embryo development in the region where the SAM will be formed. In vegetative development, *STM* expression marks the entire SAM with the exception of regions where new organs will be formed (Clark et al., 1996; Long et al., 1996). Strong loss-of-function *stm* mutants show no signs of early SAM development and germinating seedlings terminate growth with differentiated shoot apices and fused cotyledons (Endrizzi et al., 1996). Weak *stm* mutants lack a SAM and form a single leaf primordium, suggesting that the stem cells have been consumed by the formation of that primordium (Endrizzi et al., 1996). These mutant phenotypes and expression patterns indicate that *STM* is required for the establishment and maintenance of the SAM.

Two further *KNOX* genes, *BREVIPEDICELLUS* (*BP*) and *Kn1-like Arabidopsis thaliana 6* (*KNAT6*), have a redundant role with *STM* in the maintenance of the SAM. When either *BP* or *KNAT6* are mutated in weak *stm* mutants, residual meristematic activity is abolished or further reduced (Byrne et al., 2002). Consistent with the role of *KNOX* genes in maintaining indeterminacy in the SAM, *STM* and *BP* are expressed in the replum (Alonso-Cantabrana et al., 2007) and when levels of *STM* expression are reduced, defects in placenta development are observed (Scofield et al., 2007).

1.3.1.2 The role of WUS in maintenance of the stem cell population in the SAM

WUS is first expressed during embryogenesis, at the 16 cell stage, in the apical domain of the embryo (Mayer et al., 1998). During vegetative development *WUS* is expressed below the central zone of the SAM (Mayer et al., 1998) (Figure 1.4) and in *wus* loss-of-function mutants the central zone of the SAM cannot be maintained (Laux

et al., 1996). Thus, *WUS* is necessary for maintaining the stem cell population of the SAM. Given that the region of the SAM that is affected in *wus* mutants is not the same as the region where the gene is expressed, WUS action on stem cell activity is non-cell-autonomous.

1.3.1.3 The role of hormones in promoting SAM activity

CK and GA are plant hormones involved in SAM activity. CK activates cell division, whereas GA promotes cell elongation and differentiation. A high cytokinin to low gibberellin ratio is important to maintain the SAM in an indeterminate state (Hamant and Pautot, 2010). CK levels are increased in plants overexpressing *STM* and it has also been shown that KNOX proteins upregulate expression of the cytokinin biosynthesis gene *ISOPENTENYL TRANSFERASE7 (IPT7)* (Jasinski et al., 2005; Yanai et al., 2005). Consistent with this observation, application of CK can rescue the *stm* mutant phenotype (Jasinski et al., 2005). Conversely, KNOX proteins regulate gibberellin levels by directly binding to promoters of the biosynthetic gene, *GA 20-OXIDASE1 (GA20ox1)* and the catabolic gene *GA 2-OXIDASE1 (GA2ox1)* (Sakamoto et al., 2001; Bolduc and Hake, 2009). Collectively, these observations demonstrate that KNOX proteins regulate the balance between GAs and CKs in the meristem to control stem cell activity.

1.3.2 Processes involved in the acquisition of determinate fate by stem cell derivatives

The initiation of organs (organogenesis) by the SAM requires the establishment of boundaries between indeterminate and determinate growth regions, and the regulation of indeterminate versus determinate cell fate.

1.3.2.1 Establishment of boundaries between indeterminate and determinate regions of the SAM

1.3.2.1.1 SAM and cotyledon boundary

In eudicot species (angiosperms that have two embryonic leaves (cotyledons) as opposed to the single leaf found in monocots), a boundary between the two cotyledons needs to be established during embryogenesis. This boundary is of great importance because it is the site where the SAM develops. In *Arabidopsis*, three redundant genes *CUPSHAPED COTYLEDON 1*, *2*, and *3* (*CUC1*, *2* and *3*) that encode NAC (*NO APICAL MERISTEM*, *ATAF* and *CUP-SHAPED COTYLEDON*) transcription factors are required for cotyledon separation (Olsen et al., 2005). *CUC* genes are first expressed throughout the entire apical portion of the embryo but expression is subsequently restricted to the boundary between the SAM and cotyledons. Plants lacking functional *CUC* genes form a fused goblet-shaped cotyledon and fail to specify a SAM (Aida et al., 1997; Takada et al., 2001; Vroemen et al., 2003; Hibara et al., 2006). The overexpression of *CUC* genes results in ectopic *STM* expression in cotyledons while loss of *CUC* function results in loss of *STM* expression (Aida et al., 1999; Takada et al., 2001). This suggests that *CUC* genes play a role in the formation of the SAM by activating *STM*.

Similar to *CUC*, *STM* is initially expressed throughout the apical region of the embryo before being localized to the SAM. Loss of *STM* activity results in expansion of the *CUC* gene expression domain (Aida et al., 1999). Indicating that *STM* participates in the formation of the boundary between cotyledons by repressing *CUC* genes. It has also been shown that *KNAT6* acts redundantly with *STM* to establish the cotyledon boundary (Belles-Boix et al., 2006). The boundary between determinate cotyledons and the indeterminate SAM is thus maintained by a feedback loop between *CUC* and

KNOX genes. *CUC* activates *KNOX*, and *KNOX* spatially restricts the expression domain of *CUC* to the boundary region.

1.3.2.1.2 Central zone and peripheral zone boundary

Organogenesis takes place in the peripheral zone of the SAM. However, cells that differentiate into the various organs are initially recruited from the central zone. This raises the issue of how the boundary between the central and the peripheral zone is established. In *Arabidopsis* there are three *CLAVATA* genes, *CLV1*, *CLV2* and *CLV3* that fulfill this role. *CLV3* encodes a small secreted protein and is expressed in the L1 and L2 layers of the central zone while *CLV1* and *CLV2* genes encode receptor-like proteins and are expressed mainly in the L3 (Clark et al., 1997; Jeong et al., 1999). *CLV3* interacts with *CLV1* and *CLV2* (Muller et al., 2008) to limit the *WUS* expression zone and consequently to limit the size of the stem cell/central zone (Reddy and Meyerowitz, 2005) (Figure 1.4). This interaction is revealed by the expansion of *WUS* expression zone and increased meristem size in *clv* loss-of-function mutants (Brand et al., 2000; Schoof et al., 2000). *WUS* positively regulates *CLV3* activity, generating a feedback loop that stabilizes stem cell activity in the SAM in an analogous manner to the *CUC/KNOX* feedback loop in the embryo.

1.3.3 Transition from an indeterminate to a determinate cell fate

1.3.3.1 The role of auxin in leaf primordium initiation

Leaf formation on the flanks of the SAM initially requires the recruitment of leaf founder cells from the peripheral zone. Notably, the emergence of leaf primordia is associated with regions of high auxin accumulation (Reinhardt et al., 2003) (Figure 1.4) and auxin transport has been implicated in the positioning of leaves about the stem (phyllotaxy) (Benkova et al., 2003; Reinhardt et al., 2003). The emergence of a new leaf primordium on the flanks of the SAM changes the distribution of auxin in the

surrounding area, and as a result only regions distant from an existing leaf are able to accumulate auxin and initiate a subsequent primordium.

1.3.3.2 Importance of *KNOX* repression in leaf development

Once a new leaf primordium has been initiated at plastochron zero (P0) (where a plastochron is the time interval between initiation of successive primordia), repression of *KNOX* gene expression is necessary for determinate leaf development (Hay and Tsiantis, 2010). The initial repression at P0 appears to be mediated by an unknown post-transcriptional mechanism because when native *KNOX* genes are constitutively overexpressed in *Arabidopsis*, repression still occurs. From P1 onwards, however, the *KNOX*-off state is maintained by the activity of *ARP* genes (named after *Arabidopsis ASYMMETRIC LEAF1* (*AS1*), maize *ROUGH SHEATH2* (*RS2*), and *Antirrhinum PHANTASTICA*) that encode MYB-type transcription factors (Timmermans et al., 1999; Tsiantis et al., 1999; Byrne et al., 2002; Tattersall et al., 2005). The repression of *KNOX* gene expression in leaf primordia results in a low CK to high GA ratio that promotes the transition from an indeterminate to a determinate cell fate (Jasinski et al., 2005; Hay and Tsiantis, 2010).

1.3.3.3 *KNOX* repression during valve development

The *ARP* gene *AS1* also restricts expression of the *KNOX* gene *BP* during carpel development. Repression of *BP* in the replum is necessary to promote correct valve differentiation, suggesting the presence of a conserved *KNOX/ARP* regulatory mechanism in the SAM and FM (Alonso-Cantabrana et al., 2007). Given that carpels are thought to have evolved from modified leaves (Weigel and Clark, 1996) it is possible that the *KNOX/ARP* module was recruited from a SAM/leaf pathway in non-flowering plants to operate in a FM/carpel context in flowering plants.

1.4 The role of *KNOX* genes in the evolution of the sporophyte

1.4.1 The *KNOX* gene family

KNOX genes encode homeodomain-containing transcription factors (TF). The homeodomain is a DNA-binding 60 amino acid motif consisting of three helices (Qian et al., 1989; Billeter et al., 1993). Homeodomain TFs were first discovered in the fruit fly *Drosophila melanogaster* (McGinnis et al., 1984) and subsequently in other animals, fungi and plants (Gehring, 1987; Gehring, 1992; Gehring, 1993; Derelle et al., 2007). In all cases they have been shown to be key regulators of developmental processes (Derelle et al., 2007; Hay and Tsiantis, 2010). *KNOX* genes belong to the three amino acid loop extension (TALE) superclass of homeodomain TFs that are characterized by three extra amino acids between homeodomain helix 1 and helix 2 (Bertolino et al., 1995; Burglin, 1997).

Maize *KNOTTED1* (*KN1*) was the first homeobox gene discovered in plants (Hake et al., 1989; Vollbrecht et al., 1991). Since the characterization of *KN1*, *KNOX* genes have been discovered in all plant species studied, encompassing lineages as phylogenetically diverse as mosses and angiosperms (Hay and Tsiantis, 2010). *KNOX* genes can be divided into two main subclasses based on sequence similarity, intron position and expression patterns: class 1 *KNOX* genes (e.g. *STM*, *BP* and *KNAT6*), which have high sequence similarity to *KN1*; and class 2 *KNOX* genes (Kerstetter et al., 1994). Both *KNOX* gene classes have common structural characteristics such as a MEINOX domain upstream of the homeodomain that is composed of two blocks (*KNOX1* and *KNOX2*) (Burglin, 1997), and a shorter ELK (Glutamic acid (E), Leucine (L), Lysine (K)) motif adjacent to the homeodomain (Hamant and Pautot, 2010) (Figure 1.5A). Recently a eudicot-specific third class of *KNOX* genes was identified (Kimura et al., 2008; Magnani and Hake, 2008). Class 3 has a unique member (called

KNATM in *Arabidopsis* and *PETROSELINUM* in tomato (*Solanum lycopersicon*)) that does not contain a homeodomain.

KNOX proteins interact with a second group of TALE proteins, known as the BEL1-like (BELL) homeodomain proteins (Bellaoui et al., 2001; Muller et al., 2001; Smith et al., 2002; Hackbusch et al., 2005; Cole et al., 2006). KNOX/BELL interactions depend on the MEINOX domain (Figure 1.5B) and different heterodimer combinations regulate different downstream genes (Hamant and Pautot, 2010).

A



B

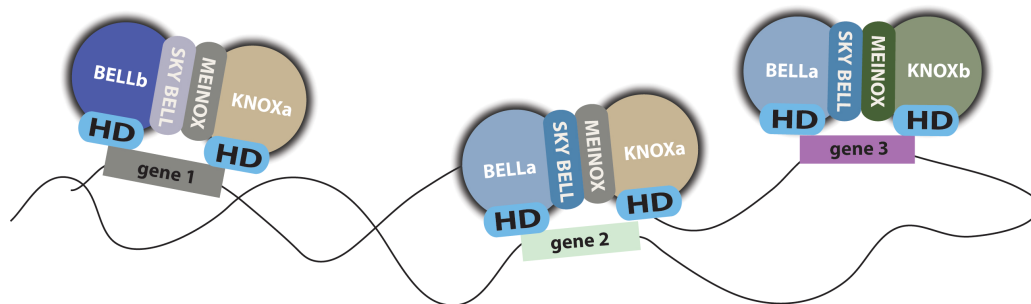


Figure 1.5

Schematic representation of KNOX protein structure and KNOX/BELL interactions. A) KNOX protein structure showing conserved functional domains. The homeodomain (HD), which has three extra amino acids (Proline (P) – Tyrosine (Y) – Proline (P) – indicated by a red arrowhead above the HD) between the first and the second helix, is located in the C-terminus of the protein. Adjacent to the homeodomain is the ELK domain which probably contains a nuclear localisation signal (Hamant and Pautot, 2010). The GSE (Glycine (G), Serine (S), Glutamic acid (E)) domain has been implicated in regulating protein stability (Hamant and Pautot, 2010). The MEINOX domain mediates interactions with BELL proteins and is composed of two subdomains, KNOX1 and KNOX2 (Sakamoto et al., 1999). B) BELL proteins contain a MEINOX interacting domain composed of the SKY and BELL regions. Depicted is a hypothetical example of how different KNOX/BELL heterodimers bind to different gene targets. The same KNOX protein (e.g. KNOXa) can form heterodimers with different BELL proteins (e.g. BELLa or BELLb) and thus regulate different gene targets (gene 1 and gene 2). Similarly, the same BELL protein (BELLa) can form heterodimers with different KNOX proteins (e.g. KNOXa or KNOXb) and thus regulate different genes (gene 2 and gene 3).

1.4.2 Significance of gene duplications in the evolution of *KNOX* gene function

Phylogenetic analysis of *KNOX* genes indicates a monophyletic origin (Bharathan et al., 1999; Reiser et al., 2000) (Figure 1.6). Representatives of both class 1 and class 2 *KNOX* genes have been found in early diverging land plants such as mosses. However, only one *KNOX* gene, that possesses features of class 2 *KNOX* genes, was isolated from chlorophytes (Lee et al., 2008). A single *KNOX* gene has also been identified in the liverwort *Marchantia polymorpha* (Dolan, personal communication). These observations suggest that an ancient duplication occurred after the colonization of land and probably after the divergence of liverworts and mosses (Table 1.1).

Further *KNOX* gene duplications took place during the evolution of land plants as indicated by the greater number of *KNOX* genes within more recently diverged lineages (Table 1.1 and Figure 1.6). The number of *KNOX* genes is greatest in angiosperms. These gene duplications, possibly coupled with regulatory element changes (Figure 1.7) would have provided a mechanism for significant divergence of *KNOX* gene function during land plant evolution.

class 1 KNOX

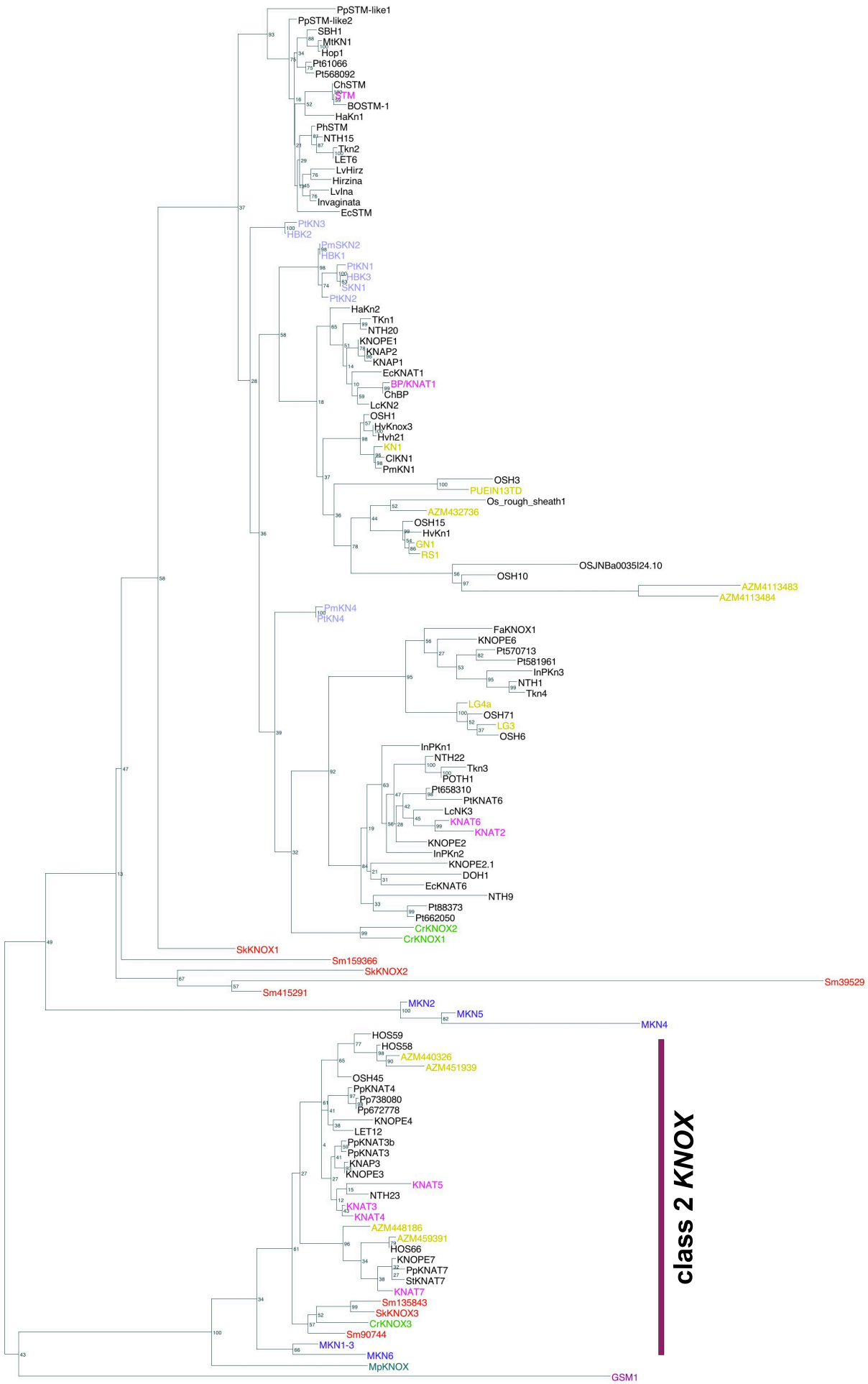


Figure 1.6.

KNOX gene phylogeny. Different species or groups are highlighted with different colours: *Chlamydomonas reinhardtii* (purple) *Marchantia polymorpha* (dark green) *Ceratopteris richardii* (green) lycophytes (red) gymnosperms (light blue), maize (yellow) and *Arabidopsis* (pink). Numbers on the branches represent bootstrap support values. Scale bar: mean number of nucleotide substitutions per site. Species and sequences used in the phylogenetic analysis are listed in Appendix 1.

	Chlorophytes	Bryophytes		Lycophytes	
	<i>C. reinhardtii</i>	<i>M. polymorpha</i>	<i>P. patens</i>	<i>S. kraussiana</i>	<i>S. moellendorffii</i>
Class 1 <i>KNOX</i>	0	?	3	2	3
Class 2 <i>KNOX</i>	1	1	2	1	2
<i>BELL</i>	1	1	4	?	2

	Monilophytes	Monocots		Eudicots	
	<i>C. richardii</i>	Maize	Rice	<i>Arabidopsis</i>	Poplar
Class 1 <i>KNOX</i>	2	9	8	4	6
Class 2 <i>KNOX</i>	1	4	4	4	9
<i>BELL</i>	?	17	14	13	19

Table 1.1

Numbers of class 1 *KNOX*, class 2 *KNOX*, and *BELL* genes identified in different plant lineages. Numbers for *Arabidopsis thaliana*, *Populus trichocarpa* (Poplar), *Zea mays* (maize), *Oryza sativa* (rice), *Selaginella moellendorffii* and *Physcomitrella patens* were sourced from Mukherjee et al., 2009. Additional information was sourced from Lee et al., 2008 (*Chlamydomonas reinhardtii*); Liam Dolan personal communication (*Marchantia polymorpha*); Harrison et al., 2005 (*Selaginella kraussiana*) and Sano et al., 2005 (*Ceratopteris richardii*). A question mark indicated that the number of genes is currently not known.

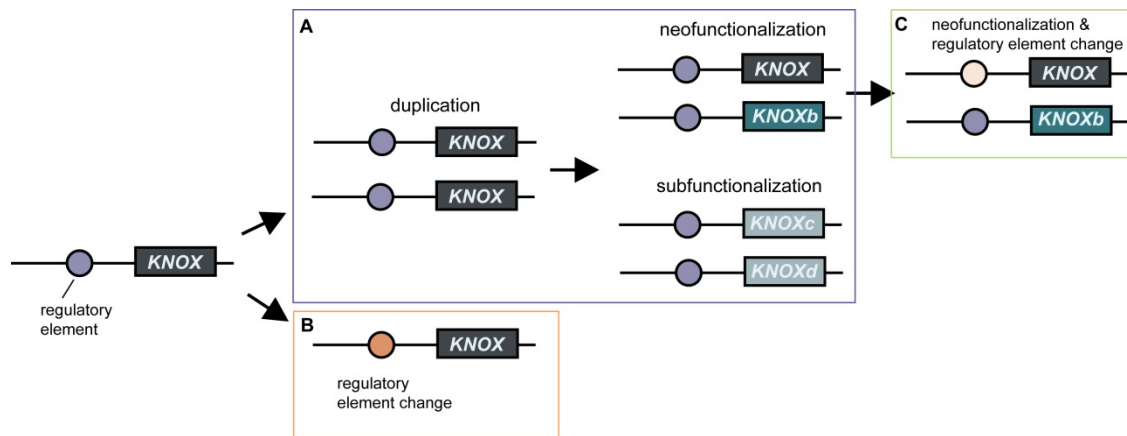


Figure 1.7

Importance of gene duplications in the evolution of *KNOX* gene function (adapted from Carroll, 2005). The main drivers of morphological evolution are coding sequence changes and regulatory element changes. A) Coding sequence changes emerge mainly through gene duplication. The subsequent two gene paralogues can have two main fates: either 1) one may gain an entirely new function while the other retains its original function (neofunctionalization) or 2) duplicate genes may divide the function of the progenitor gene between them (subfunctionalization) (Wagner, 2001). B) Regulatory element changes are also important in the evolution of morphology. C) Both structural and regulatory changes can drive evolution of new morphologies.

1.4.3 Two distinct expression patterns of class 1 *KNOX* genes

Most of our current understanding about the role of *KNOX* genes in plant development comes from studies of class 1 *KNOX* genes in angiosperms. Class 1 *KNOX* genes are expressed predominantly in meristematic tissues. This contrasts with class 2 *KNOX* genes that are expressed in tissues throughout the plant body (Hake et al., 2004; Hay and Tsiantis, 2009, 2010) and class 3 *KNOX* genes that are expressed in organ primordia and at the boundary of mature organs but not in the SAM (Magnani and Hake, 2008).

Class 1 *KNOX* gene expression during vegetative growth in angiosperms displays two distinct patterns (Figure 1.8). In the first, transcripts accumulate throughout the SAM but not in P0 cells where the leaf primordia initiate. This pattern is exhibited by *Arabidopsis STM* (Long et al., 1996), maize *KN1* (Vollbrecht et al., 1991) and rice *ORYZA SATIVA HOMEBOX 1 (OSH1)* (Kano-Murakami et al., 1993). In the second, transcripts accumulate in subdomains of the SAM, a pattern exhibited by *Arabidopsis BP* (Lincoln et al., 1994), maize *ROUGH SEATH1 (RS1)* (Kerstetter et al., 1994) and rice *ORYZA SATIVA HOMEBOX 15 (OSH15)* (Sato et al., 1998).

The two distinct *KNOX* gene expression patterns observed correlate with two different categories of function during vegetative growth. In the first category, genes with *STM*-like expression patterns control the establishment and maintenance of the SAM (Figure 1.9A and B). In the second category genes with *BP*-like expression patterns control internode growth. This second category, is exemplified by *bp* mutants that have reduced height, shortened internodes and downwards pointing siliques and flowers (Figure 1.9C) (Byrne et al., 2002; Douglas et al., 2002; Venglat et al., 2002; Smith and Hake, 2003). Similar to *bp*, *osh15* mutants also have reduced height, however, only the lower internodes are shortened (Figure 1.9D) (Sato et al., 1999).

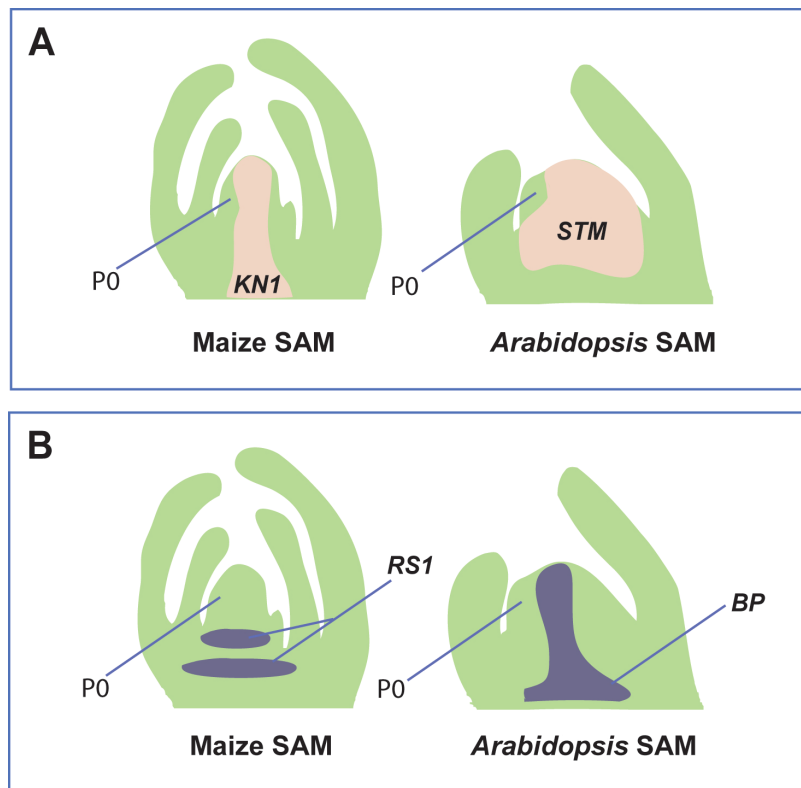


Figure 1.8

Expression patterns of class 1 *KNOX* genes in angiosperms. Two principal class 1 *KNOX* gene expression patterns have been observed in angiosperms. In one pattern, class 1 *KNOX* genes are expressed throughout the SAM except for P0 primordia (A), and in the other, class 1 *KNOX* genes are expressed within subdomains of the SAM (B) (modified from Hake, 2004 and Ragni et al., 2007).



Figure 1.9

Loss-of-function class 1 *KNOX* mutant phenotypes. *Arabidopsis stm* seedlings have fused cotyledons because they fail to form a SAM (arrowhead) (A). Similarly, maize *kn1* mutants (indicated by arrows) produce no leaves or just a single leaf (B). *Arabidopsis bp* mutants have shorter internodes than wild-type and downwards pointing siliques and flowers (C) and the *osh15* mutation in rice results in dwarf plants (D). (A: Hay and Tsiantis, 2010, B,C & D: Hake et al., 2004).

1.5 Evolution of *KNOX* gene function

The key role of *KNOX* genes in angiosperm shoot development raises the question of how this role evolved.

1.5.1 Function of *KNOX* genes in green algae

Insight into the role of *KNOX* genes in lineages that diverged before plants colonized the land was provided by a study in the chlorophyte alga *Chlamydomonas reinhardtii* (Lee et al., 2008). *C. reinhardtii* has two types of gametes, plus and minus, that express the *KNOX* protein GAMETE SPECIFIC MINUS 1 (GSM1) and the BELL protein GAMETE SPECIFIC PLUS 1 (GSP1), respectively. GSM1 and GSP1 proteins accumulate in the gametes, physically interact with each other upon gametic fusion and translocate from the cytosol to the nucleus to initiate zygote development.

1.5.2 Function of *KNOX* genes in mosses

After the colonization of land by plants, *KNOX* genes diversified both in number and function. Studies on the role of the three class 1 and the two class 2 *KNOX* genes in the moss *Physcomitrella patens* inform our understanding about *KNOX* gene function in early diverging land plants.

As in angiosperms, *KNOX* gene expression in *P. patens* is sporophyte-specific (Champagne and Ashton, 2001; Sakakibara et al., 2008). However, unlike angiosperms, the moss sporophyte lacks an indeterminate meristem. The shoot initially grows from an apical cell that results from the first zygotic cell division, but apical cell activity ceases after approximately 12 cell divisions. A new multicellular meristem, called the seta meristem, is then formed in the middle of the sporophyte (Sakakibara et al., 2008). Meristematic activity of the moss sporophyte terminates with formation of the sporangium.

When all three class 1 *KNOX* genes are mutated in *P. patens*, cell division planes in the apical cell of the embryo are disrupted. Defects in seta meristem cell divisions and growth are also observed (Figure 1.10) (Sakakibara et al., 2008). This suggests that class 1 *KNOX* genes in *P. patens* control meristematic activity in the sporophyte. In contrast, when class 2 *KNOX* genes are mutated, sporophytes make the transition to gametophyte development in the absence of meiosis. This suggests that class 2 *KNOX* genes suppress the gametophyte developmental program in the sporophyte (Sakakibara et al., 2013).

Based on our understanding of *KNOX* gene function in *P. patens*, it can be hypothesized that after the colonization of land, class 1 *KNOX* genes acquired functions to control meristematic cell divisions in the sporophyte and class 2 *KNOX* genes evolved to maintain the diploid state through suppression of the gametophytic development program. This diversification in *KNOX* function could have been facilitated by the observed parallel increase in *BELL* gene number (Table 1.1). A higher number of *KNOX* and *BELL* proteins would pave the way for new protein-protein interactions and for new heterodimers to control new developmental pathways.

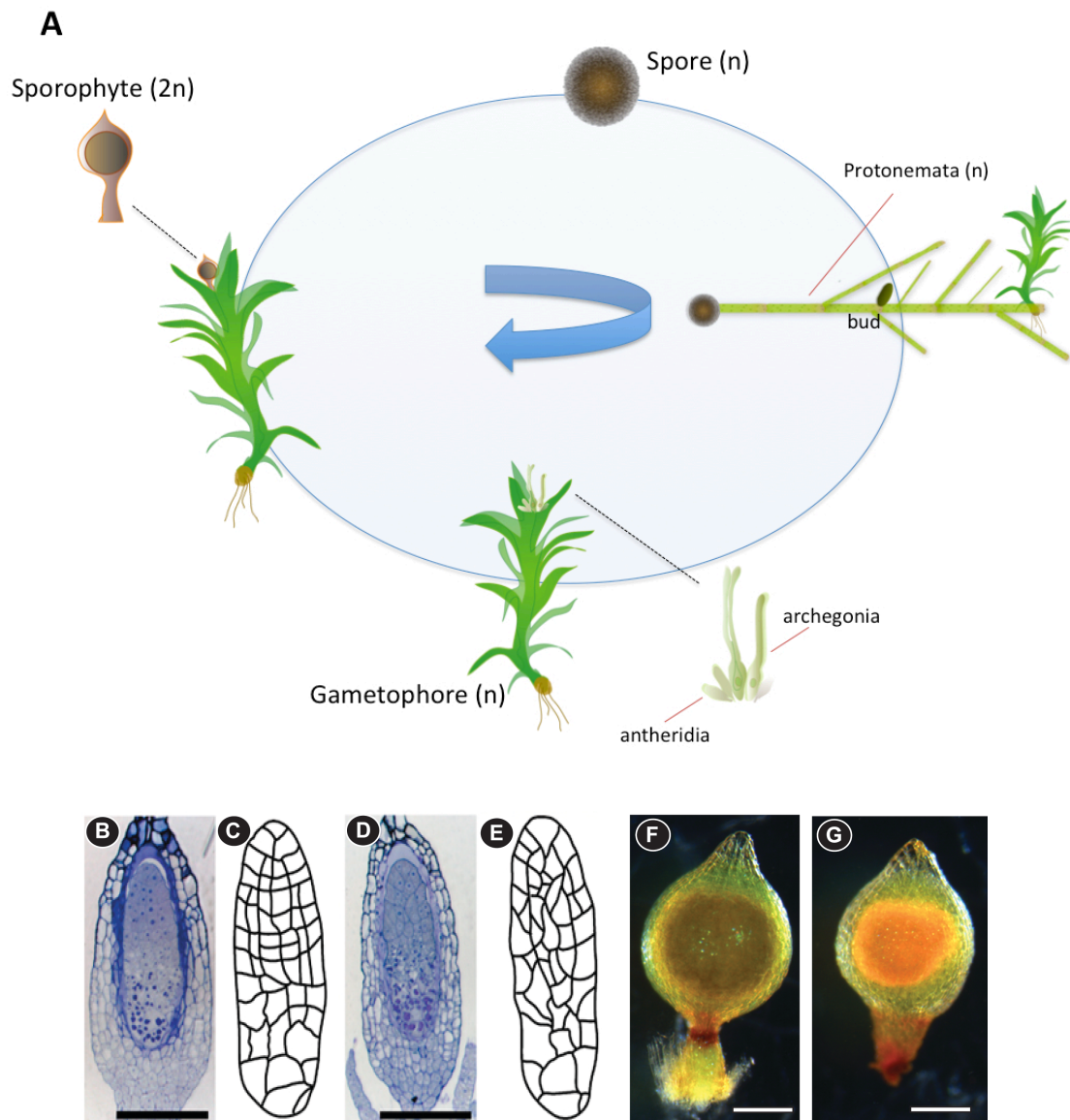


Figure 1.10

Life cycle of the moss *P. patens* and the phenotype of class 1 *KNOX* mutants. A) Moss *P. patens* life cycle. After spore germination, filamentous structures, called protonemata, develop. Buds then develop on protonemata and differentiate into leafy shoots (gametophore). Male (antheridia) and female (archegonia) reproductive organs develop at the apex of the shoot. After fertilization, the egg develops into a small unbranched diploid (2n) sporophyte that produces a sporangium at its apex. Meiosis occurs within the sporangium, leading to spore formation. B-G) Developing embryo sections showing the difference between planes of mitotic divisions in wild type (B&C) and class 1 *KNOX* mutants (D&E) (Sakakibara et al., 2008). F) Mature wild-type sporophyte and G) sporophyte of class 1 *KNOX* mutant. B-E scale bar: 10 μ M. F&G scale bar: 250 μ M.

1.5.3 *KNOX* genes in lycophytes

The function of *KNOX* genes further diversified during the evolution of vascular plants. In the lycophyte *Selaginella kraussiana* two class 1 *KNOX* genes have been identified that have similar expression patterns to the two distinct patterns seen in angiosperms: *SkKNOX1* is expressed in cells subtending the apical cells, whereas *SkKNOX2* is expressed in internodal regions (Figure 1.11) (Harrison et al., 2005). As in angiosperms, *S. kraussiana KNOX* genes are not expressed in leaf primordia whereas a *S. kraussiana ARP* homologue (*SkARP1*) is. Notably, *SkARP1* can complement *Arabidopsis as1* mutants. These observations, in combination with the proposed independent origin of leaves in lycophytes and seed plants, suggest that a similar mechanism for regulating the distinction between indeterminate shoot and determinate leaf growth evolved independently in both lycophytes and angiosperms.

1.5.4 *KNOX* genes in monilophytes

It is also possible that functional conservation exists between class 1 *KNOX* genes of monilophytes and angiosperms. In the fern *Ceratopteris richardii* two class 1 *KNOX* genes have been identified, *CrKNOX1* and *CrKNOX2* (Sano et al., 2005). The expression patterns of both *CrKNOX* genes are similar, although *CrKNOX2* transcripts accumulate at lower levels than *CrKNOX1*. Both genes are expressed in the shoot apical cell of the sporophyte, but weak expression is also detected in leaf primordia (Sano et al., 2005). Expression of *KNOX* genes in fern leaf primordia has also been reported in the ferns *Anogramma chaeophylla* (Bharathan et al., 2002) and *Osmunda regalis* (Harrison et al., 2005). It has been postulated that this expression pattern reflects the delayed determinacy of fern leaves, given that no *KNOX* expression was detected in mature leaves of *O. regalis* (Harrison et al., 2005). *CrKNOX1* and *CrKNOX2* overexpression in *Arabidopsis* has similar effects to that of overexpression of the native *Arabidopsis KNOX* genes (Sano et al., 2005). Together these

observations suggest that the role of class 1 *KNOX* genes in meristem and leaf development was most likely established in the common ancestor of monilophytes and angiosperms.

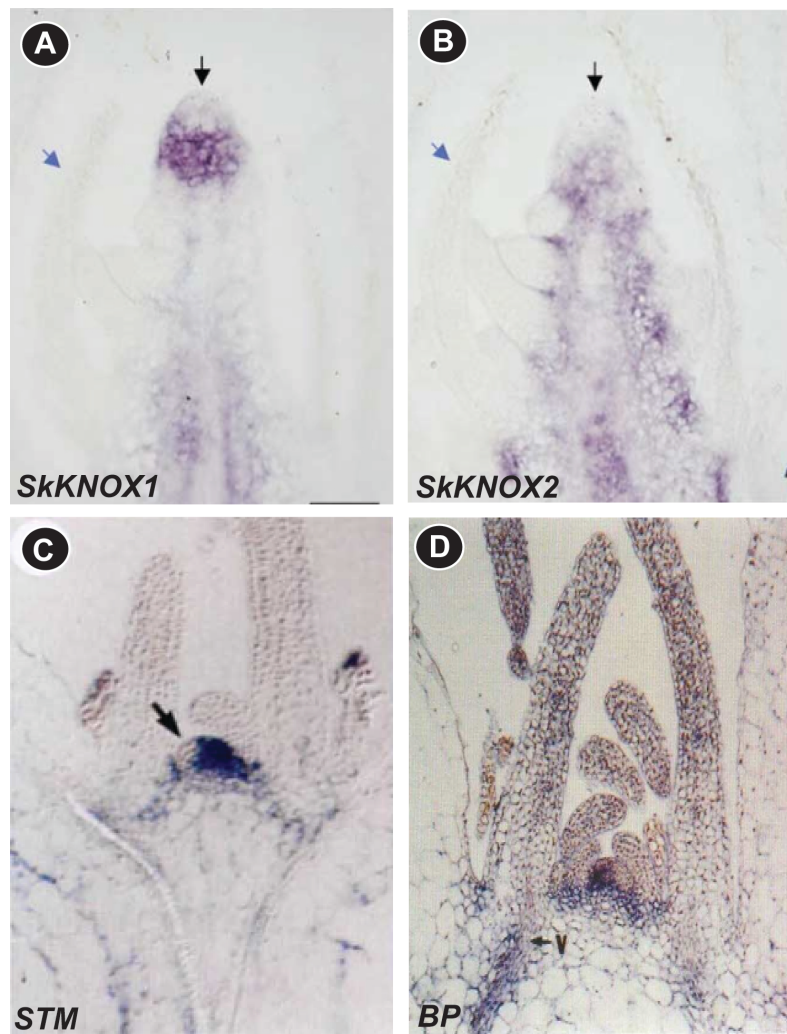


Figure 1.11

Comparison of class 1 *KNOX* gene expression patterns in *S. kraussiana* and *Arabidopsis*. *SkKNOX1* expression pattern (A) is similar to *STM* (C) while *SkKNOX2* (B) is similar to *BP* (D) (Harrison et al., 2005; Long et al., 1996; Lincoln et al., 1994). Black arrows and blue arrows indicate the location of apical cells and leaf primordia respectively in (A) and (B). The SAM in (C) is indicated by a black arrow. Vascular tissue in (D) is indicated by a "V".

1.5.5 *KNOX* genes in gymnosperms

Consistent with the suggestion that the role of class 1 *KNOX* genes in meristematic activity was established prior to the divergence of monilophytes and angiosperms, *KNOX* gene function in gymnosperms is similar to that seen in angiosperms. Four class 1 *KNOX* genes have been identified in Norway spruce (*Picea abies*), *HBK1*, *HBK2*, *HBK3* and *HBK4* (Sundas-Larsson et al., 1998; Hjortswang et al., 2002; Guillet-Claude et al., 2004). *HBK1* is expressed in the stem and the apical meristem but not in organ primordia. When *HBK1* is expressed ectopically in *Arabidopsis*, it causes phenotypes similar to those seen with ectopic expression of *Arabidopsis* native class 1 *KNOX* genes (Sundas-Larsson et al., 1998). The role of *HBK2*, *HBK3* and *HBK4* genes is less clear. Interestingly, *HBK2*, *HBK3* and *HBK4* genes were also found to be expressed in roots, something not seen for class 1 *KNOX* genes in angiosperm species (Hjortswang et al. 2002). The conservation of class 1 *KNOX* gene function in shoots of gymnosperms and angiosperms is further supported by findings in the gymnosperm *Welwitschia mirabilis*. *W. mirabilis* produces just two leaves that grow from a basal meristem for up to 1500 years. Despite this atypical growth habit, *KNOX* expression was detected in the basal meristem but not in leaf primordia (Pham and Sinha 2003).

1.6 Aims of the project

The role of *KNOX* genes in apical growth of all extant land plants thus far examined suggests that modifications to *KNOX* gene function played a key role in land plant evolution. However, there are important gaps in our knowledge about the nature of those modifications. What is the function of *KNOX* genes in bryophytes other than mosses? How was *KNOX* gene function modified as shoots evolved from determinate structures in bryophytes to indeterminate structures in vascular plants? To what extent are *KNOX* genes in different groups equivalent? The aim of this work was to shed light on these questions by:

- investigating the extent of conservation of class 1 *KNOX* gene function between mosses, lycophytes, ferns and angiosperms
- investigating the role of *KNOX* genes in hornworts.

The work presented here is divided as follows:

A comparative analysis of class 1 *KNOX* gene function in lycophytes, ferns and angiosperms is described in Chapter 3. A comparative analysis of class 1 *KNOX* gene function in bryophytes and vascular plants is described in Chapter 4. Experimental attempts to study *KNOX* gene function in hornworts are presented in Chapter 5. Finally, in Chapter 6 the role of *KNOX* genes in sporophyte evolution is discussed in the context of the results of this study.

2 Material and Methods

2.1 General Materials

Chemicals were purchased from Melford (Suffolk, UK), Sigma (Gillingham, UK), Merck (Lutterworth, UK), Invitrogen (Paisley, UK) and Fisher (Loughborough, UK). Water was purified through MilliQ (Millipore, Darmstadt, Germany) multiple de-ionising and filtration cartridges (dH₂O). Solutions were sterilized by autoclaving (120°C for 20 min at 100 kPa) or by filter sterilization using a 0.2 µm syringe filter, Nalgene (Rochester, USA). Synthetic oligonucleotides were ordered from Eurogentec (Liege, Belgium). All radiochemicals were obtained from Perkin-Elmer Inc (Bucks, UK).

2.2 General equipment

2.2.1 Centrifuges and rotors

Centrifugations were carried out using the following centrifuges and rotors: Eppendorf 5415D, referred to as a microcentrifuge, Eppendorf 5424, Eppendorf 5430 with Eppendorf A-2-MTP swing bucket rotor. Sorvall Evolution RC (Thermo Fisher, USA) with Fiberlite F10-6x500y and Fiberlite F8-6x1000y rotors. Sorvall Econospin with four places swinging bucket rotor.

2.2.2 Microscopy

P. patens and hornwort images were captured using a QImaging Micropublishing 5.0 RTV camera on a Leica M165C microscope. Area measurements were carried out using the ImageJ software (<http://rsb.info.nih.gov/ij/>). A Leica DMRB light microscope was used to count protoplast number during moss transformations.

2.2.3 Spectrophotometry

An Eppendorf Biophotometer was used to measure the absorbance of bacterial cultures at 600 nm. The absorbance of a solution was referred to as its optical density (OD). Disposable plastic cuvettes (Sarstedt, Leicester, UK) were used for measurement of absorbance.

A ND-1000 Nanodrop spectrophotometer (Nanodrop Technologies, USA) was used to measure the absorbance of nucleic acids. DNA and RNA concentration was determined according to the manufacturer's instructions.

2.3 Plant material and growth conditions

2.3.1 *Arabidopsis*

The *Arabidopsis* ecotypes Columbia (*Col-0*) and Landsberg Erecta (*Ler*) were used. The *bp* and *stm* mutants described in this thesis refer to *bp-9* and *stm-1* alleles respectively. The *bp-9* mutant allele contains a *dSpm* insertion in (Smith and Hake, 2003). The *stm-1* mutant allele contains a point mutation within the *STM* coding sequence which causes a glutamine to stop codon change (CAA to TAA) (Long et al., 1996).

2.3.1.1 Growth of *Arabidopsis* plants on Murashige-Skoog medium

Arabidopsis seeds were sterilized in 70% ethanol for 10-15 min and then in 20-30%

(v/v) bleach for 5 min with gentle agitation. Seeds were subsequently washed 3-5 times in sterile water. Seeds were plated onto 1x MS (Murashige-Skoog) medium supplemented with Gamborg's vitamin plus 1% (w/v) sucrose, 0.05% (w/v) 2-N-Morpholino ethanesulfonic acid (MES). The pH was adjusted to 5.8 using 1M KOH and the media solidified using 0.8% (w/v) Type M agar (Sigma, UK). Plates were then sealed with micropore tape and the seeds vernalized for 48 h in the dark. Plates were subsequently transferred to a growth chamber (22–24°C, 16 h light and 8 h dark, 3500 lux light intensity).

2.3.1.2 Growth of *Arabidopsis* plants in soil

Seeds were sown in compost (Levington M2)/vermiculite mix (5:1 ratio) in a greenhouse at 20-22°C with 16 h light and 8 h dark and 3000 lux light intensity.

2.3.2 *P. patens*

The Gransden strain of *P. patens* was used for all experiments. This strain was originally cultured from a single spore of plant material collected in Gransden Wood in Cambridgeshire (Engel, 1968).

2.3.2.1 *P. patens* growth medium

P. patens was cultured on BCD medium. The composition of the medium was, 1 mM MgSO₄, 1.8 mM KH₂PO₄ (adjusted to pH 6.5 with 4 M KOH), 10 mM KNO₃, 45 µM FeSO₄, 0.001% Trace Element Solution (TES) – 10 µM H₃BO₃, 0.11 µM Al₂(SO₄)K(SO₄)₂, 0.22 µM CuSO₄, 0.24 µM KBr, 0.66 µM LiCl, 2 µM MnCl₂, 0.23 µM CoCl₂, 0.19 µM ZnSO₄, 0.17 µM KI and 0.12 µM SnCl₂, 1 mM CaCl₂, solidified with

0.8% (w/v) high gel strength agar (Sigma, UK) (Grimsley *et al.*, 1977).

2.3.2.2 *P. patens* culture conditions

To maintain protonemal cultures, 2 g of protonemal tissue was homogenized in 5-10 ml of sterile water using an IKA Ultra Turrax homogenizer and corresponding dispensing tools (Fisher, UK). The homogenized tissue was plated onto solid cellophane-overlaid (Type 325P, A. A. Pachaging Ltd., Preston, Lancashire) BCD medium supplemented with 5 mM diammonium tartrate (to promote protonemal growth). Moss was grown in an MLR-351 Sanyo Versatile Environmental Test Chamber (24°C, continuous light, 3500 lux light intensity).

2.3.2.3 *P. patens* gametophore induction

One week old protonemal tissue was transferred with a sterile scalpel onto Magenta pots containing 100 ml solidified BCD medium supplemented with 1 mM CaCl₂. The pots were then placed in a tissue culture room (24°C, 16 h light and 8 h dark, light intensity 2500 lux) for 1.5 months.

2.3.2.4 *P. patens* sporophyte induction

Magenta pots containing tall gametophores (see 2.3.2.3 above) were transferred into a Panasonic MLR-352 Versatile Environmental Test Chamber (16°C, 16 h light and 8 h dark, light intensity of 2000 lux). 5-6 ml of sterile water was added to each pot right after transferring to 16°C. Gametangia formation started within 2-3 weeks and sporophytes were visible within 5-6 weeks.

2.3.3 Hornwort strains

Anthoceros punctatus was obtained from the Royal Botanical Garden of Edinburgh (Royal Botanic Garden of Edinburgh, Edinburgh, UK, Grid: NT 24 75, v.-c. 83, Alt. c. 30 m, 20/04/2011). *Anthoceros agrestis* was obtained from Berwickshire (Berwickshire, near Fogo, Grid: NT 7700 4894v.-c. 81, Alt. c. 115 m, 30/10/2012). Axenic cultures of gametophytes were established from more than one spore of *A. punctatus* and from a single spore of *A. agrestis*.

2.3.3.1 Hornwort growth conditions

Both *A. punctatus* and *A. agrestis* thallus tissue were propagated on Lorbeer's medium (1 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g/L KH_2PO_4 2 g/L NH_4NO_3 1 g/L CaCl_2 (Proskauer, 1969)) supplemented with 1 ml of Hutner's trace elements (50 g/L EDTA disodium salt, 22 g/L $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 11.4 g/L H_3BO_3 , 5.06 g/L $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 1.61 g/L $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 1.57 g/L $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1.1 g/L $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4 \text{H}_2\text{O}$, 4.99 g/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (Hutner et al., 1950)). The medium was adjusted to pH 6.5 with 0.1M KOH and solidified using 6.5 g/L Gelzan CM (Sigma, UK). Plants were grown in a tissue culture room (23°C, 16 h of light and 8 h of dark, 3000 lux light intensity). In order to sub-culture thallus tissue, a small part of the thallus was cut and then placed on fresh medium and this was repeated on a monthly basis.

2.3.3.2 Hornwort spore sterilization

Isolated sporophytes were left to dry before removing the spore contents. Spores were placed in a 1.5 ml Eppendorf tube with 1 ml 5% (v/v) bleach (sodium hypochloride solution, ~10% (SIGMA, UK)) and gently agitated for 3 min. Spores were

briefly spun using a microcentrifuge and then washed 3-4 times with sterile water. The sterilized spores were then plated on Lorbeer's medium. Plates were placed in a tissue culture room (23-25°C, 16 h of light and 8 h of dark, 3000 lux light intensity). Germination occurred after approximately 7 days. Young gametophytes were transferred onto fresh Lorbeer's medium one month after germination.

2.3.3.3 *Hornwort sporophyte induction*

To induce sporophytes, plants were grown for one month in petri dishes on Lorbeer's medium. Afterwards, plants were transferred to Magenta pots on 1/10 Knop medium (2.5 g/L K_2HPO_4 , 2.5 g/L KH_2PO_4 , 2.5 g/L KCL, 2.5 g/L $MgSO_4 \cdot 7H_2O$, 10 g/L $Ca(NO_3)_2 \cdot 4H_2O$, 37 mg/L $FeSO_4 \cdot 7H_2O$) at pH 6.2 and containing 0.7% (w/v) agar. Pots were then placed in a Panasonic MLR-352 Versatile Environmental Test Chamber (16°C, 16 h of light and 8 h of dark, 1500 lux light intensity). After approximately one month, antheridia started to appear and after one more month sporophytes emerged.

2.4 Bacterial material, growth conditions and transformation

2.4.1 Bacteria strains

The *Escherichia coli* strain DH5 α was used for plasmid manipulation. A disarmed *Agrobacterium tumefaciens* strain GV3101 was used for *Arabidopsis* transformation.

2.4.2 Bacterial growth conditions

E. coli was grown at 37°C overnight in liquid or solidified Luria Bertani (LB) medium (1% (w/v) bacto-tryptone, 0.5% (w/v), bacto-yeast extract, 1% (w/v) NaCl) pH 7.5, containing 0.8% (w/v) agar where necessary.

A. tumefaciens was grown at 28°C for two days in liquid or solidified (with 0.8% (w/v) agar) YEP medium (1% (w/v) yeast extract, 1% (w/v) tryptone peptone, 0.5% (w/v) NaCl, pH 7.0).

Where appropriate, antibiotics were added to the media to the following final concentrations: 50-100 µg/ml ampicillin (Amp), 50 µg/ml kanamycin (Kan), 140 µg/ml spectinomycin (Spec), 12.5 µg/ml tetracycline (Tet), 100 µg/ml rifampicin (Rif) and 20 µg/ml gentamicin (Gen).

2.4.3 Bacterial methods

2.4.3.1 Preparation *E. coli* competent cells

500 ml LB medium were inoculated with 5 ml of an overnight culture of DH5α *E. coli* cells and incubated at 37°C with vigorous agitation at 180 rpm until the culture had reached an OD₆₀₀ of 0.4-0.5. The culture was incubated for 30 min on ice and spun at 4200xg for 15 min at 4°C. The supernatant was discarded, and the cell pellet was gently resuspended in 300 ml ice cold sterile water at 4°C and centrifuged at 4200xg for 15 min at 4°C. The cells were again resuspended in 120 ml ice cold sterile water and transferred to four 30 ml centrifuge tubes. The tubes were centrifuged at 1500xg for 10 min at 4°C. The cells were resuspended in 20 ml ice cold sterile 10% glycerol at 4°C, and centrifuged at 1500xg for 10 min at 4°C. The cells were resuspended in a final volume of 1 ml ice cold sterile 10% glycerol, dispensed into 40 µl aliquots and fast frozen in dry ice and then stored at -80°C.

2.4.3.2 Preparation of electrocompetent *A. tumefaciens* GV3101 cells

100 ml YEP medium were inoculated with 20 ml of an overnight culture of *A. tumefaciens* and incubated at 28°C with vigorous agitation at 180 rpm until the culture had reached an OD₆₀₀ of 0.5. The culture was incubated for 30 min on ice and spun at 4200xg for 10 min at 4°C. The pellet was resuspended in 1 ml ice-cold 20 mM CaCl₂ and 100 µl aliquots placed in eppendorf tubes. Tubes were subsequently snap-frozen in liquid nitrogen and stored at -80°C.

2.4.3.3 Bacterial transformations

2.4.3.3.1 Transformation of *E. coli* by heat-shock

10 ng of plasmid DNA was added to 50µl competent *E. coli* DH5α cells and incubated on ice for 20 min. The transformation mixture was then heat shocked for 45 sec at 42°C and then placed on ice for 2 min. 500 µl LB medium was then added to the mixture and incubated for 1 h at 37°C with shaking at 190 rpm. 250 µl transformation mixture was then spread onto LB agar plates containing the appropriate antibiotics and these were subsequently incubated overnight at 37°C.

2.4.3.3.2 Transformation of *A. tumefaciens* by electroporation

10 ng of plasmid DNA was added to 100 µl of *A. tumefaciens* electrocompetent cells in a 0.2 cm electroporation cuvette (BioRad, UK). The mixture was then electroporated using a BioRad Genepulser (2.5 kV/cm, 25 µF, 600 Ω), for 5-6 milliseconds. Afterwards, 1 ml of liquid LB medium was added and the cells were incubated at 28°C for 3 h with shaking at 150 rpm. Cells were then spread onto plates containing the appropriate antibiotics and incubated at 28°C for 2 days.

2.5 Plant transformation

2.5.1 Transformation of *Arabidopsis* using floral dipping

A 5 ml LB preculture was inoculated with a single *A. tumefaciens* colony. Gentamicin, rifampicin and the plasmid-specific selection antibiotic were added. The preculture was incubated at 28°C for 48 h at 200 rpm. A 500 ml LB culture was then inoculated with the 5 ml preculture and incubated at 28°C for 18-24 h at 180 rpm. The bacteria were sedimented by centrifugation at 2700xg for 20 min at 4°C. The pellet was resuspended in 500 ml of 0.5x MS media, 5% (w/v) sucrose and 0.05% (v/v) Silwet (LEHLE SEEDS, USA). The bacterial solution was added to a narrow 500 ml beaker, large enough to fully immerse the plant material to be transformed. *Arabidopsis* plants were inverted into the bacterial solution for approximately 1 min. Plants were then removed from the solution and laid down on a tray, covered with Saran wrap and placed at 4°C overnight. Plants were returned to normal growth conditions in the greenhouse and allowed to set seed.

2.5.2 Transformation of *P. patens*

2.5.2.1 Preparation of protoplasts

Two plates of 5 day old protonemal tissue were used to prepare protoplasts. The following procedure was carried out in a sterile laminar flow hood. Driselase enzyme (Sigma, UK) was added to 10 ml of 8% (w/v) mannitol to a final concentration of 10 mg/ml in a 15 ml centrifuge tube and incubated at room temperature for 20 min (the tube was inverted every 5 min to mix the content). The tube was then centrifuged at

1100xg for 3 min in a benchtop centrifuge at room temperature. The supernatant was filtered into a new sterile round-bottomed centrifuge tube. The protonemal tissue was added to the tube using a sterile scalpel. The tube was covered with aluminium foil and then was incubated at room temperature for 45 min with gentle shaking. The digested tissue was filtered through a 100 µm nylon filter (Fisher, Loughborough, UK) to remove undigested material from the cell suspension. The filtrate was collected in a sterile 14 ml tube and left for an additional 10 min at room temperature. The tube was centrifuged in a benchtop centrifuge at 140 x g for 3 min and the supernatant was discarded. Protoplasts were washed twice in 6 ml 8% (w/v) mannitol. Protoplasts were resuspended in 6 ml of mannitol and the protoplast yield was estimated using a haemocytometer.

2.5.2.2 PEG-based transformation

Protoplasts were spun down in a benchtop centrifuge at 120 x g for 3 min and resuspended in 1 – 2 ml of MMM solution (9.1% (w/v) mannitol, 15 mM MgCl₂, 0.1 % (w/v) MES pH 5.6) to obtain a protoplast density of 1.5x10⁶/ml. 300 µl of the protoplast suspension was added to a 14 ml tube containing 25 ng of digested DNA in a 30 µl volume of water. 300 µl of PEG solution (0.12 g of PEG6000 dissolved in 7.2% (w/v) mannitol, 0.1 M Ca(NO₃)₂, 0.01 M tris-HCL pH 8.0) was slowly added to the tubes in drops. The solution was mixed by swirling the tubes, and the cells were incubated at 45°C for 5 min and then allowed to recover at room temperature for another 5 min. Each mix was gradually diluted by the addition of 300 µl 8% (w/v) mannitol 5 times at 4 min intervals and then by the addition of 1 ml of 8% (w/v) mannitol 5 times at 4 min intervals. The tubes were then centrifuged in a benchtop centrifuge at 140 x g for 3 min and the protoplasts were resuspended in 5 ml liquid BCD, mannitol and glucose (1 mM MgSO₄, 1.8 mM KH₂PO₄ (adjusted to pH 6.5 with 4 M KOH), 10 mM KNO₃, 45 µM FeSO₄, 10 µM H₃BO₃, 0.11 µM Al₂(SO₄)K(SO₄)₂, 0.22

μM CuSO_4 , 0.24 μM KBr , 0.66 μM LiCl , 2 μM MnCl_2 , 0.23 μM CoCl_2 , 0.19 μM ZnSO_4 , 0.17 μM KI and 0.12 μM SnCl_2 , 1 mM CaCl_2 , 8% (w/v) mannitol and 0.5% (w/v) glucose). The tubes were incubated at 24°C overnight in the dark.

2.5.2.3 Regeneration of protoplasts

Protoplasts were spun down by centrifugation in a benchtop centrifuge at 140 x g for 3 min. Pellets were resuspended in 2-3 ml of 8% (w/v) mannitol. After mixing well, 1 ml of the suspension was plated out onto cellophane-overlaid protoplast regeneration medium (PRMB) composed of BCD minimal medium supplemented with 5 mM ammonium tartrate, 60 g/L mannitol, 10 mM CaCl_2 , 0.5% (v/v) glucose and 8 g/l agar. Plates were incubated at 24°C, with continuous light and light intensity of 3500 lux for 5 days.

2.5.2.4 Screening of *P. patens* transformants

After 5 days of growth, cellophane discs were transferred to BCD medium supplemented with 5 mM ammonium tartrate containing blasticidin (Melford, UK). Antibiotic resistance was conferred by the transformed construct and therefore only transformed colonies survived in the presence of antibiotic. The colonies were maintained with antibiotic selection for two weeks and were then transferred to BCD media supplemented with 5 mM ammonium tartrate but without antibiotic. Transformants were allowed to grow for one week. The cellophane discs with regenerating moss colonies were transferred again to BCD medium supplemented with 5 mM ammonium tartrate and the antibiotic of selection. Colonies were allowed to grow for two more weeks in order to obtain stable transformants.

2.6 Nucleic acid extraction methods

2.6.1 Plasmid DNA

Plasmid DNA was extracted from *E. coli* cells using a Qiagen Midiprep kit (Qiagen, Chatsworth, CA). 40 µg of plasmid DNA was digested with 250 units (50u/µl) *NotI* (*P. patens*) or *BamHI* (*Arabidopsis*) restriction enzyme (NEB, UK) in a total volume of 150 µl at 37°C overnight. The enzyme was deactivated with heating at 65°C for 20 min and then the digested plasmid DNA was concentrated by ethanol precipitation. 0.1 volumes of 3 M NaOAc pH 4.5 and 2.5 volumes of 100% ethanol were added and after mixing well the mix was incubated at -80°C for 1 h and then centrifuged in a microcentrifuge at 16100xg for 15 min at 4°C. The pellet was washed with 500 µl of 70% ethanol and air dried for 5–10 min. The pellet was resuspended in 30 µl of sterile water.

2.6.2 *P. patens* and *Arabidopsis* genomic DNA extraction

350 mg of tissue was harvested and frozen in liquid N₂. Tissue was ground to a fine powder using a chilled mortar and pestle and then added to 3 ml of CTAB extraction buffer (100 mM Tris-HCl pH 8, 1.4 M NaCl, 20 mM EDTA pH 8, 2% (w/v) hexadecyltrimethylammonium bromide – C₁₉H₄₂NBr (CTAB), 0.3% (v/v) β-mercaptoethanol). 30 µl of RNase A (100 mg/ml) was added, the solution was mixed well and incubated at 65°C for 30 min and then removed from heat. 3 ml of chloroform:isoamylalcohol (24:1) was added, mixed well by inversion and then the solution was centrifuged at 12000xg for 10 min at 4°C. The upper aqueous phase was transferred to a clean 30 ml centrifugation tube and 0.7 volumes of isopropanol was added and mixed well, and then the mix was centrifuged at 23400xg for 15 min at 4°C.

The pellet was transferred to a new 1.5 ml tube and washed with 70% ethanol. The sample was centrifuged at 20000xg for 2 min at room temperature and the pellet was air dried for 5 min and resuspended in 80 µl of water.

2.6.3 Hornwort genomic DNA extraction

A modified CTAB protocol (from Porebski et al. 1997) was used for hornwort genomic DNA extraction. 0.5 g of tissue was harvested and frozen in liquid nitrogen. Tissue was ground into a fine powder using a chilled mortar and pestle and then added to 10 ml of DNA extraction buffer (100mM Tris-HCl pH 8, 1.4M NaCl, 20mM EDTA pH 8, 2% (w/v) CTAB, 0.3% (v/v) β-mercaptoethanol and 100 mg of polyvinylpyrrolidone (PVP)/g of tissue) that had been prewarmed at 60°C, 100 µl of RNase A (100 mg/ml) was added and the solution was mixed well. The mix was incubated at 60°C for 30 min and then removed from heat and allowed to cool at room temperature for 4-6 min. 12 ml of chloroform:isoamylalcohol (24:1) was added, mixed well by inversion and then centrifuged at 12000xg for 10 min at room temperature. The upper aqueous phase was transferred to a new 30 ml centrifugation tube and 10 ml of chloroform:isoamyl alcohol (24:1) was added, mixed well by inversion and then centrifuged at 5860xg for 10 min at room temperature to remove any remaining PVP in the aqueous phase. The upper aqueous phase was transferred to a new 30 ml centrifugation tube and ½ volume of 5M NaCl was added. 2 volumes of cold (-20°C) 95% ethanol were also added and the contents of the tube were mixed well by inversion. The tube was spun at 20000xg for 6 min. The pellet was resuspended in 2 ml of TE buffer and the previous step was repeated. The pellet was washed with cold 70% ethanol. The pellet was dried and dissolved in 500 µl of TE buffer and then stored at 4°C.

2.6.4 *Arabidopsis*, *S. kraussiana* and *P. patens* RNA extraction

All solutions used during this protocol were prepared using Diethyl pyrocarbonate (DEPC) treated sterile water. To prepare DEPC-treated water, DEPC was added to dH₂O to a final concentration of 0.1% (w/v) and the solution was incubated at room temperature overnight and then autoclaved.

0.5 g of tissue was harvested and frozen in liquid nitrogen. Tissue was ground to a fine powder using a chilled mortar and pestle and then added to a 30 ml centrifuge tube containing 700 µl of Lysis buffer (100 mM Tris-HCl pH 8.6, 2 % (w/v) sarcosyl, 4 M guanidine thiocyanate, 25 mM EDTA, 25 mM EGTA, 100 mM β-mercaptoethanol and 20 mM diethylcarbamic acid) and 350 µl of phenol. 350 µl of chloroform:isoamylalcohol (24:1) was also added and then mixed well. The mix was transferred to a 2 ml tube and centrifuged in a microcentrifuge at 7650xg for 5 min at 4°C. 700 µl of the aqueous phase was transferred to a new 1.5 ml tube, and 700 µl of phenol:chloroform:isoamylalcohol (25:24:1) was added and mixed well. The tube was centrifuged at 7659xg for 10 min at 4°C and the aqueous phase was transferred to a new 1.5 ml tube. 3.5 µl of acetic acid and 0.6 volumes of isopropanol were added, mixed well and the tube was placed for 1 – 2 h at -80°C. The tube was then centrifuged at 20000xg for 12 min at 4°C, the pellet washed with 95% ethanol, air-dried and then resuspended in 400 µl DEPC-treated dH₂O. 100 µl of 8 M LiCl was added to the sample and RNA was precipitated overnight at 4°C. The samples were centrifuged at 20000xg for 12 min at 4°C and then the pellet was dissolved in 400 µl DEPC-treated dH₂O. 40 µl of NaOAc pH 4.5 and 1 ml 95% ethanol were added and the mixture was incubated at -80°C for 30 min before centrifugation at 16100xg for 10 min at 4°C. The pellet was washed with 95% ethanol, air-dried, resuspended in 40 µl DEPC-treated dH₂O and stored at -80°C.

2.6.5 Hornwort RNA extraction

All solutions used during this protocol were prepared using DEPC water (see section 2.6.4).

0.5 g of tissue was harvested and frozen in liquid N₂. Tissue was ground into a fine powder in a chilled mortar and pestle and then added to 10ml of DNA extraction buffer (100mM Tris-HCl pH 8, 1.4M NaCl, 20mM EDTA pH 8, 2% (w/v) CTAB, 0.3% (v/v) β-mercaptoethanol and 100 mg of PVP/g of tissue). The mix was incubated at 60°C for 30 min and the 12 ml of chloroform:isoamylalcohol (24:1) was added, mixed well by inversion and then centrifuged at 12000xg for 10 min at room temperature. The upper aqueous phase was transferred to a new 30 ml centrifugation tube and 10 ml of chloroform:isoamyl alcohol (24:1) was added, mixed well by inversion and then centrifuged at 7650xg for 10 min at room temperature to remove remaining any PVP in the aqueous phase. The upper aqueous phase was transferred to a new 30 ml centrifugation tube and ½ volume of 5M NaCl was added. 2 volumes of cold (-20°C) 95% ethanol were also added and the contents of the tube were mixed well by inversion. The tube was spun at 20000xg for 6 min. The pellet was resuspended in 2 ml of TE buffer and the previous step was repeated. The pellet was washed with cold (0-4°C) 70% ethanol, air-dried and resuspended in 400 µl DEPC-treated dH₂O. 100 µl of 8 M LiCl was added to the sample and RNA was precipitated overnight at 4°C. The samples were centrifuged at 20000xg for 12 min at 4°C and then the pellet was dissolved in 400 µl DEPC-treated water. 40 µl of NaOAc pH 4.5 and 1 ml 95% ethanol were added and the mixture was incubated at -80°C for 30 min and centrifuged at 16100xg for 10 min at 4°C. The pellet was washed with 95% ethanol, air-dried, resuspended in 40 µl DEPC-treated water and stored at -80°C.

2.7 Nucleic acid manipulation

Vectors

2.7.1 Molecular Cloning

Plasmid DNA isolation was carried out using either a QIAGEN plasmid Miniprep kit or a QIAGEN plasmid Maxiprep kit according to the manufacturer's instructions (Qiagen). Restriction digests of plasmid DNA were performed according to the manufacturer's instructions (NEB, UK and PROMEGA, UK). Restriction fragments were isolated from agarose gels using GenElute Gel Extraction Kit (Sigma, UK). Ligation of DNA fragments was performed using T4 DNA ligase according to the manufacturer's instructions (Promega, UK). TaKaRa DNA ligation kit LONG (TaKaRa, Japan) was used when the size of the insert was bigger than 3.5 kb. The molar ratio of insert to vector was between 3:1 and 5:1. Ligations were incubated at 16°C overnight or at 22°C for 3 h. 10 µl of the ligation reaction was used to transform competent *E. coli* cells. Transformed colonies were initially screened by colony PCR and then verified by restriction digestion.

2.7.2 Ligation of PCR products

PCR products were ligated into the pJET1.2 vector. Reactions were performed according to the manufacture's instructions (Thermo Scientific, UK).

2.7.3 Site directed mutagenesis

Site directed mutagenesis was performed using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit following manufacturer's instructions (Agilent Technologies, USA).

2.8 Nucleic acid analysis

2.8.1 Polymerase Chain Reaction (PCR)

To amplify DNA fragments from template DNA a series of Polymerase Chain Reaction variations were used.

2.8.1.1 *Standard PCR*

100-150 ng of genomic DNA or 5-10 ng of plasmid DNA was used as a template in a PCR reactions using either GoTaq Polymerase enzyme mix (Promega, UK) or Phusion proof reading polymerase enzyme mix (Fermentas, UK). The following cycling conditions were used:

1 cycle 94°C for 5 min

30-35 cycles 94°C for 30 sec, 50-60°C for 30 sec, 72°C 30 sec - 5 min

1 cycle 72°C for 10 min

The annealing temperature was varied depending on the desired stringency of primer binding. The extension time was proportional to the size of the fragment being amplified using the relationship of 1 min of extension time per 1 kb to be amplified for the GoTaq polymerase and 15 sec of extension time per 1 kb to be amplified for the Phusion polymerase.

2.8.1.2 Colony PCR

Colony PCR was used to screen for plasmid inserts directly from *E. coli* colonies. GoTaq Polymerase enzyme mix was prepared on ice and then pipetted into 0.2 ml PCR tubes. A pipette tip was then used to pick a single colony of bacteria, which was first streaked on an LB agar plate and then the remaining bacteria on the tip were put inside the 0.2 ml PCR tube. After pipetting up and down a few times to resuspend the bacteria in the PCR mix, the PCR was carried out with the following conditions:

1 cycle	94°C for 5 min
30 cycles	94°C for 30 sec, 50-60°C for 30 sec, 72°C 30 sec - 2 min
1 cycle	72°C for 10 min

The annealing temperature was varied depending on the desired stringency of primer binding. The extension time was proportional to the size of the fragment being amplified.

2.8.1.3 Nested PCR

Nested PCR was used to amplify DNA products that were difficult to amplify by other methods. For nested PCR, two pairs of PCR primers (external and internal) were used. The first pair amplified a primary DNA fragment that was then used as a template for the second pair of primers to amplify the desired DNA sequence. Nested PCR was carried out in two steps. First a standard PCR was performed as described above. External oligonucleotides were used as primers for the first reaction. Then a second PCR round was performed using the internal pair of oligonucleotides as primers and 1 µl of the first round PCR as template. PCR was carried out with the following conditions:

1 cycle	94°C for 5 min
30 cycles	94°C for 30 sec, 50-60°C for 30 sec, 72°C 30 sec - 2 min

1 cycle 72°C for 10 min

2.8.1.4 Touchdown PCR

Touchdown PCR was used to increase specificity of PCR reactions. The PCR reaction was set up using GoTaq Polymerase enzyme mix (Promega, UK). All primers were used at a final concentration of 0.3 µM. The annealing temperature was up to 10 degrees higher than the calculated annealing temperature of the primers. The annealing temperature was then gradually decreased to the calculated annealing temperature of the primers and the following steps of amplification were carried out using this temperature. The following cycling conditions were used:

1 cycle 94°C for 5 min

10 cycles 94°C for 30 sec, 65-55°C (-1 per cycle) for 30 sec, 72°C 30 sec - 5 min

30-35 cycles 94°C for 30 sec, 50-60°C for 30 sec, 72°C 30 sec - 5 min

1 cycle 72°C for 10 min

The annealing temperature was varied depending on the desired stringency of primer binding. The extension time was proportional to the size of the fragment being amplified.

2.8.1.5 Reverse-transcription PCR (RT-PCR)

RNA was DNase treated according to the manufacturer's instructions (Ambion, Life Technologies, UK) and then converted to cDNA using Superscript III Reverse Transcriptase according to the manufacturer's instructions (Invitrogen, Life Technologies, UK).

PCR was carried out using 1 µl of the RT reaction product. Reaction mixtures were as detailed for the standard PCR and the cycling conditions were as follows:

1 cycle 94°C for 5 min
30 cycles 94°C for 30 sec, 50-60°C for 30 sec, 72°C 30 sec - 2 min
1 cycle 72°C for 10 min

The number of cycles in the second stage was varied between 30 and 45, depending on the abundance of cDNA in the reaction mix.

2.8.2 DNA sequencing

50-100 ng of PCR products and 250-350 ng of plasmid DNA were used as template for sequencing reactions. The following reaction was set up: 1µl Big Dye mix, 1.5µl 5X BigDye buffer, 0.5µl of 10 µM primer and dH₂O up to 10µl. The PCR programme for sequencing was as follows:

1 cycle 96 °C for 1min
40 cycles 96 °C for 10 sec, 50 °C for 5 sec, 60 °C for 4 min

Products were then precipitated by adding 0.1 volume of 100% ethanol and 2.5 volumes of 3 M NaOAc pH 5.2. The sample was mixed well and incubated at -20°C for 10-15 min. The sample was then centrifuged at 2200xg for 40 min at room temperature. The pellet was washed using 70% ethanol, centrifuged at 2200xg for 15 min at room temperature and then air dried in the dark. Products were run and resolved using an ABI 310 Sequencer (Department of Zoology, University of Oxford).

2.8.3 Southern blotting

2.8.3.1 Genomic DNA restriction digest

DNA was digested in 80 µl volume containing 6 µg of genomic DNA, 3 µl 10 x enzyme

buffer, 3 μ l 10 mM spermidine, 3 μ l 1 mg/ml BSA, 60 units of enzyme and water. Digests were incubated overnight at 37°C.

2.8.3.2 Gel electrophoresis

Digested genomic DNA was electrophoresed through a 0.8% (w/v) agarose and DNA was visualized by a FirstLight UV Illuminator and an image was taken using a Biospectrum 2D Imaging System (UVP, Cambridge, UK).

2.8.3.3 Capillary transfer of DNA onto nylon membrane

After electrophoresis, DNA gels were photographed alongside a ruler and the DNA was denatured by soaking the gel for 1 h in denaturing solution (1.5 M NaCl, 0.5 M NaOH) and then neutralized by soaking the gel for 1 h in neutralization buffer (3 M NaCl, 0.5 M Tris pH 7.5). The gel was then incorporated into the blotting apparatus. The blotting apparatus was set up within a glass dish containing 20 x SSC (3 M NaCl, 0.3 M tri-sodium citrate, pH 7.6) and a glass plate placed across the dish. The glass plate was covered with two sheets of Whatman 3mm paper that had their ends immersed into the 20x SSC and therefore acted as a wick. One piece of Whatman paper, cut to the size of the gel, was presoaked in 20x SSC and placed on the covered glass plate. The gel was placed upside down on this piece of Whatman paper. A piece of nylon membrane (GE Healthcare, Little Chalfont, UK), cut to the size of the gel, was presoaked in 2x SSC (0.3 M NaCl, 0.03 M tri-sodium citrate, pH 7.6) and placed over the gel. Two pieces of Whatman paper, presoaked in 2x SSC, were placed on top of the membrane. Dry towels were stacked on top and a weight was placed on top of the stack. After overnight transfer, the membrane was washed using 2x SSC at room temperature for 10 sec and allowed to air-dry. The DNA was then UV-

crosslinked to the membrane using a UV Stratalinker 1800 (Stratagene, Cambridge, UK).

2.8.3.4 DNA Radiolabelling

DNA probes for Southern blot analysis were radiolabelled using the Random Primer DNA Labeling System (Invitrogen, Life Technologies, Carlsbad, CA). 30 ng of DNA was made up to 21 μ l with dH₂O, heated at 100°C for 5 min and rapidly cooled on ice. The reaction was set up in a 1.5 ml tube with the denatured DNA, 15 μ l buffer, 2 μ l dATPs, 2 μ l dGTPs, 2 μ l dTTPs, 1 μ l Klenow fragment and 5 μ l of α^{32} P-dCTPs (10 μ Ci/ μ l). The tube was incubated for 1 h at 25°C. The probe was added to a 50 ml falcon tube containing 35 ml of hybridization buffer (see 2.8.3.5 below) and was boiled for 10 min to denature the DNA prior to hybridization.

2.8.3.5 Hybridization

Membranes were treated with 3 x SSC pre-hybridization buffer (0.45 M NaCl, 45 mM Na₃C₆H₅O₇, 1% (w/v) SDS, 0.1% (w/v) Na₄P₂O₇, 5x Denhardt's [0.1% (w/v) Ficoll, 0.1% (w/v) (C₆H₉NO)_n, 0.1% (w/v) BSA], 200 μ g/ml sheared single-stranded herring sperm DNA) for 1-3 h at 65°C with shaking. Membranes were then hybridized with the probe in 3 x SSC hybridization buffer (0.45 M NaCl, 45 mM Na₃C₆H₅O₇, 1% (w/v) SDS, 0.1% (w/v) Na₄P₂O₇, 5x Denhardt's [0.1% (w/v) Ficoll, 0.1% (w/v) (C₆H₉NO)_n, 0.1% (w/v) BSA], 10% (w/v) dextran sulphate, 200 μ m/ml sheared single-stranded herring sperm DNA) at 65°C overnight with shaking. The membrane was washed at 65°C with 2 X SSC wash solution (0.3 M NaCl, 30 mM Na₃C₆H₅O₇, 1% (w/v) SDS, 0.1% (w/v) Na₄P₂O₇) for 5 min and then with 1X SSC wash solution (0.15 M NaCl, 15 mM Na₃C₆H₅O₇, 0.1% (w/v) SDS, 0.1% (w/v) Na₄P₂O₇) for 10 min. The membrane

was air dried and covered with Saran wrap before it was exposed to Carestream Kodak Biomax XAR film (Sigma, UK) in a cassette with intensifying screens overnight.

2.9 Bioinformatics

2.9.1 Sequence alignment

Amino acid sequences were aligned using MergeAlign (Collingridge and Kelly, 2012) and alignments were trimmed to remove poorly aligned regions using trimAl (Capella-Gutierrez et al., 2009).

2.9.2 Phylogenetic analysis

ProtTest (Abascal et al., 2005) was used to select the best model of sequence evolution for the tree generation (JTT model of amino acid substitution for both *KNOX* and *ApKNOX* trees). The trees were generated using RaxML (Stamatakis, 2006) to perform a rapid bootstrap analysis with 100 trees and then search for the best-scoring maximum likelihood tree. Trees were visualized and edited using Dendroscope 3 (Huson and Scornavacca, 2012).

3 Comparative analysis of class 1 *KNOX* gene function in lycophytes, ferns and angiosperms

3.1 Introduction

KNOX gene expression patterns in the shoot apical meristem appear to be conserved between different families of vascular plants. For example, in the lycophyte *S. kraussiana*, the *SkKNOX1* gene expression pattern resembles the expression pattern of the *Arabidopsis STM* gene whereas *SkKNOX2* expression is similar to that of the *Arabidopsis BP* gene (Harrison et al., 2005). *CrKNOX1* and *CrKNOX2* genes in the fern *C. richardii* are also expressed in shoot apical cells (Sano et al., 2005). *KNOX* gene expression in the SAM of the lycophyte *S. kraussiana*, the fern *C. richardii* and of all angiosperms examined (Harrison et al., 2005; Sano et al., 2005) suggests that gene function may be conserved across these groups. To date there are no established lycophyte and fern experimental systems, and as such the necessary genetic tools to study *KNOX* gene function *in planta* are lacking. However, it is possible to inform our understanding of *KNOX* function in lycophytes and ferns by performing a series of cross-species complementation and overexpression experiments in *Arabidopsis*.

3.2 Aims

The aim of the work in this chapter was to determine the extent of functional conservation between class 1 *KNOX* genes from the lycophyte *S. kraussiana*, the fern *C. richardii* and the angiosperm *Arabidopsis*. The objectives were to investigate:

- the potential of class 1 *KNOX* genes from *S. kraussiana* and *C. richardii* to complement the phenotypes of class 1 *KNOX* mutants in *Arabidopsis* when expressed under the control of the endogenous *Arabidopsis BP* and *STM* class 1 *KNOX* gene promoters.

- the functional potential of *S. kraussiana* and *C. richardii* class 1 *KNOX* genes when constitutively expressed in wild type *Arabidopsis*.

3.3 Generation of constructs for transformation

All constructs for transformation were generated using the pART27 binary vector (Gleave, 1992). pART27 confers kanamycin resistance and contains both right and left transfer-DNA (T-DNA) border sequences necessary to integrate T-DNA into the plant genome following *Agrobacterium* mediated transformation (Figure 3.1).

3.3.1 Expression under the *Arabidopsis BP* promoter

To study the potential of *S. kraussiana* and *C. richardii* class 1 *KNOX* genes to complement *Arabidopsis bp-9* mutants, five constructs were designed to express *S. kraussiana* and *C. richardii* class 1 *KNOX* genes under the endogenous *BP* promoter (hereafter called *pBP::KNOX*) (Figure 3.1). The final constructs consisted of the *BP* 5' *cis*-regulatory sequence (from -5431bp to -1bp relative to the ATG) fused to the coding sequence of *SkKNOX1*, *SkKNOX2*, *CrKNOX1*, *CrKNOX2* or *BP*. Fused to the 3' end of the various *KNOX* gene coding sequences was the *octopine synthase* (*ocs3'*) terminator sequence to facilitate termination of transcription.

3.3.2 Expression under the *Arabidopsis STM* promoter

To study the potential of *S. kraussiana* and *C. richardii* class 1 *KNOX* genes to complement *Arabidopsis stm-1* mutants, five constructs were designed to express *S. kraussiana* and *C. richardii* class 1 *KNOX* genes under the endogenous *STM*

promoter (hereafter called *pSTM::KNOX*) (Figure 3.2). The final construct consisted of the *STM* 5' *cis*-regulatory sequence from (-5743bp to -1bp relative to the ATG) (Heisler et al., 2005) fused to the coding sequence of *SkKNOX1*, *SkKNOX2*, *CrKNOX1*, *CrKNOX2* or *STM*. Fused to the 3' end of the various *KNOX* gene coding sequences was the *ocs3'* terminator sequence to facilitate termination of transcription.

3.3.3 Constitutive expression

Five constructs were designed to constitutively express *S. kraussiana* and *C. richardii* class 1 *KNOX* genes (hereafter called *35S::KNOX*) (Figure 3.3) in both wild type (*Ler*) ecotype *Arabidopsis* plants and *stm-1 Arabidopsis* plants. *SkKNOX1*, *SkKNOX2*, *CrKNOX1*, *CrKNOX2*, *BP* and *STM* gene coding sequences were cloned downstream of the *CaMV* 35S promoter in the pART7 vector (Gleave, 1992), prior to insertion in pART27.

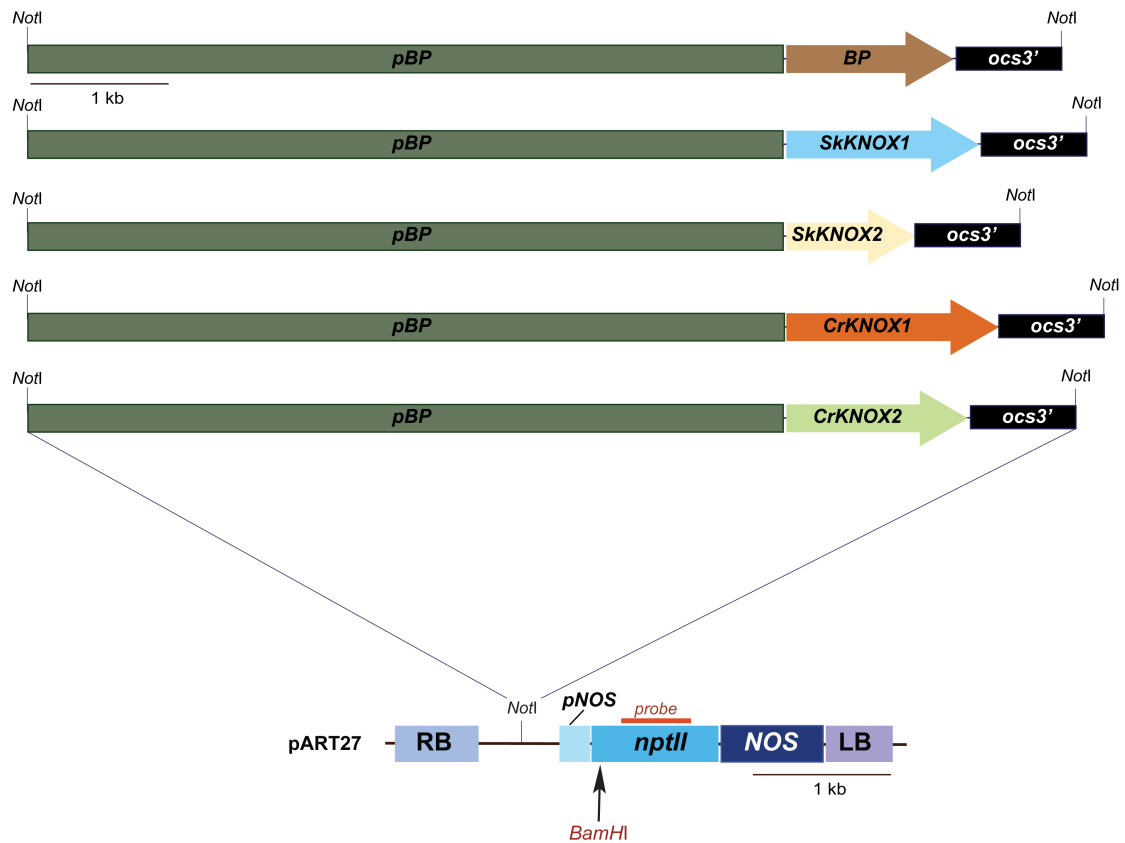


Figure 3.1

Schematic representation of *pBP::BP*, *pBP::SkKNOX1*, *pBP::SkKNOX2*, *pBP::CrKNOX1* and *pBP::CrKNOX2* constructs. Diagram represents *BP* gene promoter sequence (dark green boxes), various *KNOX* gene coding sequences (coloured arrows) and the *ocs3'* terminator (black boxes). All constructs were subcloned into the *NotI* restriction site of the pART27 binary vector. RB: right border, *pNOS*: *NOPALINE SYNTHASE* (*NOS*) promoter, *nptII*: *neomycin phosphotransferase II* gene (kanamycin resistance gene), *NOS*: *NOS* terminator and LB: left border. Kanamycin resistance cassette single *BamHI* restriction site is indicated with an arrow and the probe used for Southern blot hybridization is represented with a red box above the *nptII* gene.

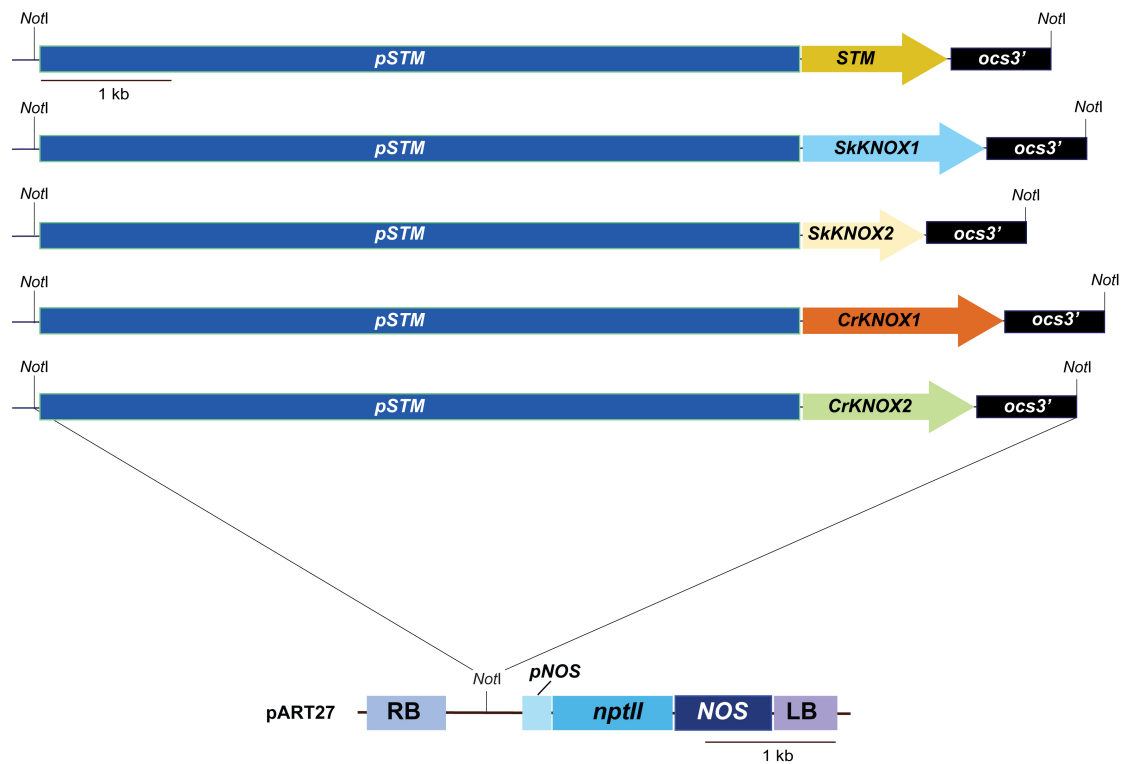


Figure 3.2

Schematic representation of *pSTM::STM*, *pSTM::SkKNOX1*, *pSTM::SkKNOX2*, *pSTM::CrKNOX1* and *pSTM::CrKNOX2* constructs. Diagram represents *STM* gene promoter sequence (blue boxes), various *KNOX* gene coding sequences (coloured arrows) and the *ocs3'* terminator (black boxes). All constructs were subcloned into the *NotI* restriction site of the pART27 binary vector. RB: right border, *pNOS*: *NOPALINE SYNTHASE* (*NOS*) promoter, *nptII*: *neomycin phosphotransferase II* gene (kanamycin resistance gene), *NOS*: *NOS* terminator and LB: left border

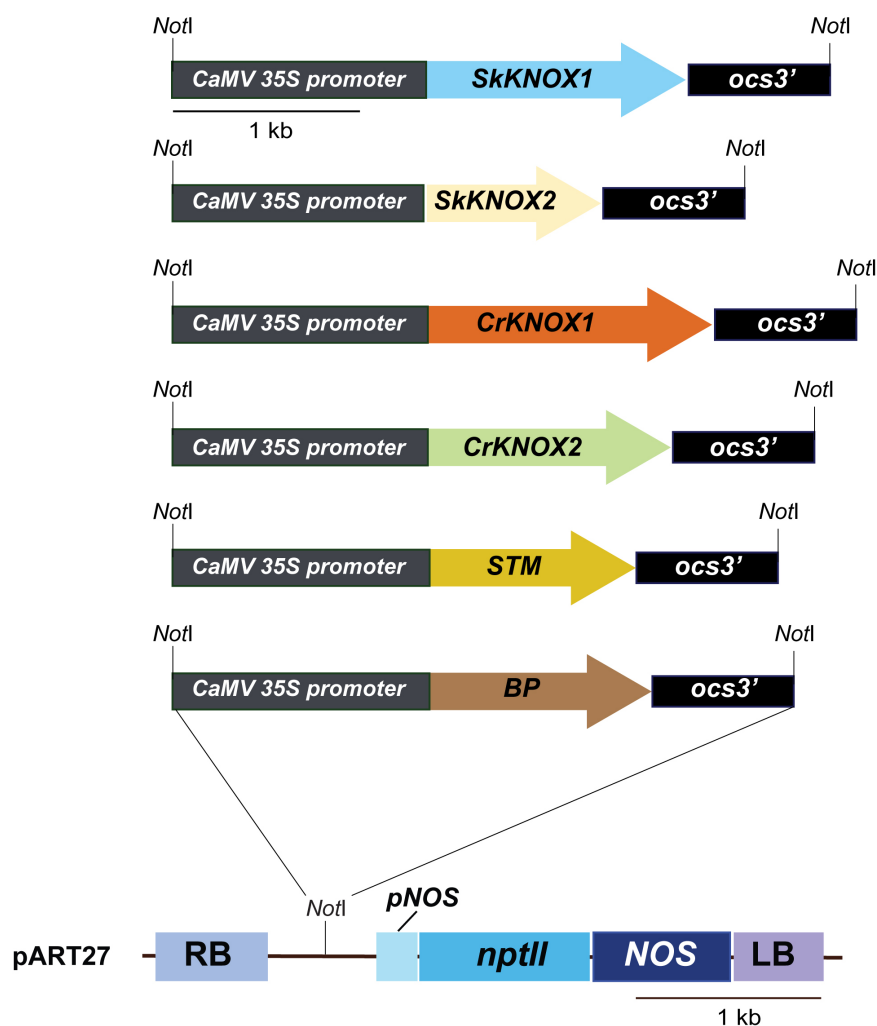


Figure 3.3

Schematic representation of *35S::SkKNOX1*, *35S::SkKNOX2*, *35S::CrKNOX1*, *35S::CrKNOX2*, *35S::STM* and *35S::BP* constructs. Diagram represents *CaMV 35S* gene promoter sequence (grey boxes), various *KNOX* gene coding sequences (coloured arrows) and the *ocs3*, terminator (black boxes). All constructs were subcloned into the *NotI* restriction site of pART27 binary vector. RB: right border, pNOS: NOPALINE SYNTHASE (NOS) promoter, *nptII*: neomycin phosphotransferase II gene (kanamycin resistance gene), NOS: NOS terminator, LB: left border

3.4 Generation of transformed lines

Arabidopsis bp-9 and *stm-1* mutants, plus wild type (*Ler*) *Arabidopsis* plants, were transformed by the floral dip method with the *pBP::KNOX*, *pSTM::KNOX* and *35S::KNOX* constructs respectively. *stm-1* mutant plants were also transformed with the *35S::KNOX* constructs. After setting seed, T₁ progeny were screened on MS kanamycin plates. Putative transformed seedlings were selected on the basis of survival, and were allowed to complete another generation. The T₂ seed collected from these T₁ plants were again screened on MS plates containing kanamycin, and segregation ratios of sensitive versus resistant plants were checked to identify plants carrying a single transgene copy. Unfortunately, several attempts to transform *stm-1* mutants with *pSTM::KNOX* and *35S::KNOX* constructs were unsuccessful. Generation of wild-type plants overexpressing the *35S::KNOX* constructs is still ongoing, having been deprioritized in favour of the *bp* complementation experiments. The number of independent lines selected for further analysis is summarized in Tables 3.1 and 3.2.

Construct	Numbers of plants	
	Initial T ₁ transformants	T ₂ lines segregating 3:1 ratios
<i>pBP::BP</i>	50	8
<i>pBP::SkKNOX1</i>	28	5
<i>pBP::SkKNOX2</i>	35	5
<i>pBP::CrKNOX1</i>	44	9
<i>pBP::CrKNOX2</i>	33	5

Table 3.1

Summary of *pBP::BP*, *pBP::SkKNOX* and *pBP::CrKNOX* constructs transformed into *bp-9* mutants. Numbers of T₁ individuals (initial transformants) obtained and numbers of T₂ lines segregating 3:1 ratios.

Construct	Line	Kan ^R	Kan ^S	Total	Segregation ratio in T2	T2	T3
<i>pBP::BP</i>	1	31	106	137	3.41:1		✓
	4	35	92	127	2.62:1		✓
	9	36	98	134	2.72:1		✓
	35	20	59	79	2.95:1	✓	
	37	12	36	48	3:1	✓	
	39	19	56	75	2.94:1	✓	
	42	22	54	76	2.45:1	✓	
	46	23	63	86	2.73:1	✓	
<i>pBP::SkKNOX1</i>	6	30	106	133	3.53:1		✓
	11	35	100	135	2.85:1		✓
	16	20	55	75	2.75:1	✓	
	21	18	47	65	2.61:1	✓	
	24	14	45	59	3.2:1	✓	
<i>pBP::SkKNOX2</i>	1	40	102	142	2.55:1		✓
	3	26	84	110	3.23:1		✓
	4	33	86	119	2.60:1		✓
	8	36	95	131	2.63:1		✓
	22	18	55	73	3.05:1	✓	
<i>pBP::CrKNOX1</i>	3	34	101	135	2.97:1		✓
	6	45	136	181	3.02:1		✓
	14	35	99	134	2.82:1		✓
	15	33	94	127	2.84:1		✓
	25	19	60	79	3.15:1	✓	
	26	16	41	57	2.56:1	✓	
	36	18	56	74	3.11:1	✓	
	39	13	47	60	3.61:1	✓	
	44	22	69	91	3.13:1	✓	
<i>pBP::CrKNOX2</i>	11	21	81	102	3.85:1		✓
	13	25	70	95	2.8		✓
	20	13	48	61	3.69:1	✓	
	21	19	60	79	3.16:1	✓	
	33	17	52	69	3.05:1	✓	

Table 3.2

Summary of segregation ratios for *pBP::BP*, *pBP::SkKNOX* and *pBP::CrKNOX* lines. Twelve days after germination, the number of kanamycin resistant (Kan^R) and kanamycin sensitive (Kan^S) plants were counted and the segregation ratio of resistant to sensitive was determined. Transgenic lines that segregated at a 3:1 ratio were allowed to self-fertilize to obtain homozygous transgenic lines in the T₃ generation. In some cases, due to time constraints, only T₂ segregating lines were obtained.

3.5 Genetic analysis of transgenic lines

3.5.1 Analysis of transgene copy number

To verify transgene copy number, Southern blot experiments were carried out on plants from the T₃ generation, using probes that specifically hybridize to the kanamycin resistance cassette (illustrated in Figure 3.1). Total genomic DNA was extracted from *pBP::SkKNOX1*, *pBP::SkKNOX2*, *pBP::CrKNOX1*, *pBP::CrKNOX2*, *pBP::BP*, wild type and *bp* mutant lines and 5 µg of each were digested with *Bam*HI. Figure 3.4 shows that all of the lines checked exhibited multiple hybridizing fragments indicating the presence of multiple inserts per line.

3.5.2 Transgene expression analysis

In order to confirm *BP* gene promoter activity in transformed plants, RT-PCR analyses were carried out using RNA extracted from two to five independent lines per transgene. Total RNA was extracted from approximately 10 seedlings and then cDNA was synthesized using 1 µg of total RNA. Primer pairs (Appendix 1) designed to amplify part of the N-terminal domain from each *KNOX* transgene were used to check transcript accumulation, and primers specific for the *Arabidopsis ACTIN 2 (ACT2)* (At3g18780) gene were used for loading and amplification controls. Figure 3.5 shows that transgene expression was detected in all of the lines. However, the levels of expression differed slightly between individual lines of the same transgene (notwithstanding the inherent limitations of quantification in assays such as this).

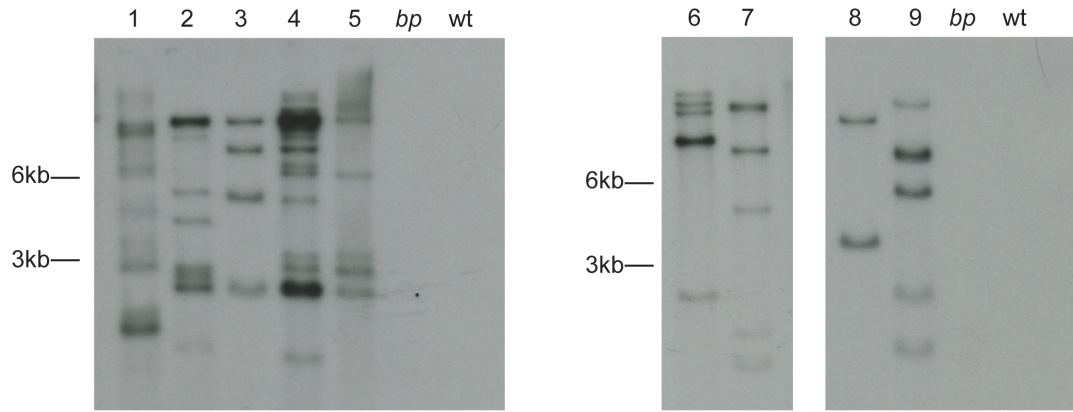


Figure 3.4

Southern blot analysis of *Bam*HI digested genomic DNA from *bp* mutant plants transformed with the *pBP::KNOX*. 1: *pBP::SkKNOX1* line 6, 2: *pBP::SkKNOX2* line 1, 3: *pBP::SkKNOX2* line 3, 4: *pBP::SkKNOX2* line 4, 5: *pBP::SkKNOX2* line 8, 6: *pBP::CrKNOX1* line 3, 7: *pBP::CrKNOX1* line 13, 8: *pBP::CrKNOX2* line 11, 9: *pBP::BP* line 4.

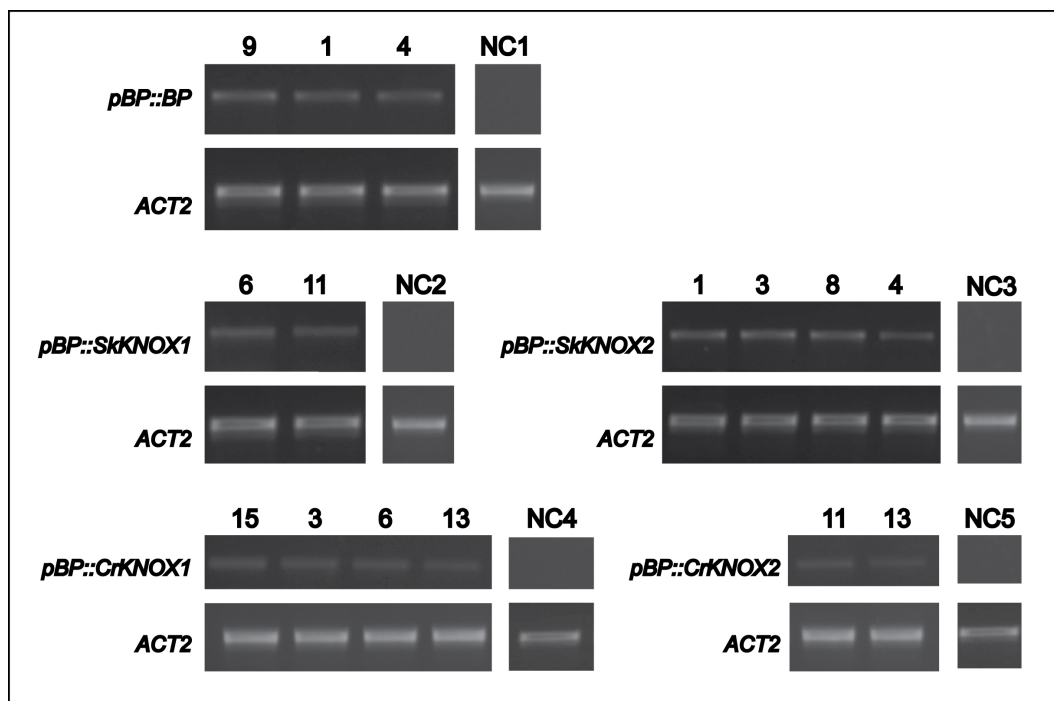


Figure 3.5

RT-PCR showing transcript accumulation in *bp* mutant plants transformed with *pBP::KNOX* constructs. *ACT2* was used as an internal amplification control. PCR assays were performed at 36 cycles, with both sets of primer pairs in each reaction. RT-PCR assays for *pBP::SkKNOX* and *pBP::BP* represent assays that were performed separately from those for *pBP::CrKNOX*. The identification number for each independent line is indicated. NC1-5: negative controls (run on a different gel); NC1: RT-PCR assays using primers for *BP* N-terminal domain and *bp* mutant genomic DNA, NC2: RT-PCR assays using primers for *SkKNOX1* N-terminal domain and *bp* mutant genomic DNA, NC3: RT-PCR assays using primers for *SkKNOX2* N-terminal domain and *bp* mutant genomic DNA, NC4: RT-PCR assays using primers for *CrKNOX1* N-terminal domain and *bp* mutant genomic DNA, NC5: RT-PCR assays using primers for *CrKNOX2* N-terminal domain and *bp* mutant genomic DNA.

3.6 Phenotypic characterization of plants

3.6.1 Cross-species complementation of *bp* mutants

Arabidopsis bp-9 mutants exhibit dwarfism and have downwards pointing siliques and flowers due to fewer cell divisions in the shoot internodes and pedicels (Venglat et al., 2002). Consequently, in order to determine whether *SkKNOX* and *CrKNOX* genes can complement the *bp* mutant phenotype, silique orientation (downwards pointing versus upwards pointing) and plant height were examined. Because insufficient T_3 homozygous lines were obtained, T_2 segregating lines were also used in this study.

3.6.2 Analysis of homozygous *pBP::KNOX* lines

Two to four transgenic lines per construct were grown for 8 weeks in a 16 h light : 8 h dark regime. Representatives of *pBP::KNOX*, wild type and *bp-9* mutant plants at 8 weeks post-germination are illustrated in Figure 3.6. Phenotypic complementation was observed in both of the *pBP::SkKNOX1* lines, three of the four *pBP::SkKNOX2* lines, only one of the four *pBP::CrKNOX1* lines and neither of the two *pBP::CrKNOX2* lines examined. Only one of three *pBP::BP* control transgenic lines examined showed partial complementation whereas the other two lines showed no potential to complement.

The angle of the siliques of the primary inflorescence (Figure 3.7) and the height (Figure 3.8) of ten plants per independent line was measured. Based on the lines that showed phenotypic complementation, *SkKNOX1*, *SkKNOX2* and *CrKNOX1* have the potential to rescue the *bp* mutant phenotype whereas *CrKNOX2* cannot complement.

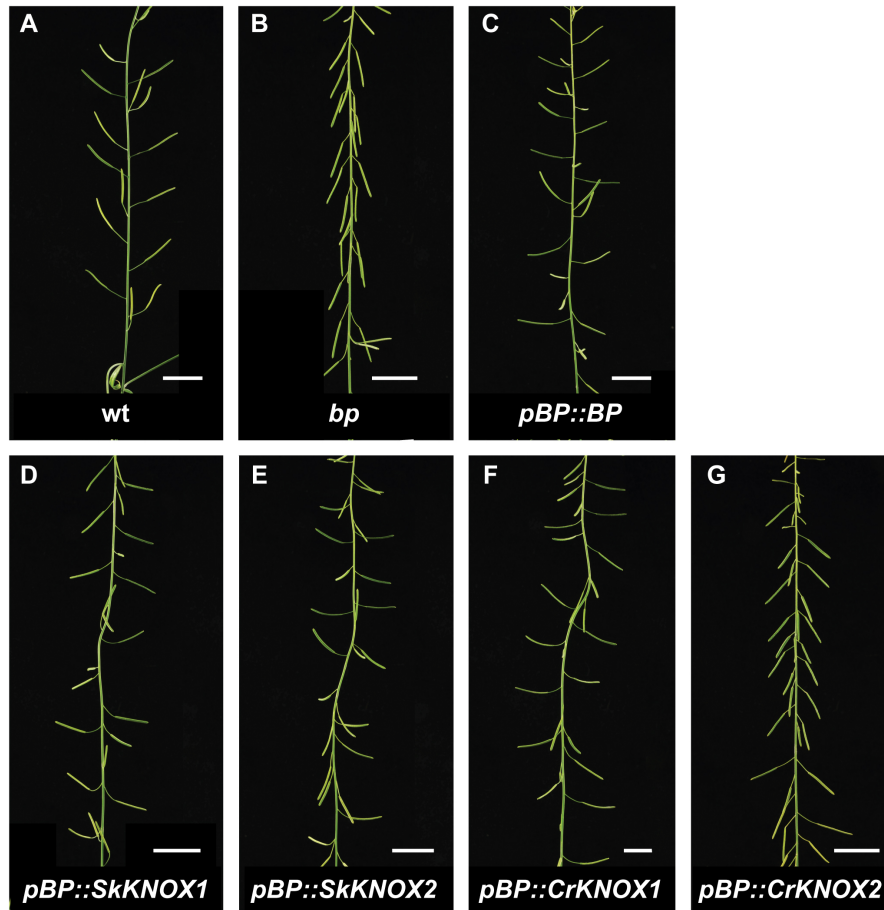
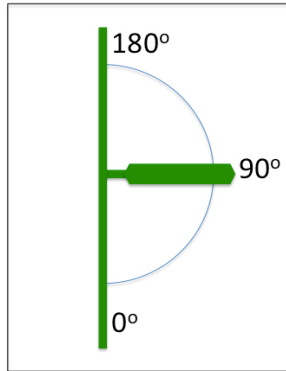
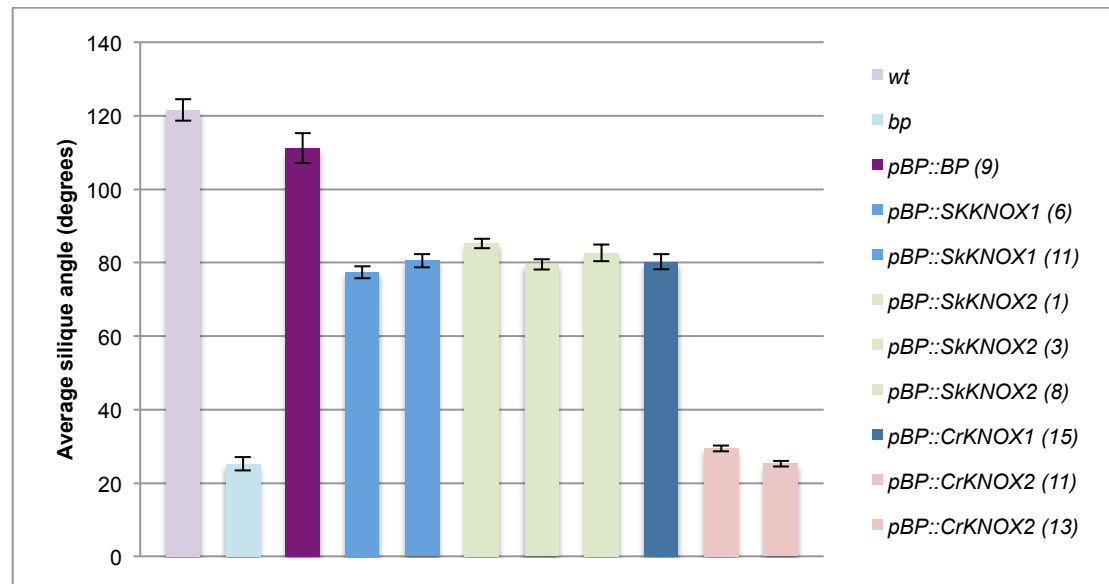


Figure 3.6

A) Wild type plant. B) *bp* mutant plant. C-G) Phenotype of homozygous *bp* mutant plants transformed with *pBP::BP* (C), *pBP::SkKNOX1* (D), *pBP::SkKNOX2* (E), *pBP::CrKNOX1* (F) and *pBP::CrKNOX2* (G). Plants were grown for 8 weeks. Only representatives of lines with evident phenotypic complementation are shown. The phenotype of the rest of the lines is similar to that of *bp* mutants. Scale bar 2 cm.

A**B****Figure 3.7**

A) Schematic to illustrate measurement of silique angles. B) Average silique angle (primary inflorescence) of wild type, *bp*, *pBP::BP*, *pBP::SkKNOX* and *pBP::CrKNOX* mutant plants. Ten plants per line were analyzed. Plants were grown for 8 weeks. Line identification number is indicated within a parenthesis next to the construct name. Error bars indicate standard error of the mean.

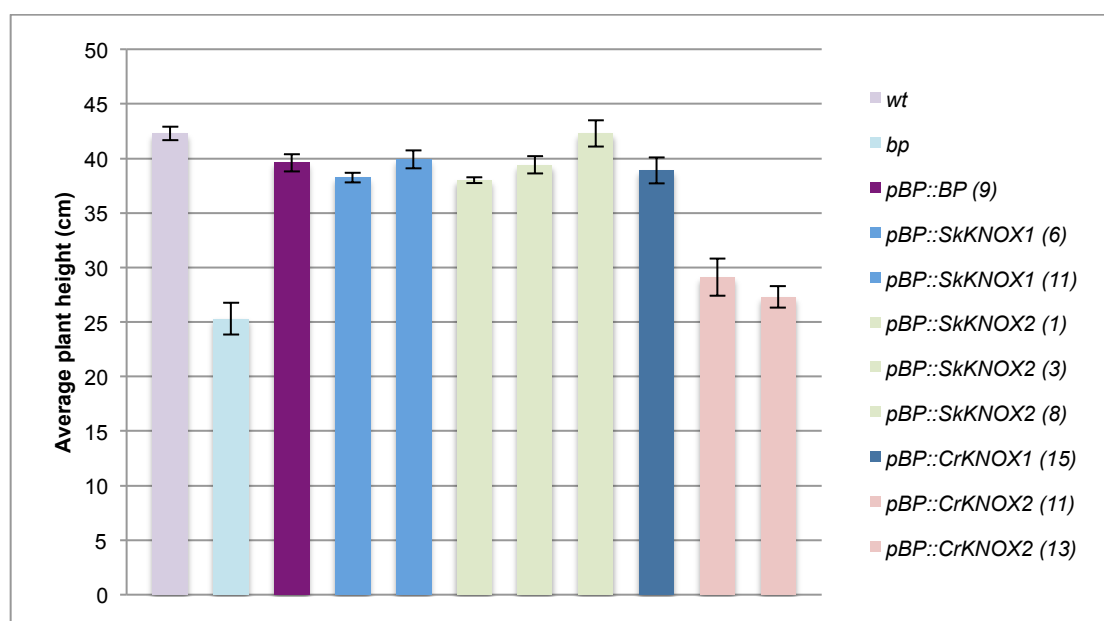


Figure 3.8

Average height of wild type, *bp*, *pBP::BP*, *pBP::SkKNOX* and *pBP::CrKNOX* mutant plants (primary inflorescence length was measured). Ten plants per line were analyzed. Plants were grown for 8 weeks. Line identification number is indicated within a parenthesis next to the construct name. Error bars indicate standard error of the mean.

3.6.3 Analysis of T₂ segregating *pBP::KNOX* lines

For the T₂ segregating *pBP::KNOX* lines, transgene carrying individuals were selected first on kanamycin and resistant plants (putative hemizygous and homozygous for the transgene) were transferred onto soil and grown for 8 weeks in a 16 h light : 8 h dark regime. Some lines manifested two phenotypic classes (Table 3.3), probably correlating with homozygous and hemizygous plants. In one class the potential of transgenes to rescue the mutant phenotype was higher relative to the potential of the second class. The reason why some lines have only one class could be the low number of plants examined (five to eight). Figure 3.9 shows representatives of *pBP::SkKNOX* and *pBP::CrKNOX* transgenic plants with the higher potential to rescue the mutant phenotype plus wild type and *bp* mutant plants at 8 weeks post germination. Five to eight plants per independent line were analysed for silique angle on the main inflorescence (Figure 3.10) and plant height (Figure 3.11).

Phenotypic complementation was observed in two of the three *pBP::SkKNOX1* lines, the single *pBP::SkKNOX2* line, only two of the five *pBP::CrKNOX1* lines and neither of the two *pBP::CrKNOX2* lines examined. All of the five *pBP::BP* control lines showed partial complementation. *SkKNOX1*, *SkKNOX2* and *CrKNOX1* thus have the potential to rescue the *bp* mutant phenotype, whereas, *CrKNOX2* does not.

A summary of the phenotypic analyses for both homozygous and segregating T₂ lines is presented in Table 3.4.

Construct	Line	Manifesting two distinct classes of phenotypes
<i>pBP::BP</i>	35	yes
	37	yes
	39	no
	42	yes
	46	no
<i>pBP::SkKNOX1</i>	16	no
	21	yes
	24	no
<i>pBP::SkKNOX2</i>	22	no
<i>pBP::CrKNOX1</i>	25	no
	26	yes
	36	no
	39	no
	44	yes
<i>pBP::CrKNOX2</i>	20	no
	21	no
	33	no

Table 3.3

Summary of *pBP::BP*, *pBP::SkKNOX* and *pBP::CrKNOX* T₂ segregating lines manifesting two different phenotypic classes. In one class the potential to complement the *bp* mutant phenotype was higher relative to the other class.



Figure 3.9

Phenotype of representatives from T₂ segregating *pBP::KNOX* lines that exhibit a higher potential to complement the *bp* mutant phenotype (putative homozygous plants). A) Wild type plant, B) *bp* mutant plant, C) *pBP::BP*, D) *pBP::SkKNOX1*, E) *pBP::SkKNOX2*, F) *pBP::CrKNOX1* and G) *pBP::CrKNOX2*. Plants were grown for 8 weeks. The phenotype of the rest of the lines is similar to that of *bp* mutants. Scale bar 2 cm.

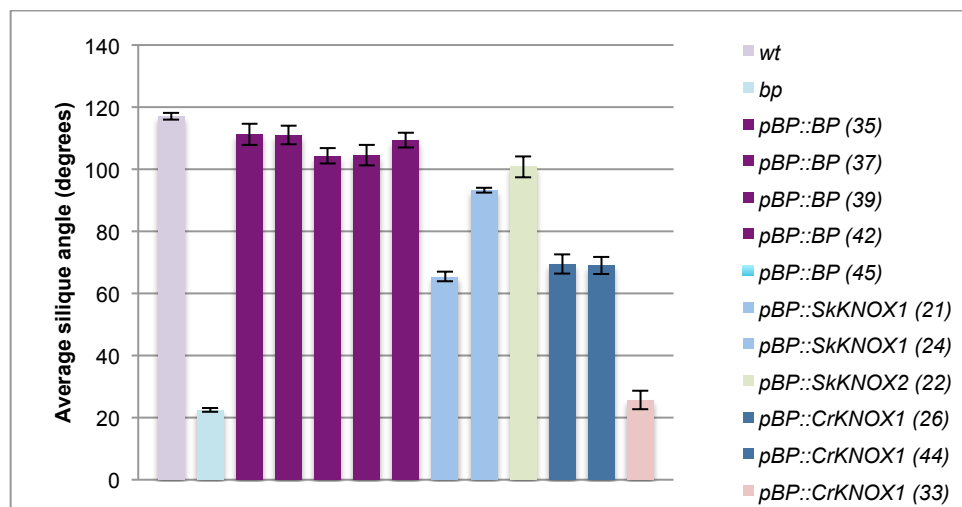


Figure 3.10

Average silique angle (primary inflorescence) of wild type, *bp*, *pBP::BP*, *pBP::SkKNOX* and *pBP::CrKNOX* mutant plants. Five – eight plants per line were analyzed. Plants were grown for 8 weeks. Line identification numbers are indicated within a parenthesis next to the construct name. Error bars indicate standard error of the mean.

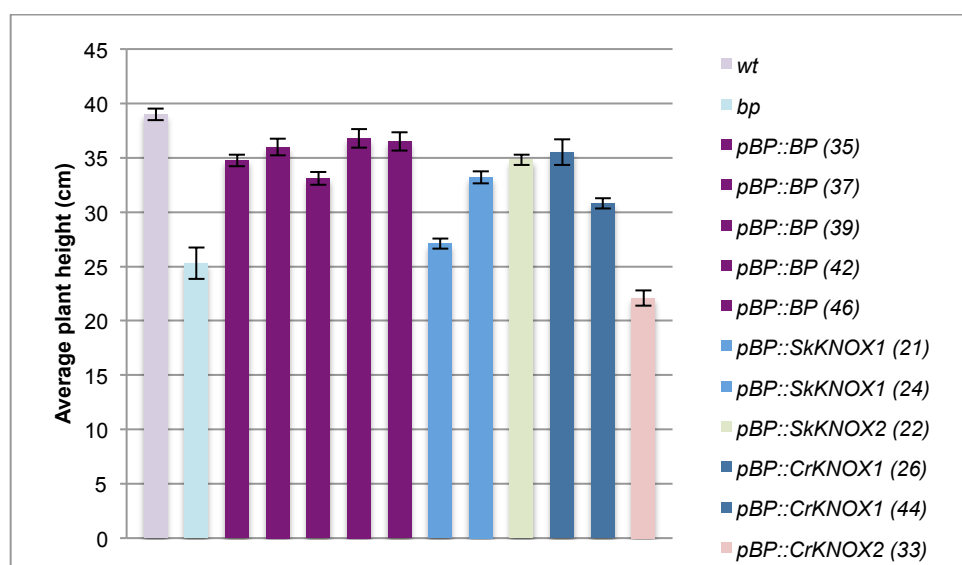


Figure 3.11

Average height of wild type, *bp*, *pBP::BP*, *pBP::SkKNOX* and *pBP::CrKNOX* mutant plants (primary inflorescence length was measured). Five – eight plants per line were analyzed. Plants were grown for 8 weeks. Line identification numbers are indicated within a parenthesis next to the construct name. Error bars indicate standard error of the mean.

Construct	Line	Number of inserts	Expression	Complementation
<i>pBP::BP</i>	1	not checked	++++	no
	4	5	++++	no
	9	not checked	++++	yes
	35	not checked	not checked	yes
	37	not checked	not checked	yes
	39	not checked	not checked	yes
	42	not checked	not checked	yes
	46	not checked	not checked	yes
<i>pBP::SkKNOX1</i>	6	not checked	++++	yes
	11	5	+++	yes
	16	not checked	not checked	no
	21	not checked	not checked	yes
	24	not checked	not checked	yes
<i>pBP::SkKNOX2</i>	1	5	++++	yes
	3	4	++++	yes
	4	multiple	+++	no
	8	5	++++	yes
	22	not checked	not checked	yes
<i>pBP::CrKNOX1</i>	3	5	+++	no
	6	not checked	+++	no
	13	5	+++	no
	15	not checked	++++	yes
	25	not checked	not checked	no
	26	not checked	not checked	yes
	36	not checked	not checked	no
	39	not checked	not checked	no
	44	not checked	not checked	yes
<i>pBP::CrKNOX2</i>	11	2	+++	no
	13	not checked	+++	no
	20	not checked	not checked	no
	21	not checked	not checked	no
	33	not checked	not checked	no

Table 3.4

Comparison of T₃ homozygous (highlighted with blue) and T₂ segregating (highlighted with purple) lines examined in this study. Table shows line number, number of inserts, relative expression levels (++++: higher in comparison to the rest of the lines, +++: lower in comparison to the rest of the lines) and complementation potential.

3.7 Discussion

The phenotypic characterization of two independent homozygous *pBP::SkKNOX1* and three independent homozygous *pBP::SkKNOX2* lines indicates that both *SkKNOX* genes can complement the phenotype of *Arabidopsis bp* mutants. This conclusion is further supported by the phenotypes observed in two independent T₂ segregating *pBP::SkKNOX1* lines and one T₂ segregating *pBP::SkKNOX2* line. However, more independent homozygous lines need to be characterized to substantiate this finding and to determine whether both *SkKNOX* genes have equivalent potential to rescue the *bp* mutant phenotype. The observation that some homozygous lines do not rescue the *bp* mutant phenotype also has to be taken into consideration. One explanation is that the transgene is silenced. *BP* is expressed during both the vegetative and the reproductive phases of *Arabidopsis* development. The RNA used for the RT-PCR assays was extracted from the vegetative phase of plants because it was easier to harvest tissue at this stage. However, the *bp* mutant phenotype is exhibited after bolting. Consequently, if silencing occurred only during the reproductive stage, phenotypic complementation would not be observed. Southern blot analysis showed that multiple transgene inserts were present in the lines characterized. Given that phenotypes do not differ significantly between lines, it could be inferred that there is no significant dosage effect of transgene expression. One reason for this could be that KNOX proteins have to form heterodimers with BELL proteins to perform their function. Given that the number of available BELL proteins is not altered, excessive KNOX protein accumulation would not necessarily have an effect.

In the case of *CrKNOX* genes it is more difficult to interpret the results. Only one of the four homozygous lines and two of the five T₂ segregating lines investigated showed the potential to complement the *bp* mutant phenotype.

Notably, *BP* did not fully complement the *bp* mutant phenotype in all control transgenic lines. This could be due to the fact that the promoter sequence cloned and used in the constructs does not contain all of the *cis*-regulatory elements that control *BP* expression. Additionally, it has to be taken into consideration that only the coding sequence of *BP* was used in the construct and thus potential intron regulatory elements were not present to achieve native levels of *BP* expression. However, given that transgene transcripts were detected in lines that were not able to complement the *bp* mutant phenotype, it is also possible that gene silencing is mediated through dicer-mediated inhibition of translation (Hutvagner and Simard, 2008; Ghildiyal and Zamore, 2009). Finally, the difference between wild type and *pBP::BP* transgenic phenotypes could be explained by the fact that the wild type *Arabidopsis* (*Col-0* ecotype) plants used as a reference were different from the original wild type *Arabidopsis* (*Col-0* ecotype) plants that were used to generate the *bp* mutant in the USA (Smith and Hake, 2003). Thus, it is possible that the Oxford wild type *Arabidopsis* line has accumulated random mutations resulting in a slightly altered morphology compared to the line used for *bp* mutant generation.

Homozygous *stm-1* mutant plants do not form a SAM and growth terminates at the seedling stage. In contrast, heterozygous *stm-1* mutant plants do not exhibit any developmental defects (Long *et al.*, 1996). *pSTM::KNOX* and *35S::KNOX* constructs were therefore transformed into segregating *stm-1* mutant plants (population of heterozygous and wild type plants). Since wild type plants were also included in the plants used for transformation, it is unlikely that the inability to obtain transformed plants is due to mutant resistance to transformation.

In summary, the preliminary results presented in this chapter suggest that the function of class 1 KNOX proteins is likely to be conserved between lycophytes and angiosperms. However, class 1 KNOX function in ferns remains unclear.

4 Comparative analysis of class 1 *KNOX* genes in bryophytes and vascular plants

4.1 Introduction

In the moss *P. patens*, sporophyte development requires three class 1 *KNOX* genes; Moss *KNOTTED1-like 2* (*MKN2*), 4 and 5, as indicated by defects in sporophyte morphology observed when all three genes are mutated (Champagne and Ashton, 2001; Sakakibara et al., 2008). *MKN2* appears to be the dominant paralogue as *mkn2* single mutant and triple *mkn2;mkn4;mkn5* mutant phenotype are similar (Sakakibara et al., 2008). When *MKN2* is expressed in *Arabidopsis* under the control of the *CaMV* 35S promoter, the effect on leaf morphology is similar to that seen when the *Arabidopsis* class 1 *KNOX* genes *BP* or *KNAT2* are overexpressed. However, in *P. patens mkn2;mkn4;mkn5* triple mutant plants there is no effect on the expression of homologs of the *GA2ox1* and *GA20ox1* genes that regulate GA biosynthesis and which in *Arabidopsis* are downstream targets of class 1 *KNOX* activity (Sakakibara et al., 2008). This observation suggests that the relationship between class 1 *KNOX* genes and their downstream effectors may have changed during land plant evolution.

4.2 Aims

The aim of the work described in this chapter was to determine the extent to which class 1 *KNOX* gene function is conserved between the angiosperm *Arabidopsis*, the lycophyte *S. kraussiana*, the fern *C. richardii* and the moss *P. patens*. The objective was to determine whether *BP*, *STM*, *SkKNOX1*, *SkKNOX2*, *CrKNOX1* and *CrKNOX2* genes are able to complement *P. patens mkn2* single knockout and *mkn2;mkn4;mkn5* triple knockout mutant phenotypes (Sakakibara et al., 2008), when expressed under the control of the endogenous *MKN2* gene promoter.

4.3 Generation of constructs for transformation

Seven constructs were designed to express *MKN2*, *BP*, *STM*, *SkKNOX1*, *SkKNOX2*, *CrKNOX1* and *CrKNOX2* genes under the control of the endogenous *MKN2* gene promoter. For the generation of *mkn2* single mutants (that was used as a background for the generation of the triple *mkn2;mkn4;mkn5* mutants) the entire genomic sequence of *MKN2* and an approximately 200 bp fragment of 5' *MKN2* flanking sequence was replaced by the G418 resistance cassette (Sakakibara et al., 2008). In order to reintroduce this 200 bp fragment into the endogenous locus, this fragment in addition to 1 kb sequence of 5' *MKN2* flanking sequence, was used as homologous sequence for recombination. The 1.2 kb fragment was fused with the coding sequence of *MKN2*, *BP*, *STM*, *SkKNOX1*, *SkKNOX2*, *CrKNOX1* and *CrKNOX2* genes. The *NOPALINE SYNTHASE* (*NOS*) terminator was fused 3' to the coding sequence followed by the blasticidin resistance cassette and then a 1.8 kb fragment of 3' *MKN2* flanking sequence as homologous sequence for recombination. Expression of the blasticidin resistance gene was driven by the 35S *CaMV* promoter. Upon integration into the moss genome, expression of the various *KNOX* genes should be controlled by the *MKN2* *cis*-regulatory elements and the *NOS* terminator sequences. Figure 4.1 shows the *pMKN2::MKN2*, *pMKN2::BP*, *pMKN2::STM*, *pMKN2::SkKNOX1*, *pMKN2::SkKNOX2*, *pMKN2::CrKNOX1* and *pMKN2::CrKNOX2* constructs (hereafter called *pMKN2::KNOX*).

4.4 Generation of transgenic lines

Single and triple mutants were originally generated by the insertion of G418, hygromycin and zeomycin antibiotic resistance cassettes into the *MKN2*, *MKN4* and *MKN5* loci respectively (Sakakibara et al., 2008). In each case, the entire genomic sequence was replaced by the antibiotic resistance cassette.

In five to seven independent transformation experiments, the *pMKN2::KNOX* constructs were introduced into *mkn2* single and *mkn2;mkn4;mkn5* triple mutant protoplasts (Figure 4.2). Following transformation and regeneration under blasticidin selection, a series of stable transformed lines were isolated (Table 4.1).

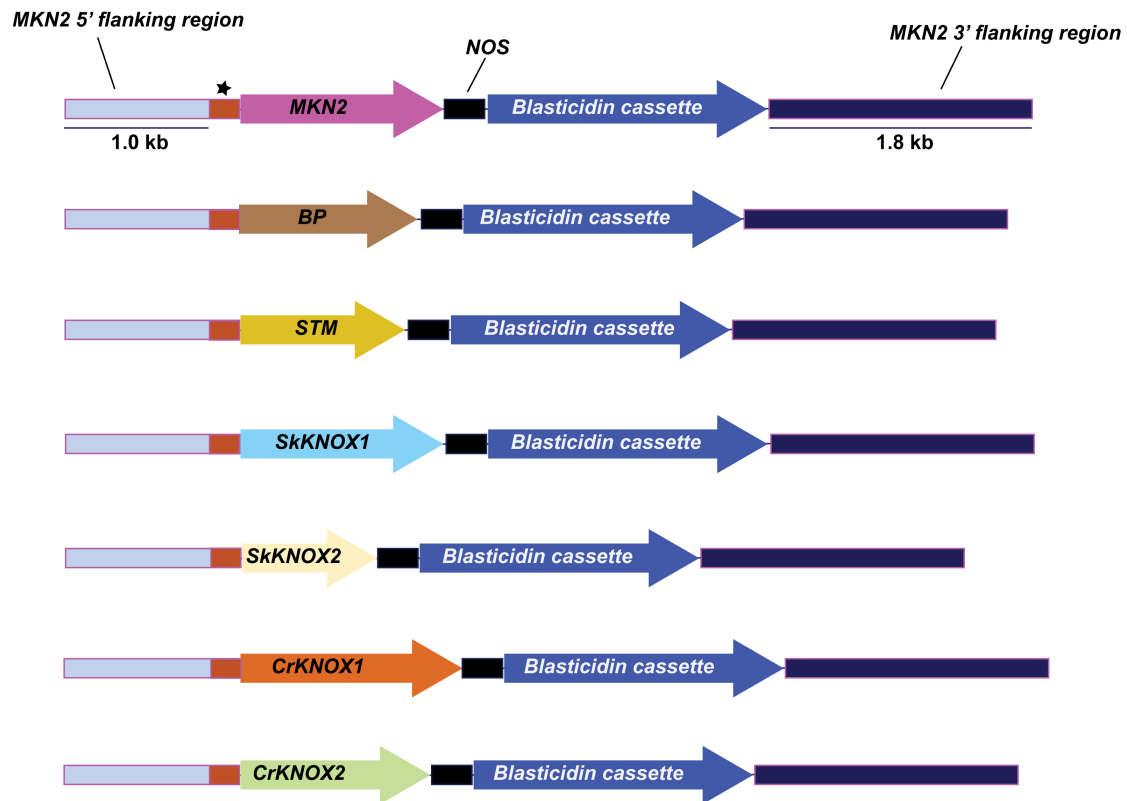


Figure 4.1

Schematic representation of *pMKN2::MKN2*, *pMKN2::BP*, *pMKN2::STM*, *pMKN2::SkKNOX1*, *pMKN2::SkKNOX2*, *pMKN2::CrKNOX1* and *pMKN2::CrKNOX2* constructs. *MKN2* 5' (blue boxes) and 3' flanking regions (dark purple boxes), various *KNOX* gene coding sequences (coloured arrows), the *NOS* terminator (black boxes) and the blasticidin resistance cassette (dark blue arrow) are indicated. The re-introduced 200 bp fragment of 5' *MKN2* flanking sequence that was replaced by the G418 resistance cassette is indicated by a red box with a black star above.

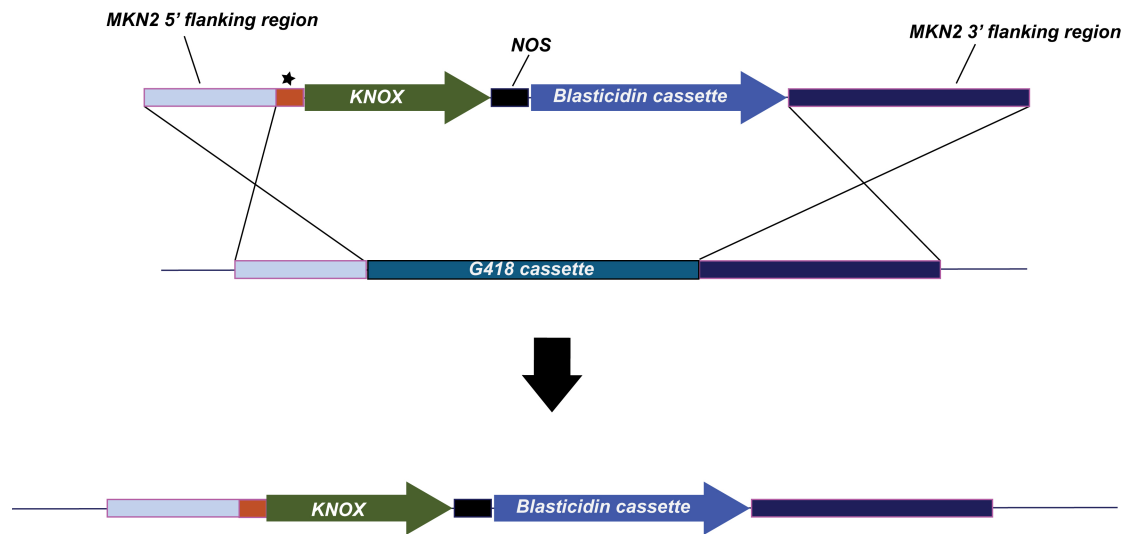


Figure 4.2

Schematic representation of *pMKN2::KNOX* integration. For the generation of *mkn2* mutants the entire coding sequence of *MKN2* and an approximately 200 bp fragment of 5' *MKN2* flanking sequence (indicated by a red box with a black star above) was replaced by the G418 resistance cassette. For the generation of the *pMKN2::KNOX* lines the kanamycin resistant cassette was replaced via homologous recombination by the *pMKN2::KNOX* constructs. *MKN2* 5' (blue boxes) and 3' flanking regions (dark purple boxes), *KNOX* gene coding sequences (green arrows), the *NOS* terminator (black boxes), the *blasticidin* resistance cassette (dark blue arrow) and the G418 cassette (blue box) are indicated.

4.5 Analysis of transformed lines

4.5.1 Genetic analysis

After generating stable transgenic moss lines, each one was analysed by PCR to check that the construct was integrated into the desired locus, and not elsewhere in the genome. In the *mkn2* background, 9 lines with the correct integration event were isolated for *pMKN2::MKN2* construct, 7 for *pMKN2::BP* construct, 9 for *pMKN2::STM*, 11 for *pMKN2::SkKNOX1*, 7 for *pMKN2::SkKNOX2*, 5 for *pMKN2::CrKNOX1* and 10 for *pMKN2::CrKNOX2*. In the *mkn2;mkn4;mkn5* background, 8 lines with the correct integration event were isolated for the *pMKN2::BP* construct, 11 for *pMKN2::MKN2*, 8 for *pMKN2::STM*, 15 for *pMKN2::SkKNOX1*, 10 for *pMKN2::SkKNOX2*, 15 for *pMKN2::CrKNOX1* and 14 for *pMKN2::CrKNOX2* (Table 4.1). Lines transformed with the *pMKN2::BP* and *pMKN2::STM* constructs were deprioritized for phenotypic analysis in favour of lines transformed with the *pMKN2::SkKNOX* and *pMKN2::CrKNOX* constructs, and so were not included in subsequent experiments.

To determine the transgene copy number in each line, a 510 bp fragment of the blasticidin resistance gene was hybridized to *HindIII* restricted genomic DNA on Southern blots (Figure 4.3). If the construct was correctly targeted to the *MKN2* locus, then a single fragment should be detected after hybridization. Analyses of insertion copy number were performed on 10 stable lines, with the exception of *pMKN2::MKN2*, *pMKN2::SkKNOX2* and *pMKN2::CrKNOX1* into the *mkn2* mutant background for which 9, 7 and 5 lines were examined, respectively. Genomic DNA from the *mkn2* mutant and the *mkn2;mkn4;mkn5* triple mutant lines was used as negative controls (Figures 4.4 – 4.8).

For each construct, two independent lines with a single transgene insertion were selected for further characterization (hereafter called *pMKN2::SkKNOX1a* and *b*, *pMKN2::CrKNOX1a* and *b*, *pMKN2::CrKNOX2a* and *b* and *pMKN2::MKN2a* and *b*), with the exception of *mkn2* mutants transformed with the *pMKN2::SkKNOX2* construct for which 3 lines with 2 inserts (one at the *MKN2* locus and another one in a different unknown locus) were selected for further characterization (hereafter called *pMKN2::SkKNOX1a*, *b* and *c*) (Figure 4.9).

Mutant Background	Construct	Number of stable lines	Number of lines with correct targeting	Number of lines with single transgene insert
<i>mkn2</i>				
	<i>pMKN2::MKN2</i>	44	9	2
	<i>pMKN2::BP</i>	34	7	not checked
	<i>pMKN2::STM</i>	43	9	not checked
	<i>pMKN2::SkKNOX1</i>	36	11	2
	<i>pMKN2::SkKNOX2</i>	38	7	0
	<i>pMKN2::CrKNOX1</i>	33	5	4
	<i>pMKN2::CrKNOX2</i>	47	10	4
	Total number	275	58	12
<i>mkn2;mkn4;mkn5</i>				
	<i>pMKN2::MKN2</i>	56	11	5
	<i>pMKN2::BP</i>	36	8	not checked
	<i>pMKN2::STM</i>	62	8	not checked
	<i>pMKN2::SkKNOX1</i>	48	15	7
	<i>pMKN2::SkKNOX2</i>	53	10	5
	<i>pMKN2::CrKNOX1</i>	70	15	6
	<i>pMKN2::CrKNOX2</i>	68	14	3
	Total number	393	81	26

Table 4.1

Summary of genotypic analysis of *mkn2* and *mkn2;mkn4;mkn5* mutants transformed with the *pMKN2::KNOX* constructs.

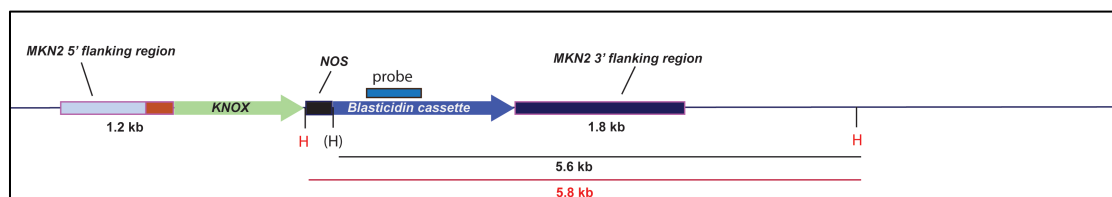


Figure 4.3

Schematic representation of *pMKN2::KNOX* constructs targeted into the *MKN2* locus. *Hind*III restriction sites are represented with the capital letter "H". A *Hind*III restriction site that is present only in the *pMKN2::SkKNOX1* construct is indicated with "(H)". The probe used for Southern blot hybridization is represented with a blue box above the blasticidin cassette.

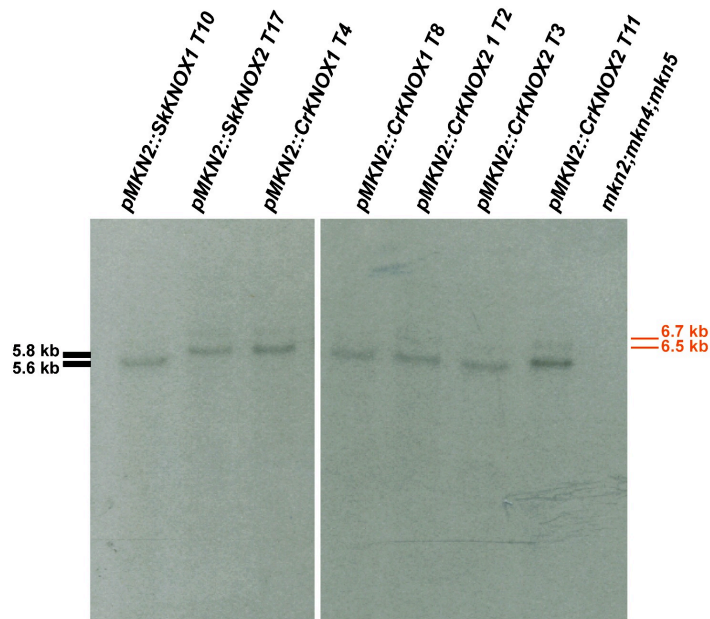


Figure 4.4

Southern blot analysis of *Hind*III digested genomic DNA from *mkn2;mkn4;mkn5* mutants transformed with the *pMKN2::KNOX* constructs. One fragment that corresponds to the expected 5.8 kb fragment (and 5.6 kb fragment for the *pMKN2::SkKNOX1* construct) is detected. An additional fragment that is about 0.9 kb larger was also detected in all of the lines examined. The stoichiometry of the detected fragments, and the fact that there is a second *Hind*III restriction site about 0.9 kb downstream of the first *Hind*III restriction 5' to the construct integration locus, suggest that the restriction digest is incomplete. Multiple attempts to achieve complete digestion of genomic DNA were not successful. The letter next to the name of each construct indicates the mutant background (T: *mkn2;mkn4;mkn5*) and the accompanying number is the identification number for each independent line.

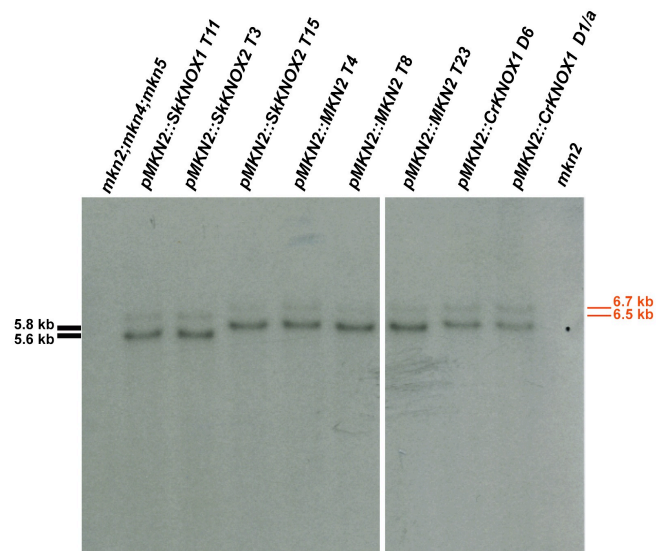


Figure 4.5

Southern blot analysis of *Hind*III digested genomic DNA from *mkn2;mkn4;mkn5* and *mkn2* mutants transformed with the *pMKN2::KNOX* constructs. The letter next to the name of each construct indicates the mutant background (D: *mkn2* and T: *mkn2;mkn4;mkn5*) and the number is the identification number for each independent line. *pMKN2::CrKNOX1 D1* hereafter called *pMKN2::CrKNOX1a*.

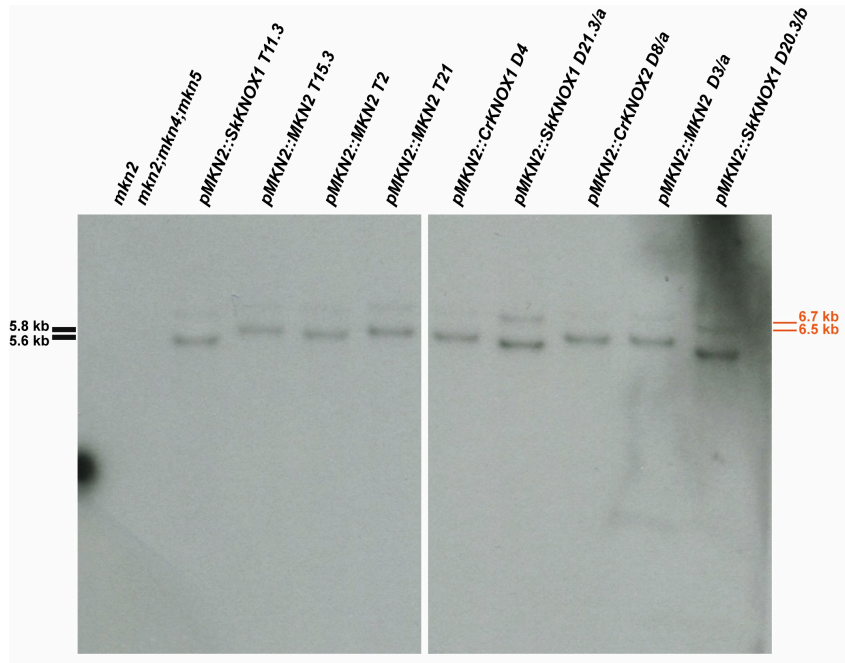


Figure 4.6

Southern blot analysis of *Hind*III digested genomic DNA from *mkn2;mkn4;mkn5* and *mkn2* mutants transformed with the *pMKN2::KNOX* constructs. *pMKN2::SkKNOX1 D21.3* hereafter called *pMKN2::SkKNOX1a*, *pMKN2::SkKNOX1 D20.3* hereafter called *pMKN2::SkKNOX1b*, *pMKN2::CrKNOX2 D8* hereafter called *pMKN2::CrKNOX2a*, *pMKN2::MKN2 D3* hereafter called *pMKN2::MKN2a* (*mkn2;mkn4;mkn5* and *mkn2* controls loaded in the same lane).

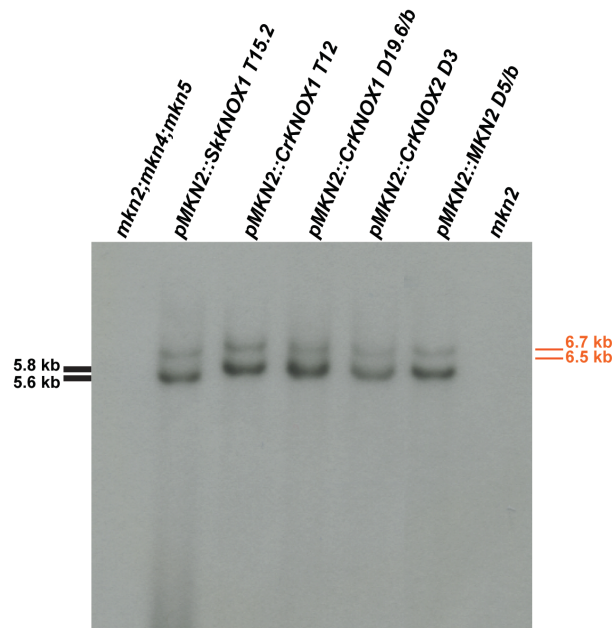


Figure 4.7

Southern blot analysis of *Hind*III digested genomic DNA from *mkn2;mkn4;mkn5* and *mkn2* mutants transformed with the *pMKN2::KNOX* constructs *pMKN2::CrKNOX1 D19.6* hereafter called *pMKN2::CrKNOX1b* and *pMKN2::MKN2 D5* hereafter called *pMKN2::MKN2b*.

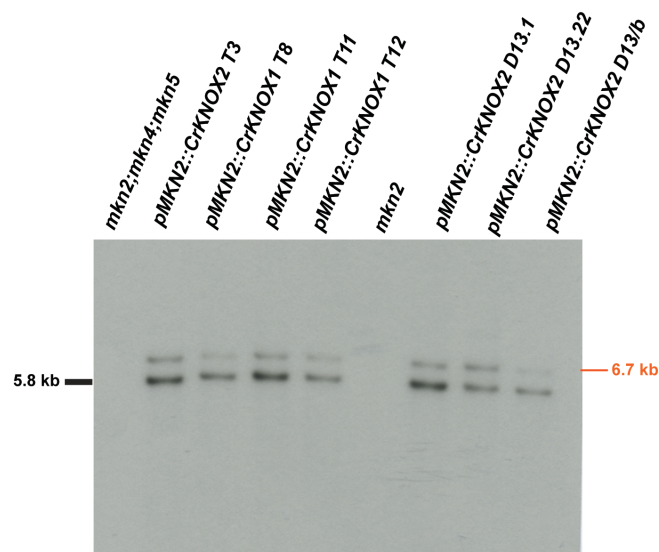


Figure 4.8

Southern blot analysis of *Hind*III digested genomic DNA from *mkn2;mkn4;mkn5* and *mkn2* mutants transformed with the *pMKN2::KNOX* constructs. *pMKN2::CrKNOX2 D13* hereafter called *pMKN2::CrKNOX2b*.

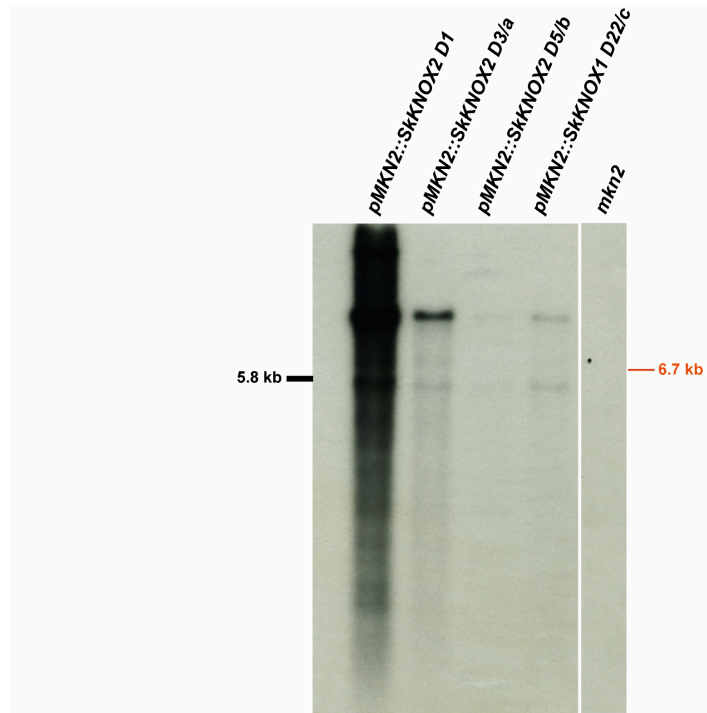


Figure 4.9

Southern blot analysis of *Hind*III digested genomic DNA from *mkn2* mutants transformed with the *pMKN2::SkKNOX2* construct. *pMKN2::SkKNOX2 D5*, *D3* and *D22* hereafter called *pMKN2::SkKNOX2a*, *b* and *c* respectively.

4.5.2 Expression analysis

During wild type development in *P. patens*, *MKN2* is initially expressed in the unfertilized egg. After fertilization *MKN2* is expressed in the apical part of the developing sporophyte and then at the seta meristem. Gene expression ceases as the sporophyte matures and growth ceases (Sakakibara et al., 2008).

To determine whether the *S. kraussiana* and *C. richardii* class 1 *KNOX* genes were expressed in the respective transgenic moss lines, RT-PCR analyses were carried out for each independent line using RNA extracted from approximately 300 archegonia, that contained either unfertilized eggs or early developing embryos. cDNA was synthesized using approximately 400 ng of total RNA per sample. Primer pairs (Appendix 1) designed to amplify an approximately 500 bp fragment of the N-terminal domain of the *S. kraussiana* and *C. richardii* class 1 *KNOX* genes, and *MKN2*, were used to check transcript accumulation. Amplification of the *Glycolytic glyceraldehyde-3-phosphate dehydrogenase* (*GAPC1*) gene (X72381) was used as an internal control. Each PCR was performed with both sets of primer pairs and adjusted so that the reactions were in the exponential phase. However, optimization of RT-PCR conditions to generate identical levels of *GAPC1* amplification was not possible due to the difficulty of extracting sufficient amounts of RNA.

Figure 4.10 shows that *KNOX* transcripts accumulate in all lines examined but not in the *mkn2* control. In the *pMKN2::SkKNOX2b* and *c* lines the fragment amplified was smaller than expected, suggesting that part of the sequence was probably lost upon integration (data not shown).

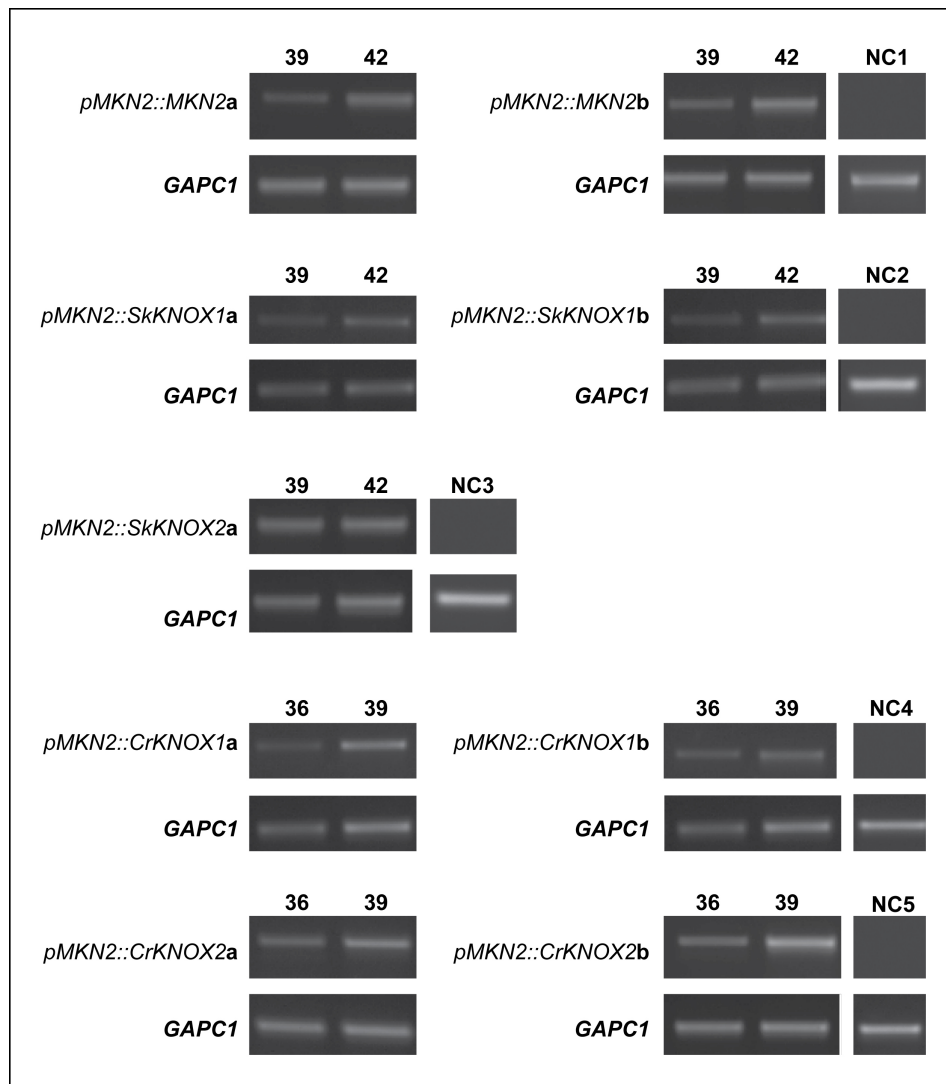


Figure 4.10

RT-PCR showing transcript accumulation in *mkn2* mutant plants transformed with the *pMKN2::MKN2*, *pMKN2::SkKNOX1*, *pMKN2::SkKNOX2*, *pMKN2::CrKNOX1* and *pMKN2::CrKNOX2* constructs. The *GAPC1* gene was used as an amplification and loading control. PCR assays were performed for an appropriate number of cycles (number of cycles is indicated above the gel photos) to achieve amplification in the exponential phase. RT-PCR assays for *pMKN2::SkKNOX1a*, *pMKN2::SkKNOX2a*, *pMKN2::CrKNOX1a*, *pMKN2::CrKNOX2a* and *pMKN2::MKN2a* were performed in an independent experiment from RT-PCR assays performed for *pMKN2::SkKNOX1b*, *pMKN2::CrKNOX1b*, *pMKN2::CrKNOX2b* and *pMKN2::MKN2b*. NC1-5: negative controls; NC1: RT-PCR assays using primers for *MKN2* N-terminal domain and *mkn2* mutant cDNA, NC2: RT-PCR assays using primers for *SkKNOX1* N-terminal domain and *mkn2* mutant cDNA, NC3: RT-PCR assays using primers for *SkKNOX2* N-terminal domain and *mkn2* cDNA, NC4: RT-PCR assays using primers for *CrKNOX1* N-terminal domain and *mkn2* mutant cDNA, NC5: RT-PCR assays using primers for *CrKNOX2* N-terminus domain and *mkn2* mutant cDNA.

4.6 Phenotypic characterization of transgenic plants

Wild type *P. patens* sporophytes exhibit radial symmetry, a distinct border between the seta and foot, and an even distribution of spores within the sporangium. When all three class 1 *KNOX* genes are knocked out in *P. patens* the mature sporophyte is smaller in size compared to wild type, radial symmetry is abolished, the border between the seta and foot is hard to distinguish, fewer spores are present, and spores are unequally distributed within the sporangium. A similar phenotype is observed in the single *mkn2* mutant (Sakakibara et al., 2008).

Phenotypic characterization of *mkn2;mkn4;mkn5* mutants transformed with the *pMKN2::KNOX* constructs was not possible because plants failed to fertilize and produce sporophytes. The lack of fertilization is likely due to technical issues rather than transgene effects given that the plants produced reproductive organs and that wild type control plants in the same batch also failed to fertilize. All phenotypic characterization was thus carried out with transgenic *mkn2* lines.

The *mkn2* mutant exhibits two phenotypic classes: a severe phenotype where the sporophyte is significantly malformed and a mild phenotype where the reduction in spore number is less obvious and the seta – foot border is more distinct (Figure 4.11). To assess the complementation potential of *SkKNOX* and *CrKNOX* genes, the phenotypes of mutant and transgenic moss sporophytes were compared. Photographs of approximately 40-50 sporophytes were taken and then the total surface area of the sporophyte was compared to the area occupied by spores (Figure 4.12A).

In all of the *pMKN2::SkKNOX1*, *pMKN2::SkKNOX2*, *pMKN2::CrKNOX1* and *pMKN2::CrKNOX2* lines examined, the frequency of sporophytes with a severe

mutant phenotype is lower than in the *mkn2* mutant control (Table 4.2). Sporophytes with a mild mutant phenotype are still present in all of the lines examined (Figure 4.12B). *pMKN2::CrKNOX2b* sporophytes with a mild phenotype show a reduced number of spores, compared to *pMKN2::CrKNOX2a* but this is possibly due to the lower number of sporophytes examined in comparison to the other lines (only twenty). Importantly, the *pMKN2::MKN2* control lines were indistinguishable from wild type sporophytes (Figure 4.11). Given time constraints, it was not possible to determine the precise level of phenotypic complementation, however, some general observations could be made (Figure 4.12B).

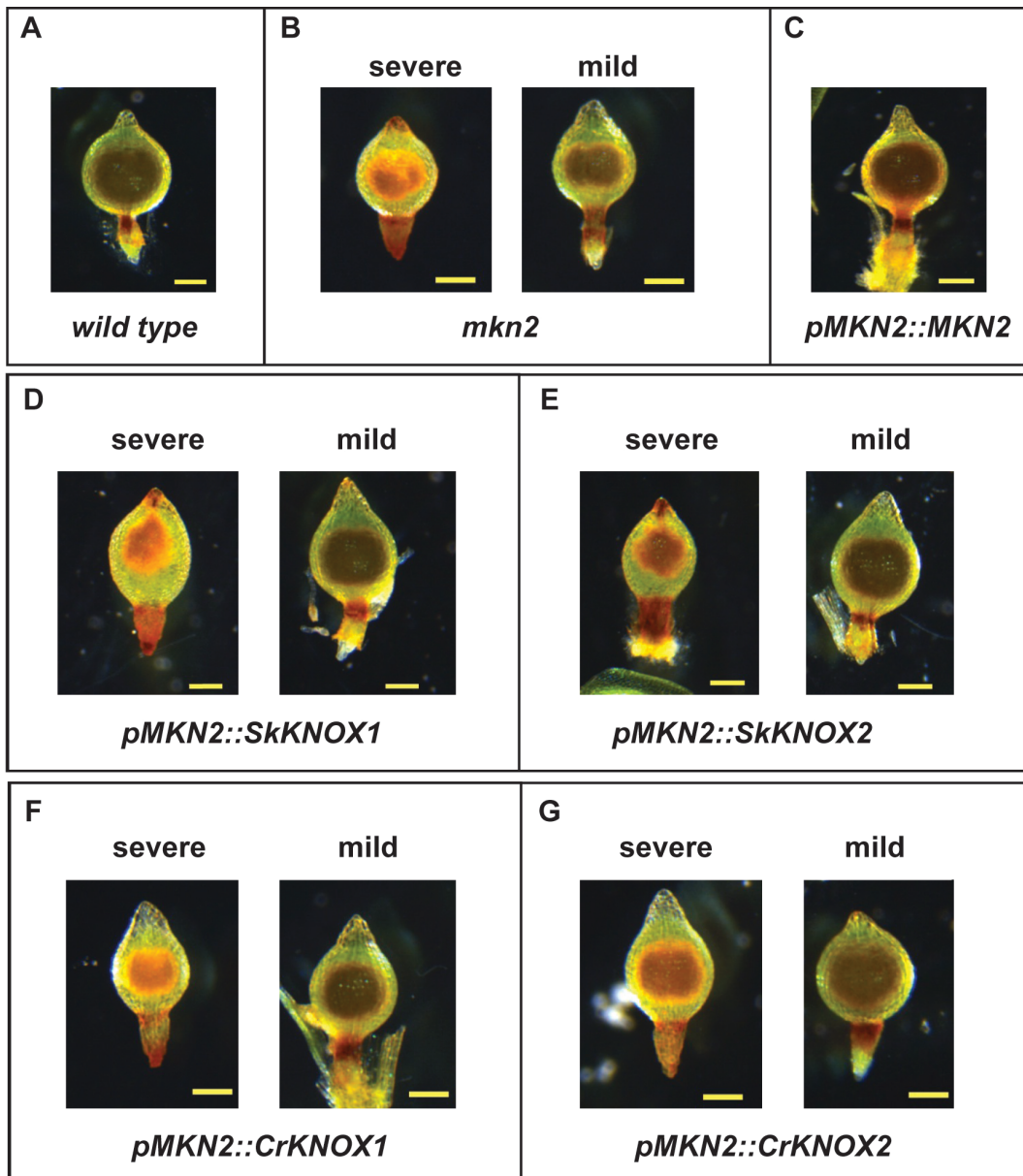


Figure 4.11

Phenotype of *mkn2* mutant plants transformed with the *pMKN2::MKN2*, *pMKN2::SkKNOX1*, *pMKN2::SkKNOX2*, *pMKN2::CrKNOX1* and *pMKN2::CrKNOX2* constructs. Representatives of sporophytes with strong and mild phenotypes are shown. Scale bar 200µm.

	Total number of sporophytes examined	Sporophytes with severe mutant phenotype	Percentage of sporophytes with severe mutant phenotype	Sporophytes with mild mutant phenotype	Percentage of sporophytes with mild mutant phenotype
wild type	43	0	0%	0	0%
<i>mkn2</i>	48	29	60.41%	19	39.58%
<i>pMKN2::MKN2a</i>	42	0	0%	0	0%
<i>pMKN2::MKN2b</i>	55	0	0%	0	0%
<i>pMKN2::SkKNOX1a</i>	45	2	4.44%	43	95.56%
<i>pMKN2::SkKNOX1b</i>	43	12	27.90%	31	72.09%
<i>pMKN2::SkKNOX2a</i>	47	21	44.68%	26	55.32%
<i>pMKN2::CrKNOX1a</i>	52	9	17.31%	43	82.69%
<i>pMKN2::CrKNOX1b</i>	30	6	20%	24	80%
<i>pMKN2::CrKNOX2a</i>	31	7	22.58%	24	77.42%
<i>pMKN2::CrKNOX2b</i>	21	4	19.05%	17	80.95%

Table 4.2

Summary of sporophyte numbers with severe and mild mutant phenotype.

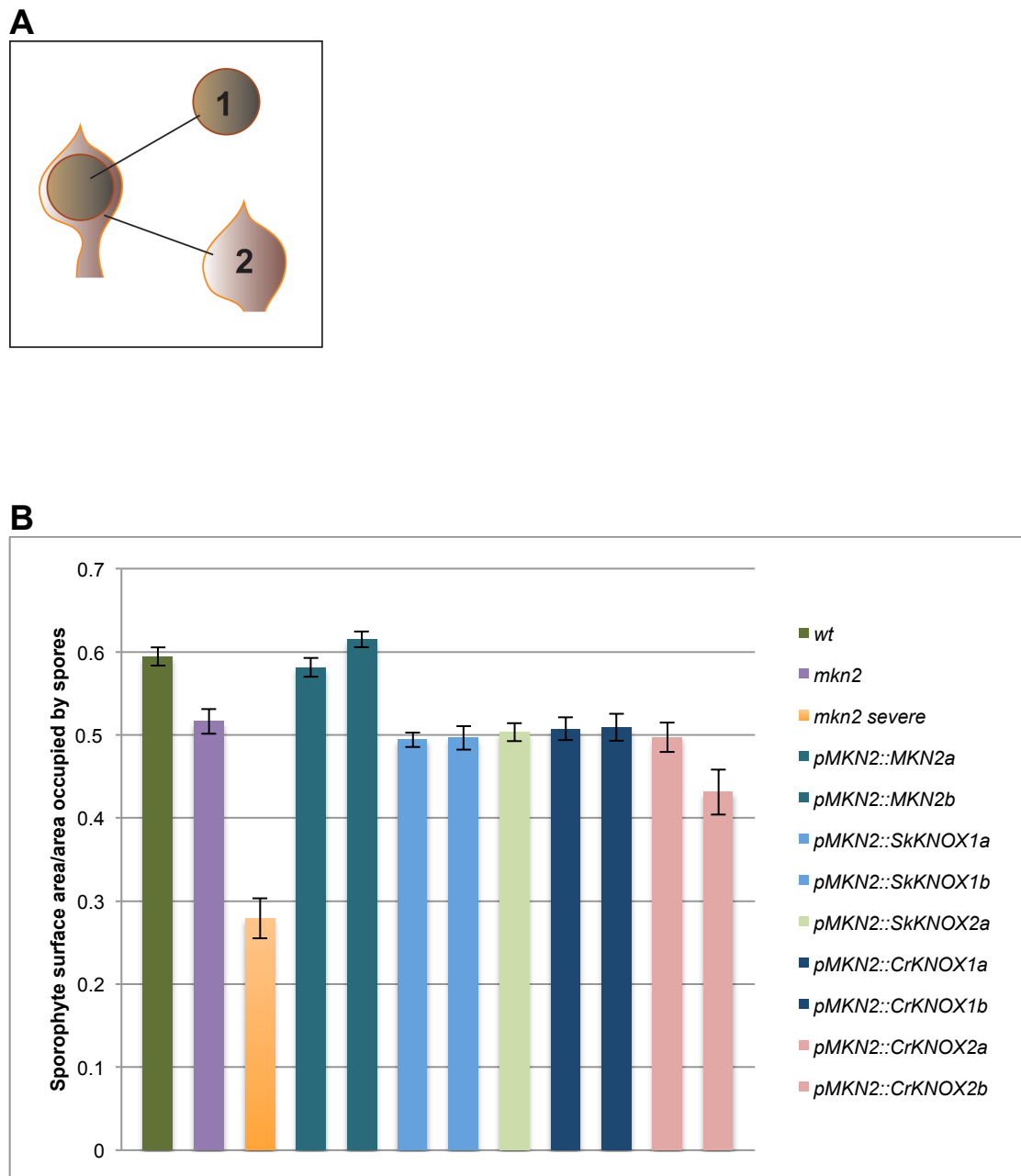


Figure 4.12

A) Schematic to illustrate measurement of sporophyte surface area (2) in comparison to the area occupied by spores (1). B) Average ratio of total sporophyte surface area to the area occupied by spores, for sporophytes with a mild phenotype. The ratio for sporophytes with a severe *mkn2* mutant phenotype is also indicated for comparison. Error bars represent standard error of the mean.

4.7 Discussion

The results presented in this chapter suggest that *S. kraussiana* and *C. richardii* *KNOX* genes cannot fully complement the *mkn2* mutant phenotype. One explanation for this observation is that the downstream targets of the respective *KNOX* genes have diversified during land plant evolution. Another possibility is that structural changes in either the DNA-binding domain or the MEINOX domain prevent SkKNOX and CrKNOX proteins from binding to downstream targets in *P. patens* and thus they were unable to complement the *mkn2* mutant phenotype. However, based on the fact that the number of sporophytes with a strong mutant phenotype is reduced in all of the lines examined, partial complementation by *SkKNOX* and *CrKNOX* does occur. Further phenotypic characterization needs to be carried out to determine the precise degree of complementation, something that was not possible in the current study due to the low fertilization rates of the lines. Specifically, four further experiments need to be carried out:

- To further analyze the degree of complementation of the mild phenotype, additional important aspects of sporophyte morphology, such as the width of the sporophyte seta, need to be examined. This will require longitudinal sectioning of the lower half of the mature sporophyte and measurement of the number of cells in the seta and foot. Additionally, a more precise method to measure the number of spores in each sporophyte should be considered. For example, spores could be removed from the sporangium, diluted in water and then counted using a haemocytometer. Sectioning of early embryos would provide insight into the potential of *SkKNOX* and *CrKNOX* genes to restore the planes of mitotic divisions in the developing embryo.
- Optimization of RT-PCR assays is necessary so that the level of transgene expression can be compared between lines. This is particularly important in

the case of *pMKN2::SkKNOX1* and *pMKN2::CrKNOX1* in the *mkn2* mutant background, where the number of sporophytes with a severe phenotype and the average ratio of total sporophyte surface area to the area occupied by spores differs between the two independent lines examined.

- *mkn2;mkn4;mkn5* triple mutants exhibit a similar but more severe mutant phenotype to that of *mkn2*. Thus, it will probably be easier to assess the level of complementation by studying the lines where *pMKN2::KNOX* constructs were transformed into the *mkn2;mkn4;mkn5* triple mutant background. This is particularly important in the case of *pMKN2::SkKNOX2* as there are lines with a single insert whereas there are only lines with multiple inserts in the *mkn2* mutant background.
- The examination of *pMKN2::BP* and *pMKN2::STM* mutant lines is needed to provide insight into the level of specialization of *KNOX* gene function between mosses and angiosperms.

5 Investigation of *KNOX* genes in hornworts

5.1 Introduction

The hornworts are the smallest clade within bryophytes, consisting of 200-250 species that are geographically widespread (Villarreal et al., 2010). The gametophyte generation is dominant, characterized by a flattened thallus that generally lacks specialized internal tissue differentiation. Uniquely amongst extant land plant lineages, hornworts have a single (or just a few) algal-like chloroplast(s) per cell and a symbiotic relationship with *Nostoc* cyanobacteria (Renzaglia et al., 2009). Comprehensive analysis of phylogenetic relationships between land plants resolves the hornworts as the sister group to vascular plants (Qiu et al., 2006). Hornworts are thus an important group to provide insight into the mechanisms that evolved to facilitate the progressive increase in sporophyte complexity in vascular plants.

At present, nothing is known in hornworts about the role of important regulators of sporophyte development, for example the *KNOX* gene family. Hornworts are particularly interesting candidates for the study of *KNOX* genes, not only due to their phylogenetic position as sister group to vascular plants, but also due to the differences in sporophyte development between hornworts and other land plants. Specifically, the hornwort sporophyte develops on the gametophyte, growing from a basal meristem that is established early in sporophyte development. At maturity the hornwort sporophyte is an elongated cylindrical structure that, unlike other bryophyte sporophytes, produces spores continuously throughout the growing season. The question arises therefore as to whether modifications to *KNOX* gene function played a role in the evolution of these unique hornwort sporophyte characters. The development of a hornwort experimental system would therefore be of significant scientific benefit.

5.2 Aims

The work presented in this chapter aimed to establish an experimental system for the study of *KNOX* genes in hornworts. The objectives were to:

- establish axenic cultures and optimize growth conditions
- develop protocols for nucleic acid extraction
- isolate *KNOX* gene sequences.

5.3 Establishment of a hornwort experimental system

5.3.1 Selection of hornwort species and establishment of axenic cultures

Anthoceros punctatus (Figure 5.1), and later *Anthoceros agrestis* (Figure 5.2), were selected as candidates for a hornwort experimental system. Selection of these two species was based on the fact that they are the species that appear most frequently in the literature, they can easily be propagated vegetatively in lab conditions, and they have small genome sizes of approximately 180 Mbp (Bainard and Villarreal, 2013).

Mature gametophytes and sporophytes of *A. punctatus* were originally collected from a glasshouse at the Royal Botanic Garden of Edinburgh, whereas mature gametophytes and sporophytes of *A. agrestis* were collected near Fogo, Berwickshire, by Dr David Long. Several sterilization protocols were tested on sporophytes, spores and thallus tissue (Table 5.1). Bleach treatment of spores successfully eliminated growth of bacterial and fungal contamination. Sterile spores were plated on Lorbeer's medium supplemented with Hutner's trace elements solution. Plates were incubated under controlled environment conditions at 23°C under long day photoperiod (16 h

light : 8 h dark, 3200 lux light intensity). Spore germination occurred successfully after approximately 7-10 days for both *A. punctatus* and *A. agrestis*.

5.3.2 Optimisation of growth conditions

Young thalloid gametophytes were transferred onto fresh Lorbeer's medium one month after germination. Axenic cultures of gametophytes were then established and were routinely propagated by sub-culturing on a monthly basis. Four different growth media were tested: Lorbeer's medium, KNOP medium, BCD medium and MS medium (see Materials and Methods). Thallus growth was fastest on Lorbeer's medium compared to the other three media, and so was adopted. Gametophyte cultures were grown on Lorbeer's medium at 23°C under long day photoperiod (16 h light : 8 h dark, 3200 lux light intensity). For sub-culturing, a small piece of thallus tissue was isolated by cutting and then placed on fresh medium.

5.3.3 Establishment of sporophyte induction conditions

To investigate the function of *KNOX* genes in hornworts, sporophyte induction in lab conditions had to be achieved. *KNOX* gene expression and function are sporophyte-specific in all other land plants examined to date, and several attempts to amplify *KNOX* genes from *A. punctatus* gametophyte cDNA were unsuccessful. Under the growth conditions established for gametophyte culture, sporophytes did not develop.

Sporophytes arise from the fertilization of egg cells within specialized archegonia by sperm produced from antheridia. Both of these structures develop on the surface of the gametophyte thallus. In their natural environment hornwort thalli develop antheridia first, and after approximately one month archegonia develop (Poskauer, 1967). In an attempt to induce these developmental changes under lab conditions, a series of different media, temperatures, photoperiods and light intensity conditions

were tested for *A. punctatus* and *A. agrestis* (Table 5.2A). Sporophytes were successfully induced for *A. agrestis* only under conditions of 16°C temperature, long day photoperiod (16 h light : 8 h dark, 2500 lux light intensity). This indicates that antheridia induction is triggered by a period of low temperature, similar to the moss *P. patens*. 1/10 KNOP medium was optimal for *A. agrestis* sporophyte induction (Table 5.2B). The number of *A. agrestis* sporophytes produced per culture was between 50 to 60 in comparison to about 10 sporophytes produced when Lorbeer's medium was used. Under the same inductive conditions only antheridia formation was achieved for *A. punctatus* and thus no sporophytes were obtained.

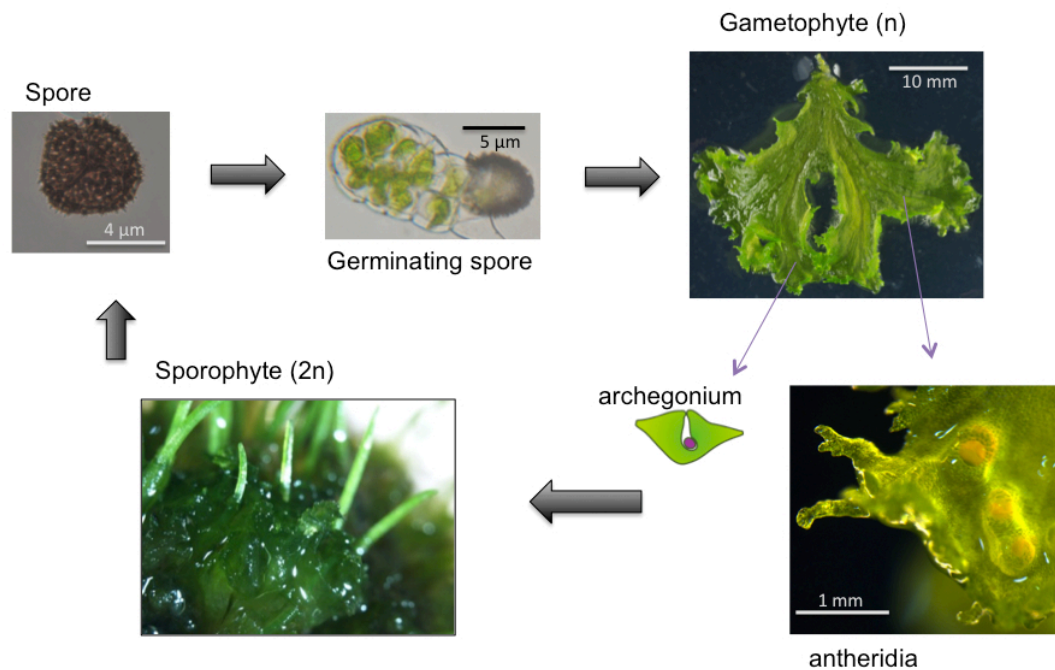


Figure 5.1

Life cycle of the hornwort *Anthoceros punctatus*. The thallus is irregularly shaped, lacks specialized internal differentiation and is composed of mucilage chambers and parenchyma cells. Male (antheridia) and female (archegonia) reproductive organs are embedded in the thallus. The sporophyte has stomata, a foot, columella, a spore layer, multicellular jacket and elaters (for spore dispersal). n: haploid, 2n: diploid. (sporophyte photo: Dr Keiko Sakakibara).

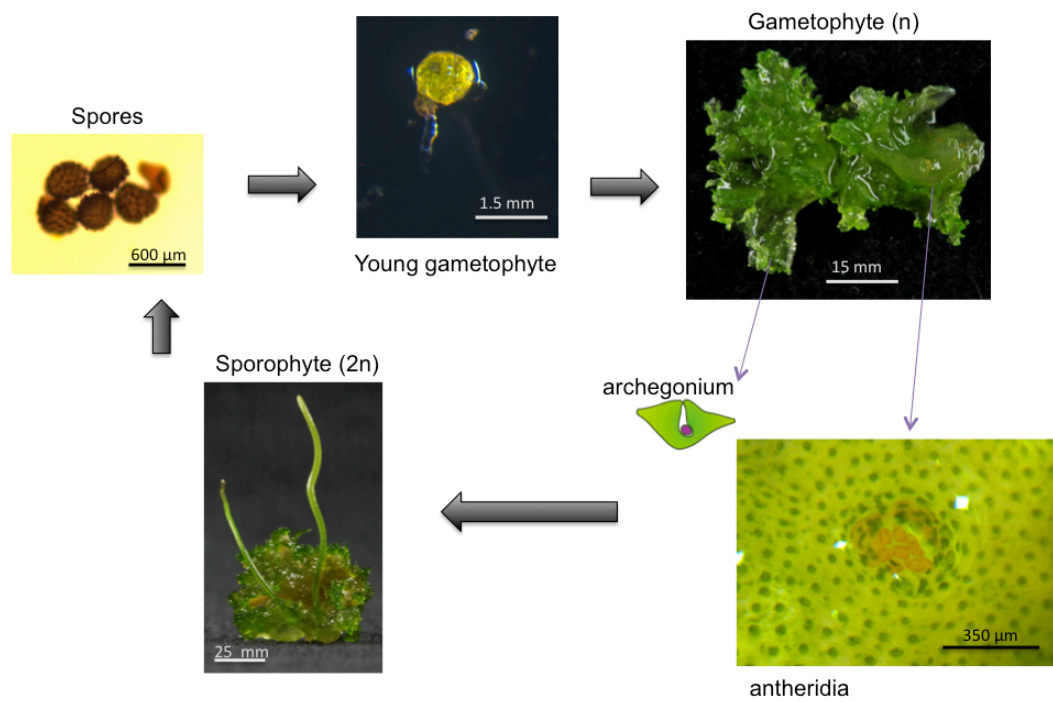


Figure 5.2

Life cycle of the hornwort *Anthoceros agrestis*. The morphology and developmental stages are similar to *A. punctatus*.

Plant material tested	Number of sporophytes/spores	Step 1: Bleach (+0.1% Triton-X) - 3min	Step 2: EtOH - 3min	Step 3: Number of washes with water	Successful sterilisation achieved
Whole sporophyte	5	5%	70%	3	yes
	5	10%	70%	3	yes
	5	20%	70%	3	yes
	5	5%	-	3	yes
	5	10%	-	3	yes
Spores	100-150	0.50%	-	3	yes
	100-150	5%	-	3	no
	100-150	10%	-	3	no
Thallus	-	0.10%	-	3	yes
	-	0.50%	-	3	yes

Table 5.1

Optimization of sterilization methods prior to the culture of sporophyte, spore and thallus tissues. Sporophytes, spores or thallus were first treated with bleach, followed by a treatment with 70% EtOH only in the case of sporophytes, and finally rinsed with water to remove cleaning agents.

A Antheridia induction				
	16°C, 16 h light : 8 h dark, 2500 lux	21°C, 16 h light : 8 h dark, 3200 lux	25°C, 16 h light : 8 h dark, 3000 lux	16°C, 24 h light : 22°C 24 h dark, 2500 lux
<i>A. punctatus</i>	yes	no	no	no
<i>A. agrestis</i>	yes	not tested	not tested	not tested

B Sporophyte induction					
	Lorbeer's	1/10 Lorbeer's without trace elements	1/10 Knop	BCD without trace elements	peat pellets
<i>A. punctatus</i>	no	no	no	no	no
<i>A. agrestis</i>	yes	yes	yes	no	no

Table 5.2

A. Different conditions tested for antheridia induction.

B. Different media tested for sporophyte induction.

5.3.4 Development of nucleic acid extraction methods

The gametophytic tissue of hornworts is rich in polysaccharides (Renzaglia et al., 2009) and was also found to be rich in polyphenolics. Both compounds pose a problem for DNA and RNA extraction. A range of DNA and RNA extraction protocols were therefore tested to optimize the procedure and to reduce contamination levels as much as possible (Table 5.3). A modified CTAB protocol was found to be optimal for hornwort genomic DNA extraction (adapted from Porebski et al. 1997). This protocol uses polyvinyl pyrrolidone to remove polyphenolics and contains an extra ethanol precipitation step with a relatively high NaCl concentration compared standard DNA extraction protocols. At NaCl concentrations higher than 0.5 M, polysaccharides remain in solution and they do not co-precipitate with DNA. The same protocol was used for RNA extraction with the addition of an overnight RNA precipitation step with LiCl (See Materials and Methods).

DNA extraction method	Contamination (polysaccharides and polyphenolics)
CTAB based	yes
Urea based	yes
CTAB method - Forrest et al., 2011	yes
Guanidine based	yes
Modified CTAB method - Porebski et al. 1997	no

Table 5.3

Optimization of DNA extraction protocols for hornwort genomic DNA extraction.

5.3.5 Identification of a single *KNOX* gene in *A. punctatus*

To identify candidate hornwort *KNOX* genes, a *de novo* sequence assembly of the *A. punctatus* genome was carried out. Genomic DNA from gametophytic tissue was isolated in Oxford and sent to the Beijing Genomics Institute (China) for sequencing. DNA was sheared into ~400 bp fragments, and sequenced using Illumina HiSeq2000. A total of 25 million 90 bp paired-end reads (about 20X genome coverage) were generated. The reads were assembled by Dr Steven Kelly. The final assembly contained 29,582 contigs with lengths ranging from 919 bp to 76.5 kb, with 1643 contigs over 10 kb.

This draft genome was screened for homologs of *KNOX* genes using translated BLAST searches. Both class 1 and class 2 *KNOX* genes from *S. krausianna* and *P. patens* were used for BLAST searches (Altschul et al., 1997). A single *KNOX* gene was identified in *A. punctatus* (hereafter called *ApKNOX*). On the basis of sequence similarity, this gene is more closely related to class 2 than class 1 *KNOX* genes (Figure 5. 3).

5.3.6 Phylogenetic analysis of *ApKNOX* gene in land plants

To further confirm that the identified *KNOX* gene was a class 2 gene, a phylogenetic analysis was carried out to investigate the relationship between the *ApKNOX* gene and *KNOX* homologues in other land plants. The class 1 and class 2 *KNOX* genes that were included in the analysis are summarized in Table 5.3 and Table 5.4.

Gene	GenBank accession/reference	Species	Taxon
<i>STM</i>	AAC49148.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>BP/KNAT1</i>	NP_192555.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT2</i>	NP_177208.2	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT6</i>	BAB69679.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>OSH1</i>	JQ2379	<i>Oryza sativa</i>	Angiosperm
<i>OSH15</i>	NP_911093	<i>Oryza sativa</i>	Angiosperm
<i>Os_rough_sheath1</i>	BAA79225	<i>Oryza sativa</i>	Angiosperm
<i>OSH10</i>	XP_469241	<i>Oryza sativa</i>	Angiosperm
<i>OSH3</i>	BAA79223	<i>Oryza sativa</i>	Angiosperm
<i>OSH6</i>	NP_914102	<i>Oryza sativa</i>	Angiosperm
<i>OSJNBa0035I24.10</i>	AAR87205	<i>Oryza sativa</i>	Angiosperm
<i>OSH71</i>	BAA79226	<i>Oryza sativa</i>	Angiosperm
<i>lg3</i>	AAD13611	<i>Zea mays</i>	Angiosperm
<i>RS1</i>	Q41853	<i>Zea mays</i>	Angiosperm
<i>GN1</i>	AAP76320	<i>Zea mays</i>	Angiosperm
<i>PUEIN13TD</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>kn1</i>	AAP76321	<i>Zea mays</i>	Angiosperm
<i>lg4a</i>	AAP31409	<i>Zea mays</i>	Angiosperm
<i>AZM432736</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM4113483</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM4113484</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>PtKN1</i>	AAV54618	<i>Pinus taeda</i>	Gymnosperm
<i>PtKn2</i>	AY680403	<i>Pinus taeda</i>	Gymnosperm
<i>PtKN3</i>	AAV54620.1	<i>Pinus taeda</i>	Gymnosperm
<i>PtKN4</i>	AAV63998.1	<i>Pinus taeda</i>	Gymnosperm
<i>SKN1</i>	AAD00691.1	<i>Picea mariana</i>	Gymnosperm
<i>PmSKn2</i>	AAD00692.1	<i>Picea mariana</i>	Gymnosperm
<i>PmKN4</i>	AAV54621.1	<i>Picea mariana</i>	Gymnosperm
<i>HBK1</i>	AAC84001.1	<i>Picea abies</i>	Gymnosperm
<i>HBK2</i>	AAN77690.1	<i>Picea abies</i>	Gymnosperm
<i>HBK3</i>	AAN77691.1	<i>Picea abies</i>	Gymnosperm
<i>CrKNOX1</i>	BAB18582.1	<i>Ceratopteris richardii</i>	Monilophyte
<i>CrKNOX2</i>	BAB18584.1	<i>Ceratopteris richardii</i>	Monilophyte
<i>SkKNOX1</i>	AAW62517.1	<i>Selaginella kraussiana</i>	Lycophyte
<i>SkKNOX2</i>	AAW62518.1	<i>Selaginella kraussiana</i>	Lycophyte
<i>MKN2</i>	XP_001758592.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MKN4</i>	XP_001781477.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MKN5</i>	XP_001778265.1	<i>Physcomitrella patens</i>	Bryophyte - Moss

Table 5.4

List of class 1 KNOX genes used in phylogenetic analysis

Gene	GenBank accession/reference	Species	Taxon
<i>KNAT3</i>	AAM53320.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT4</i>	CAA63131.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT5</i>	AAL37042.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT7</i>	AAG40858.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>OSH45</i>	Q0J6N4.2	<i>Oryza sativa</i>	Angiosperm
<i>HOS59</i>	BAA77822.2	<i>Oryza sativa</i>	Angiosperm
<i>HOS66</i>	BAB55660	<i>Oryza sativa</i>	Angiosperm
<i>HOS58</i>	BAB55658	<i>Oryza sativa</i>	Angiosperm
<i>AZM440326</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM451939</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM459391</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM448186</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>PtKNAT4</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PtKNAT3</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt672778</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt738080</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PtKNAT3b</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PpKNAT7</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>NTH23</i>	BAA25921.1	<i>Nicotiana tabacum</i>	Angiosperm
<i>Let12</i>	NP_001233945.1	<i>Solanum lycopersicum</i>	Angiosperm
<i>KNOPE3</i>	ABH03531.1	<i>Prunus persica</i>	Angiosperm
<i>KNOPE4</i>	ABO26062.1	<i>Prunus persica</i>	Angiosperm
<i>KNOPE7</i>	AFF27503.1	<i>Prunus persica</i>	Angiosperm
<i>CrKNOX3</i>	BAB18585.1	<i>Ceratopteris richardii</i>	Monilophyte
<i>SkKNOX3</i>	AAW62519.1	<i>Selaginella kraussiana</i>	Lycophyte
<i>ApKNOX</i>	--	<i>Anthoceros punctatus</i>	Bryophyte - Hornwort
<i>MKN1</i>	AAK61309.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MKN6</i>	XP_001765575.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MpKNOX</i>	--	<i>Marchantia polymorpha</i>	Bryophyte - Liverwort
<i>GSM1</i>	ABS71849.1	<i>Chlamydomonas reinhardtii</i>	Chlorophyte

Table 5.5

List of Class 2 *KNOX* genes used in phylogenetic analysis.

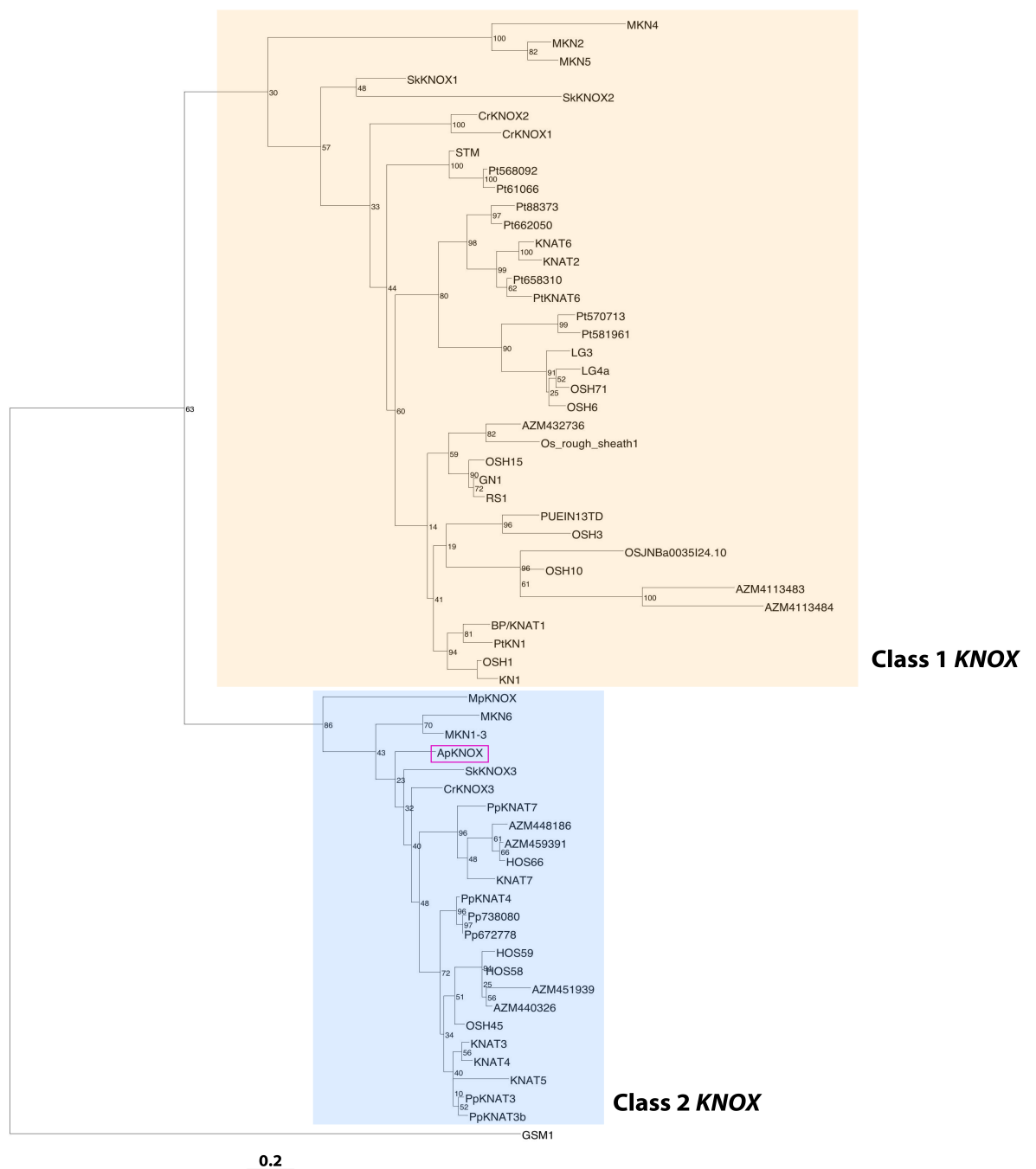


Figure 5.4

ApKNOX gene phylogeny. Class 1 *KNOX* genes are highlighted with an orange box, class 2 *KNOX* genes with a blue box and *ApKNOX* with a purple box. Numbers on the branches represent bootstrap support values. Scale bar: mean number of nucleotide substitutions per site.

5.4 Discussion

As the proposed closest extant sister group to vascular plants, the hornwort clade resides in an important phylogenetic position with regards to our understanding of the genetic basis of sporophyte evolution. The development of a hornwort experimental system will therefore greatly facilitate efforts towards a more comprehensive study of the mechanisms underpinning land plant evolution. Of the two hornwort species tested, *A. agrestis* was found to be more suitable than *A. punctatus* because its entire life cycle can be reproduced in lab conditions. Future research efforts should therefore be focussed on this species.

The isolation of only one *KNOX* gene in the hornwort *A. punctatus* is an intriguing finding given that earlier diverging bryophytes like *P. patens* carry multiple *KNOX* genes. It is believed that a whole genome duplication has occurred in *P. patens* (Rensing et al., 2008). However, the number of *P. patens* *KNOX* genes is higher than two, suggesting that the ancestral land plant condition was to possess more than one *KNOX* gene. Interestingly, the identified *A. punctatus* *KNOX* gene is most similar to class 2 *KNOX* genes. One possibility, given that the genome assembly is a rough draft, is that other *KNOX* genes are present but have not been detected. Another possible scenario is that the absence of class 1 *KNOX* genes in *A. punctatus* represents a secondary loss of class 1 *KNOX* genes within hornworts, or at least within the genus *Anthoceros*. Alternatively, it is possible that the current phylogenetic position of hornworts within the bryophytes as sister to the vascular plants is wrong. Bryophyte phylogenetic relationships are the subject of a long-standing debate. It has frequently been proposed that hornworts are the most basal group of bryophytes or sister to all embryophytes except liverworts (Mishler et al. 1994; Qiu et al. 1998; Goffinet 2000; Renzaglia and Vaughn 2000; Nickrent et al., 2000) and studies are still

ongoing to validate this suggestion (Villarreal, personal communication). If this is the case, then the absence of class 1 *KNOX* genes in *A. punctatus* can be explained more parsimoniously, given that green algae have a single *KNOX* gene, which in the case of *C. reinhardtii* is more similar to class 2 *KNOX* genes (Lee et al., 2008).

6 General discussion

6.1 Gene family origins and expansions during land plant evolution

Genome sequences of the chlorophyte *C. reinhardtii* (Merchant et al., 2007), the early divergent land plant *P. patens* (Rensing et al., 2008) and the vascular plant *S. moellendorffii* (Banks et al., 2011) have provided valuable insight into the origins and expansion of gene families during land plant evolution. It is estimated that 3814 gene families were present in the common ancestor of plants, a number that almost doubled after the colonization of land to 6820, and further increased to 8155 in angiosperms (Banks et al., 2011). Despite this significant increase, most of the approximately 50-60 transcription factor families found in angiosperms are also present in early divergent land plants, and 15-30 are present in *C. reinhardtii* (Riano-Pachon et al., 2008). However, the average number of genes in each family increased from mosses to angiosperms (Richardt et al., 2007; Riano-Pachon et al., 2008). This expansion and diversification of transcription factor families during land plant evolution was coincident with the evolution of morphological and biochemical complexity.

6.2 The role of *KNOX* genes in the evolution of the SAM

6.2.1 *KNOX* gene function in bryophytes

One of the most important transcription factor families with a key role in plant development is the *KNOX* gene family. In the chlorophyte *C. reinhardtii* a single *KNOX* gene is present that is essential for zygote development. Upon colonization of the land, *KNOX* genes diversified into two subclasses, both of which function specifically in the diploid phase of the life cycle as in *C. reinhardtii*. Unlike green algae, however,

the sporophyte in land plants is multicellular. A delay in zygotic meiosis and the intercalation of mitotic divisions has been proposed as a mechanism for the evolution of a multicellular sporophyte. Additional mechanisms to sustain mitotic divisions after zygote formation and to repress the gametophytic program during sporophyte development also had to be acquired. Notably, in the moss *P. patens*, class 1 *KNOX* genes function to control meristematic activity in the sporophyte (Sakakibara et al., 2008) whereas class 2 *KNOX* genes suppress the gametophytic developmental program (Sakakibara et al., 2013).

6.2.2 Evolution of the SAM in vascular plants

As in *P. patens*, class 1 *KNOX* function in vascular plants is meristem specific. However, in contrast to bryophytes, meristematic activity in vascular plant sporophytes is indeterminate. The question arises therefore as to how vascular plant meristems evolved from bryophyte meristems and to what extent modification of *KNOX* gene function facilitated any changes that occurred. Two main theories have been proposed to explain how the vascular plant SAM evolved:

- One theory proposes that the vascular plant SAM originated by apical displacement of a transient basal meristem similar to that found in mosses (Mishler and Churchill, 1984) or of the basal hornwort meristem (Ligrone et al., 2012).
- A second theory suggests that the vascular plant SAM originated from an existing apical meristem such as that observed in mosses (Albert, 1999).

Given that class 1 *KNOX* genes are expressed in the seta and transient apical meristem of the moss sporophyte (Sakakibara et al., 2008) it could be hypothesized

that the indeterminate vascular SAM evolved through prolonged expression of class 1 *KNOX* genes in the apical meristem (Bowman, 2013). An objection to the above scenario is that in the draft *A. punctatus* genome assembly, class 1 *KNOX* genes were absent (Chapter 5). However, given that apical meristems are also absent in hornworts both could represent secondary losses within hornworts.

6.3 Divergent class 1 *KNOX* gene function in bryophytes and vascular plants

As indicated by the preliminary results of the cross-species complementation experiments in moss (Chapter 4), *KNOX* function seems to have diverged between bryophytes and vascular plants such that *SkKNOX* and *CrKNOX* genes are unable to complement *mkn2* mutants. This suggests that *KNOX* function became specialized during the evolution of vascular plants, possibly through differential activity on downstream targets.

Similar specialization of function has been reported for other genes. An example is *GOLDEN2-LIKE* (*GLK*) transcription factors that have a key role in chloroplast development and exist as homologous pairs in all land plants examined to date. *Arabidopsis* and *P. patens glk1;glk2* double mutants exhibit reduced chlorophyll biosynthesis and have impaired chloroplast development. However, *Arabidopsis GLK1* cannot rescue the *P. patens glk1;glk2* double mutant phenotype and *P. patens GLK1* can only partially rescue the *Arabidopsis glk1;glk2* double mutant phenotype. (Yasumura et al., 2007; Bravo-Garcia et al., 2009).

Another example is seen with *ABSCISIC ACID INSENSITIVE 3* (*ABI3*), which confers desiccation tolerance in *P. patens* and *Arabidopsis*. *P. patens* homologues of *ABI3*,

can only partially rescue *Arabidopsis abi3* mutant phenotypes (Khandelwal et al., 2010).

6.4 Conserved class 1 *KNOX* gene function in lycophytes and angiosperms.

Based on the potential of *SkKNOX* genes to complement *Arabidopsis bp* mutants when expressed under the control of the *BP* promoter (Chapter 3), *KNOX* gene function appears to have been conserved since lycophytes and angiosperms diverged. However, it is not yet clear whether *SkKNOX1* and *SkKNOX2* have equivalent potential to complement the *bp* mutant phenotype because more independent lines need to be examined. The similarity between *BP* and *SkKNOX2* expression patterns, however, suggests that *SkKNOX2* function rather than *SkKNOX1* function should be more similar to *BP*.

The finding that *SkKNOX* genes can complement *Arabidopsis bp* mutants is in accordance with the suggestion that a conserved *KNOX/ARP* mechanism controls leaf development in *S. kraussiana* and angiosperms (Harrison et al., 2005). Such conservation indicates that despite the independent origins of microphylls and seed plant megaphylls, the recruitment of *KNOX* and *ARP* genes to new functions may have been constrained by their ancestral function in lycophyte meristems. In contrast to *KNOX* genes, class III *HD-Zip* transcription factors control the adaxial/abaxial patterning of leaves in seed plants but not in lycophytes, providing an example of an innovation that reflects the independent origin of microphylls and megaphylls (Floyd and Bowman, 2006).

6.5 Mechanisms for the specialization of *KNOX* gene function during land plant evolution

The evolution of specialized *KNOX* gene function could have been mediated through, gene duplication, changes in regulatory element sequences, changes in KNOX/BELL protein interactions, or sequence changes in the open reading frame.

6.5.1 *KNOX* gene duplication

A series of gene duplications gave rise to new *KNOX* gene copies (Mukherjee et al., 2009) something that can facilitate either the acquisition of a new function (neofunctionalization) or sub-functionalization of the ancestral function. For example, in *Arabidopsis* the four Class 1 *KNOX* genes, *STM*, *BP*, *KNAT2* and *KNAT6* acquired distinct but also redundant roles in SAM maintenance and organ separation (Byrne et al., 2002; Gallois et al., 2002; Belles-Boix et al., 2006). Redundancy ensures developmental robustness by facilitating compensatory interactions (Ragni et al., 2007).

6.5.2 *KNOX* regulatory element changes

An example of how changes in regulatory sequences of *KNOX* genes led to the evolution of different morphologies comes from *Cardamine hirsuta*. In contrast to *Arabidopsis*, where exclusion of KNOX function from leaves results in a simple leaf form, in *C. hirsuta* KNOX proteins are required in the leaf to delay cellular differentiation and to produce a dissected leaf form. These differences in *KNOX* expression arise through changes in the activity of upstream gene regulatory elements (Hay and Tsiantis, 2006).

A second example is the *BELL* gene *REPLUMLESS* (*RPL*, also known as *PENNYWISE* and *BELLRINGER*, (Byrne et al., 2002; Smith and Hake, 2003; Bhatt et al., 2004). Together with *BP*, *RPL* is a key regulator of replum development, an important structure involved in fruit opening to release seeds. In *Brassica rapa* the replum is reduced, resembling the *Arabidopsis* *rpl* mutant. It has been shown that a point mutation in a conserved *cis*-regulatory element of the *B. rapa* *RPL* homologue accounts for the loss of *RPL* expression in *B. rapa* fruit and thus the reduced replum size (Arnaud et al., 2011). Interestingly, in domesticated rice, a similar mutation in the same *cis*-regulatory element of an *RPL* homologue accounts for reduced seed dispersal potential (Konishi et al., 2006).

6.5.3 Evolution of novel KNOX/BELL interactions

Altered KNOX/BELL interactions may be one way of generating functional diversity without affecting protein structure. The KNOX DNA binding domain may stay intact, but the MEINOX domain may change to interact with a different BELL protein and a series of new functions can emerge.

A specialized KNOX/BELL interaction has been described recently that involves KNATM proteins (KNOX proteins that lack a homeodomain) (Magnani and Hake, 2008). KNATM proteins interact with BELL proteins and affect their availability to form active DNA-binding complexes. For example, in *Arabidopsis* KNATM may prevent the BELL proteins SAWTOOTH1 (SAW1) and SAW2 from forming transcriptional activation complexes by titrating them in an inactive dimer (Kumar et al., 2007; Magnani and Hake, 2008).

6.5.4 Sequence changes in open reading frames

There are numerous examples of changes in protein structure that led to the evolution of new functions. For example, in angiosperms, *LEAFY* regulates the transition from the vegetative phase to the reproductive phase by regulating the transcription of the *ABC* floral organ identity genes. Loss-of-function *lfy* mutants cause a partial loss of floral meristem identity and gain-of-function mutants cause inflorescence meristems to differentiate into floral meristems (Moyroud et al., 2010). The *P. patens* genome contains two *LFY* orthologs (*PpLFY1* and *PpLFY2*) that are required for the first cell division in the zygote (Tanahashi et al., 2005). Notably, *PpLFY* genes do not complement *LFY* mutants in *Arabidopsis* (Tanahashi et al., 2005), with *LFY* function in angiosperms having diverged from that in bryophytes by substitutions in the DNA binding domain (Maizel et al., 2005).

Another example is seen with the *FT* (*FLOWERING LOCUS T*) and *TFL* (*TERMINAL FLOWER*) regulators of flowering in *Arabidopsis*. *FT* expression induces flowering whereas *TFL* is a floral repressor that maintains indeterminate growth of the shoot apical meristem. The opposing functions of *FT* and *TFL* have been traced to a single amino acid difference in the phosphatidylethanolamine binding sequence of the protein (Hanzawa et al., 2005). A similar single amino acid difference underpins differential activity of the MADS-box transcription factors, *AGAMOUS* (*AG*), *PLENA* (*PLE*) and *FARINELLI* (*FAR*). In *Arabidopsis*, *AG* controls both male and female organ development. In contrast, in *Antirrhinum*, *PLE* contributes to the specification of both male and female organs whereas *FAR* contributes only to the specification of male organs. A single amino acid present in *FAR* limits protein function to the induction of male organs and its absence in both *PLE* and *AG*, correlates with the induction of both male and female organ development (Airoldi et al., 2010).

6.6 Future perspectives

The work presented in this thesis contributes to a growing body of evidence that suggests that the function of class 1 *KNOX* genes became more specialized as land plants evolved.

Given that *KNOX* protein function appears to be conserved between *S. kraussiana* and *Arabidopsis*, the next question to be addressed would be to what extent spatial differences in expression patterns in the ‘simple’ lycophyte versus multilayered angiosperm meristem can be attributed to *cis*-regulatory element changes. Promoter swap experiments could be performed to address this question. Given that methods for the genetic manipulation of *S. kraussiana* are not yet available, these experiments have to be performed in *Arabidopsis*. More specifically, the 5' regions of *SkKNOX1* and *SkNOX2* could be fused with *BP* and *STM* coding sequences to determine whether they can complement *Arabidopsis bp* and *stm* mutants. A fusion between *SkKNOX1* and *SkNOX2* promoters and the β -glucoronidase (*GUS*) reporter gene could also be used to test whether *SkKNOX* promoters can generate similar expression patterns in *Arabidopsis* to those seen with the native class 1 *KNOX* genes.

The recent development of stable transformation techniques for the fern *C. richardii* will significantly enhance our ability to understand *KNOX* function in ferns (Muthukumar et al., 2013; Plackett et al., 2014). Suppression of *CrKNOX1* and *CrKNOX2* expression by RNAi silencing will reveal which aspects of *C. richardii* sporophyte development are controlled by class 1 *KNOX* genes and overexpression will provide evidence about the functional potential of class 1 *KNOX* genes in ferns. The study of the single class 2 *KNOX* gene (*CrKNOX3*) identified in *C. richardii*, will further provide valuable insight into how the role of class 2 *KNOX* genes diversified during vascular plant evolution.

Finally, the establishment of sporophyte induction conditions for the hornwort *A. agrestis* will significantly facilitate our efforts towards understanding *KNOX* function in early divergent land plants. *A. agrestis KNOX* gene expression patterns should first be characterized by *in situ* hybridization to determine whether transcript accumulation is seen in the basal meristem. Ectopic expression of *A. agrestis KNOX* genes under the control of the constitutive *CaMV* 35S promoter in *Arabidopsis* would inform our understanding about the functional potential of hornwort *KNOX* genes and cross-species complementation experiments in *P. patens* will shed light on the extent of functional conservation of *KNOX* genes between bryophyte groups.

References

- Abascal, F., Zardoya, R., and Posada, D.** (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatic* **21**, 2104-2105.
- Aida, M., Ishida, T., and Tasaka, M.** (1999). Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. *Development* **126**, 1563-1570.
- Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M.** (1997). Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. *The Plant Cell* **9**, 841-857.
- Airoidi, C.A., Bergonzi, S., and Davies, B.** (2010). Single amino acid change alters the ability to specify male or female organ identity. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18898-18902.
- Albert, V.A.** (1999). Comment - Shoot apical meristems and floral patterning: an evolutionary perspective. *Trends in Plant Science* **4**, 84-86.
- Alonso-Cantabrana, H., Ripoll, J.J., Ochando, I., Vera, A., Ferrandiz, C., and Martinez-Laborda, A.** (2007). Common regulatory networks in leaf and fruit patterning revealed by mutations in the Arabidopsis ASYMMETRIC LEAVES1 gene. *Development* **134**, 2663-2671.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J.** (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.
- Arnaud, N., Lawrenson, T., Ostergaard, L., and Sablowski, R.** (2011). The same regulatory point mutation changed seed-dispersal structures in evolution and domestication. *Current Biology* **21**, 1215-1219.
- Bainard, J.D., and Villarreal, J.C.** (2013). Genome size increases in recently diverged hornwort clades. *Genome* **56**, 431-435.
- Bellaoui, M., Pidkowich, M. S., Samach, A., Kushalappa, K., Kohalmi, S. E., Modrusan, Z., Crosby W. L., and Haughn, G. W.** (2001). The Arabidopsis BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. *Plant Cell* **13**, 2455-2470.
- Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., dePamphilis, C., Albert, V.A., Aono, N., Aoyama, T., Ambrose, B.A., Ashton, N.W., Axtell, M.J., Barker, E., Barker, M.S., Bennetzen, J.L., Bonawitz, N.D., Chapple, C., Cheng, C., Correa, L.G., Dacre, M., DeBarry, J., Dreyer, I., Elias, M., Engstrom, E.M., Estelle, M., Feng, L., Finet, C., Floyd, S.K., Frommer, W.B., Fujita, T., Gramzow, L., Gutensohn, M., Harholt, J., Hattori, M., Heyl, A., Hirai, T., Hiwatashi, Y., Ishikawa, M., Iwata, M., Karol, K.G., Koehler, B., Kolukisaoglu, U., Kubo, M., Kurata, T., Lalonde, S., Li, K., Li, Y., Litt, A., Lyons, E., Manning, G., Maruyama, T., Michael, T.P., Mikami, K., Miyazaki, S., Morinaga, S., Murata, T.,**

Mueller-Roeber, B., Nelson, D.R., Obara, M., Oguri, Y., Olmstead, R.G., Onodera, N., Petersen, B.L., Pils, B., Prigge, M., Rensing, S.A., Riano-Pachon, D.M., Roberts, A.W., Sato, Y., Scheller, H.V., Schulz, B., Schulz, C., Shakhov, E.V., Shibagaki, N., Shinohara, N., Shippen, D.E., Sorensen, I., Sotooka, R., Sugimoto, N., Sugita, M., Sumikawa, N., Tanurdzic, M., Theissen, G., Ulvskov, P., Wakazuki, S., Weng, J.K., Willats, W.W., Wipf, D., Wolf, P.G., Yang, L., Zimmer, A.D., Zhu, Q., Mitros, T., Hellsten, U., Loque, D., Otilar, R., Salamov, A., Schmutz, J., Shapiro, H., Lindquist, E., Lucas, S., Rokhsar, D., and Grigoriev, I.V. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**, 960-963.

Barton, M.K. (2010). Twenty years on: the inner workings of the shoot apical meristem, a developmental dynamo. *Developmental Biology* **341**, 95-113.

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

Becker, B., and Marin, B. (2009). Streptophyte algae and the origin of embryophytes. *Annals of Botany* **103**, 999-1004.

Belles-Boix, E., Hamant, O., Witiak, S.M., Morin, H., Traas, J., and Pautot, V. (2006). KNAT6: an *Arabidopsis* homeobox gene involved in meristem activity and organ separation. *Plant Cell* **18**, 1900-1907.

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

Bennici, A. (2008). Origin and early evolution of land plants: Problems and considerations. *Communicative and Integrative Biology* **1**, 212-218.

Bertolino, E., Reimund, B., Wildt-Perinic, D., and Clerc, R.G. (1995). A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. *Journal of Biological Chemistry* **270**, 31178-31188.

Bharathan, G., Janssen, B.J., Kellogg, E.A., and Sinha, N. (1999). Phylogenetic relationships and evolution of the KNOTTED class of plant homeodomain proteins. *Molecular Biology and Evolution* **16**, 553-563.

Bharathan, G., Goliber, T.E., Moore, C., Kessler, S., Pham, T., and Sinha, N.R. (2002). Homologies in leaf form inferred from KNOX1 gene expression during development. *Science* **296**, 1858-1860.

Bhatt, A.M., Etchells, J.P., Canales, C., Lagodienko, A., and Dickinson, H. (2004). VAAMANA--a BEL1-like homeodomain protein, interacts with KNOX proteins BP and STM and regulates inflorescence stem growth in *Arabidopsis*. *Gene* **328**, 103-111.

Billeter, M., Qian, Y. Q., Otting, G., Muller, M., Gehring, W., and Wuthrich, K. (1993). Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *Journal of Molecular Biology* **234**, 1084-1093.

Bolduc, N., and Hake, S. (2009). The maize transcription factor KNOTTED1 directly regulates the gibberellin catabolism gene *ga2ox1*. *The Plant Cell* **21**, 1647-1658.

Bowman, J.L., Floyd, S.K., and Sakakibara, K. (2007). Green genes - Comparative genomics of the green branch of life. *Cell* **129**, 229-234.

Bowman, J.L. (2013). Walkabout on the long branches of plant evolution. *Current Opinion in Plant Biology* **16**, 70-77.

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

Bravo-Garcia, A., Yasumura, Y., and Langdale, J.A. (2009). Specialization of the Golden2-like regulatory pathway during land plant evolution. *New Phytologist* **183**, 133-141.

Burglin, T.R. (1997). Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Research* **25**, 4173-4180.

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

Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973.

Carroll, S.B. (2005). Evolution at two levels: on genes and form. *PLoS Biology* **3**, e245.

Champagne, C.E.M., and Ashton, N.W. (2001). Ancestry of KNOX genes revealed by bryophyte (*Physcomitrella patens*) homologs. *New Phytologist* **150**, 23-36.

Clark, S.E., Williams, R.W., and Meyerowitz, E.M. (1997). The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell* **89**, 575-585.

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

Cole, M., Nolte, C. and Werr, W. (2006). Nuclear import of the transcription factor SHOOT MERISTEMLESS depends on heterodimerization with BLH proteins expressed in discrete sub-domains of the shoot apical meristem of Arabidopsis thaliana. *Nucleic Acids Research* **34**, 1281-1292.

Collingridge, P.W. and Kelly, S. (2012) MergeAlign: improving multiple sequence alignment performance by dynamic reconstruction of consensus multiple sequence alignments. *BMC Bioinformatics* **13**, 117.

Crane, P.R., Herendeen, P., and Friis, E.M. (2004). Fossils and plant phylogeny. *American Journal of Botany* **91**, 1683-1699.

Derelle, R., Lopez, P., Le Guyader, H., and Manuel, M. (2007). Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes. *Evolution and Development* **9**, 212-219.

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

Edwards, D., and Feehan, J. (1980). Record of Cooksonia-type sporangia from late Wenlock strata in Ireland. *Nature* **287**, 41-42.

Endrizzi, K., Moussian, B., Haecker, A., Levin, J.Z., and Laux, T. (1996). The SHOOT MERISTEMLESS gene is required for maintenance of undifferentiated cells in Arabidopsis shoot and floral meristems and acts at a different regulatory level than the meristem genes WUSCHEL and ZWILLE. *Plant Journal* **10**, 967-979.

Engel, P.P. (1968) The induction of biochemical and morphological mutants in the moss, *Physcomitrella patens*. *American Journal of Botany* **55**, 438-446.

Floyd, S.K., and Bowman, J.L. (2006). Distinct developmental mechanisms reflect the independent origins of leaves in vascular plants. *Current Biology* **16**, 1911-1917.

Forrest, L.L., Wickett, N.J., Cox, C.J., and Goffinet, B. (2011). Deep sequencing of *Ptilidium* (Ptilidiaceae) suggests evolutionary stasis in liverwort plastid genome structure. *Plant Ecology and Evolution* **144**, 29-43.

Gallois, J.L., Woodward, C., Reddy, G.V., and Sablowski, R. (2002). Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. *Development* **129**, 3207-3217.

Gensel, P.G. (2008). The Earliest Land Plants. In *Annual Review of Ecology Evolution and Systematics*, 459-477.

Gehring, W. J. (1987). Homeoboxes in the study of development. *Science* **236**, 1245-1252.

Gehring, W. J. (1992). The homeobox in perspective. *Trends in Biochemical Sciences* **17**, 277-280.

Gehring, W. J. (1993). Exploring the homeobox. *Gene* **135**, 215-221.

Ghildiyal, M. and Zamore, P.D. (2009) Small silencing RNAs: an expanding universe. *Nature Reviews Genetics*, **10**, 94-108.

Goffinet, B. (2000). Origin and phylogenetic relationships of bryophytes. In *Bryophyte Biology*. Cambridge: Cambridge University Press. 124-49.

Gleave, A.P. (1992). A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Molecular Biology* **20**, 1203-1207.

Graham, L.E., Cook, M.E., and Busse, J.S. (2000). The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 4535-4540.

Grimsley, N.H., Ashton, N.W., and Cove, D.J. (1977) The production of somatic hybrids by protoplast fusion in the moss *Physcomitrella patens*. *Molecular and General Genetics* **154**, 97-100.

Guillet-Claude, C., Isabel, N., Pelgas, B., and Bousquet, J. (2004). The evolutionary implications of *knox-I* gene duplications in conifers: correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. *Molecular Biology and Evolution* **21**, 2232-2245.

Ha, C.M., Jun, J.H., and Fletcher, J.C. (2010). Shoot apical meristem form and function. *Current Topics in Developmental Biology* **91**, 103-140.

Hackbusch, J., Richter, K., Muller, J., Salamini, F. and Uhrig, J. F. (2005). A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4908-4912.

Haig, D., and Wilczek, A. (2006). Sexual conflict and the alternation of haploid and diploid generations. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* **361**, 335-343.

Hake, S., Vollbrecht, E., and Freeling, M. (1989). Cloning Knotted, the dominant morphological mutant in maize using Ds2 as a transposon tag. *EMBO Journal* **8**, 15-22.

Hake, S., Smith, H.M., Holtan, H., Magnani, E., Mele, G., and Ramirez, J. (2004). The role of *knox* genes in plant development. *Annual Review of Cell and Developmental Biology* **20**, 125-151.

Hamant, O., and Pautot, V. (2010). Plant development: a TALE story. *Comptes Rendus Biologies* **333**, 371-381.

Hanzawa, Y., Money, T., and Bradley, D. (2005). A single amino acid converts a repressor to an activator of flowering. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 7748-7753.

Harrison, C.J., Rezvani, M., and Langdale, J.A. (2007). Growth from two transient apical initials in the meristem of *Selaginella kraussiana*. *Development* **134**, 881-889.

Hay, A., and Tsiantis, M. (2009). A KNOX family TALE. *Current Opinion in Plant Biology* **12**, 593-598.

Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development* **137**, 3153-3165.

Hibara, K., Karim, M.R., Takada, S., Taoka, K., Furutani, M., Aida, M., and Tasaka, M. (2006). *Arabidopsis* CUP-SHAPED COTYLEDON3 regulates postembryonic shoot meristem and organ boundary formation. *Plant Cell* **18**, 2946-2957.

Hjortswang, H.I., Larsson, A.S., Bharathan, G., Bozhkov, P.V., von Arnold, S., and Vahala, T. (2002). KNOTTED1-like homeobox genes of a gymnosperm, Norway spruce, expressed during somatic embryogenesis. *Plant Physiology and Biochemistry* **40**, 837-843.

Huson, D.H., and Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Systematic Biology* **61**, 1061-1067.

Hutner, S.H., Provasoli, L., Schatz, A. and Haskins, C.P. (1950) Some approaches to the study of the role of metals in the metabolism of micro-organisms. *Proceedings of the American Philosophical Society* **94**, 152–170.

Hutvagner, G. and Simard, M.J. (2008) Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology*, **9**, 22-32.

Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., Phillips, A., Hedden, P., and Tsiantis, M. (2005). KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Current Biology* **15**, 1560-1565.

Jeong, S., Trotochaud, A.E., and Clark, S.E. (1999). The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* **11**, 1925-1934.

Kano-Murakami, Y., Yanai, T., Tagiri, A., and Matsuoka, M. (1993). A rice homeotic gene, OSH1, causes unusual phenotypes in transgenic tobacco. *FEBS Letters* **334**, 365-368.

Karol, K.G., McCourt, R.M., Cimino, M.T., and Delwiche, C.F. (2001). The closest living relatives of land plants. *Science* **294**, 2351-2353.

Kenrick, P. (2001). Palaeontology. Turning over a new leaf. *Nature* **410**, 309-310.

Kerstetter, R., Vollbrecht, E., Lowe, B., Veit, B., Yamaguchi, J., and Hake, S. (1994). Sequence analysis and expression patterns divide the maize knotted1-like homeobox genes into two classes. *Plant Cell* **6**, 1877-1887.

Khandelwal, A., Cho, S.H., Marella, H., Sakata, Y., Perroud, P.F., Pan, A., and Quatrano, R.S. (2010). Role of ABA and ABI3 in desiccation tolerance. *Science* **327**, 546.

Kimura, S., Koenig, D., Kang, J., Yoong, F.Y., and Sinha, N. (2008). Natural variation in leaf morphology results from mutation of a novel KNOX gene. *Current Biology* **18**, 672-677.

Konishi, S., Izawa, T., Lin, S.Y., Ebana, K., Fukuta, Y., Sasaki, T., and Yano, M. (2006). An SNP caused loss of seed shattering during rice domestication. *Science* **312**, 1392-1396.

Kumar, R., Kushalappa, K., Godt, D., Pidkowich, M.S., Pastorelli, S., Hepworth, S.R., and Haughn, G.W. (2007). The Arabidopsis BEL1-LIKE HOMEODOMAIN proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. *Plant Cell* **19**, 2719-2735.

Langdale, J.A. (2008). Evolution of developmental mechanisms in plants. *Current Opinion in Genetics and Development* **18**, 368-373.

Langdale, J.A., and Harrison C.J. (2008) *Evolving Pathways: Key Themes in Evolutionary Developmental Biology* Cambridge University Press, Cambridge. 299-315.

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

Lee, J.H., Lin, H., Joo, S., and Goodenough, U. (2008). Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. *Cell* **133**, 829-840.

Leliaert, F., Smith, D.R., Moreau, H., Herron, M.D., Verbruggen, H., Delwiche, C.F., and De Clerck, O. (2012). Phylogeny and Molecular Evolution of the Green Algae. *Critical Reviews in Plant Sciences* **31**, 1-46.

Lewis, L.A., and McCourt, R.M. (2004). Green algae and the origin of land plants. *American Journal of Botany* **91**, 1535-1556.

Ligrone, R., Duckett, J.G., and Renzaglia, K.S. (2012). The origin of the sporophyte shoot in land plants: a bryological perspective. *Annals of Botany* **110**, 935-941.

Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K., and Hake, S. (1994). A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. *Plant Cell* **6**, 1859-1876.

Linkies, A., Graeber, K., Knight, C., and Leubner-Metzger, G. (2010). The evolution of seeds. *New Phytologist* **186**, 817-831.

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

Magnani, E., and Hake, S. (2008). KNOX lost the OX: the Arabidopsis KNATM gene defines a novel class of KNOX transcriptional regulators missing the homeodomain. *Plant Cell* **20**, 875-887.

Maizel, A., Busch, M.A., Tanahashi, T., Perkovic, J., Kato, M., Hasebe, M., and Weigel, D. (2005). The floral regulator LEAFY evolves by substitutions in the DNA binding domain. *Science* **308**, 260-263.

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

McGinnis, W., Hart, C. P., Gehring, W. J., & Ruddle, F. H. (1984). Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of Drosophila. *Cell*, **38**, 675-680.

McKim, S., and Hay, A. (2010). Patterning and evolution of floral structures - marking time. *Current Opinion in Genetics and Development* **20**, 448-453.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., Marshall, W.F., Qu, L.H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V.V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.L., Cognat, V., Croft,

M.T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.P., Riano-Pachon, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y., Martinez, D., Ngau, W.C., Otiilar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I.V., Rokhsar, D.S., and Grossman, A.R. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**, 245-250.

Mishler, B.D., Lewis, L.A., Buchheim, M.A., Renzaglia, K.S., Garbary, D.J., Delwiche, C.F., Zechman, F.W., Kantz, T.S., and Chapman, R.L. (1994). Phylogenetic relationships of the “green algae” and “bryophytes.” *Annals of the Missouri Botanical Garden* **81**, 451–83.

Moyroud, E., Kusters, E., Monniaux, M., Koes, R., and Parcy, F. (2010). LEAFY blossoms. *Trends in Plant Science* **15**, 346-352.

Mukherjee, K., Brocchieri, L., and Burglin, T.R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. *Molecular Biology and Evolution* **26**, 2775-2794.

Muller, J., Wang, Y., Franzen, R., Santi, L., Salamini, F., and Rohde, W. (2001). In vitro interactions between barley TALE homeodomain proteins suggest a role for protein-protein associations in the regulation of *Knox* gene function. *Plant Journal*, **27**, 13-23.

Muller, R., Bleckmann, A., and Simon, R. (2008). The receptor kinase CORYNE of *Arabidopsis* transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *Plant Cell* **20**, 934-946.

Muthukumar, B., Joyce, B.L., Elless, M.P., and Stewart, C.N., Jr. (2013). Stable transformation of ferns using spores as targets: *Pteris vittata* and *Ceratopteris thalictroides*. *Plant Physiology* **163**, 648-658.

Niklas, K.J., and Kutschera, U. (2009). The evolutionary development of plant body plans. *Functional Plant Biology* **36**, 682-695.

Niklas, K.J., and Kutschera, U. (2010). The evolution of the land plant life cycle. *New Phytologist* **185**, 27-41.

Olsen, A.N., Ernst, H.A., Leggio, L.L., and Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science* **10**, 79-87.

Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J., and Wuthrich, K. (1990). Protein--DNA contacts in the structure of a homeodomain--DNA

complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO Journal* **9**, 3085-3092.

Pham T, Sinha N. (2003). Role of KNOX genes in shoot development of *Welwitschia mirabilis*. *International Journal of Plant Sciences* **164**, 333–43.

Poethig, R.S. (1987). Clonal analysis of cell lineage patterns in plant development. *American Journal of Botany* **74**, 581-594.

Porebski, S., Bailey, L.G., Baum, B.R. (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharides and polyphenol component. *Plant Molecular Biology Reporter* **15**, 8-15.

Proskauer, J.M. (1969). Studies on Anthoceratales. VIII. Phytomorphology **19**, 52–66.

Purves, W.K., Orians, G.H., and Heller, H.C. (1995). *Life: The Science of Biology*, 4th Edition. Sinauer Associates, Sunderland, USA.

Qiu, Y.L., Li, L., Wang, B., Chen, Z., Knoop, V., Groth-Malonek, M., Dombrowska, O., Lee, J., Kent, L., Rest, J., Estabrook, G.F., Hendry, T.A., Taylor, D.W., Testa, C.M., Ambros, M., Crandall-Stotler, B., Duff, R.J., Stech, M., Frey, W., Quandt, D., and Davis, C.C. (2006). The deepest divergences in land plants inferred from phylogenomic evidence. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15511-15516.

Ragni, L., Truernit, E., and Pautot, V. (2007). KNOXing on the BELL:TALE homeobox genes and meristem activity. *International Journal of Plant Developmental Biology* **1**, 42–44.

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

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

Reiser, L., Sanchez-Baracaldo, P., and Hake, S. (2000). Knots in the family tree: evolutionary relationships and functions of knox homeobox genes. *Plant Molecular Biology* **42**, 151-166.

Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.F., Lindquist, E.A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin, I.T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W.B., Barker, E., Bennetzen, J.L., Blankenship, R., Cho, S.H., Dutcher, S.K., Estelle, M., Fawcett, J.A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K.A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D.R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P.J., Sanderfoot, A., Schween, G., Shiu, S.H., Stueber, K., Theodoulou, F.L., Tu, H., Van de Peer, Y., Verrier, P.J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A.C., Hasebe, M., Lucas, S., Mishler, B.D., Reski, R., Grigoriev, I.V., Quatrano, R.S., and Boore, J.L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**, 64-69.

Renzaglia, K.S., and Vaughn, K.C. (2000). Anatomy, development and classification of hornworts. *Bryophyte Biology*. Cambridge: Cambridge University Press. 1–35.

Renzaglia, K.S., Villarreal, J.C., and Duff, R.J. (2009) New insights into morphology, anatomy and systematics of hornworts. *Bryophyte Biology II*, Cambridge: Cambridge University Press. 139 – 171.

Riano-Pachon, D.M., Correa, L.G., Trejos-Espinosa, R., and Mueller-Roeber, B. (2008). Green transcription factors: a chlamydomonas overview. *Genetics* **179**, 31-39.

Richardt, S., Lang, D., Reski, R., Frank, W., and Rensing, S.A. (2007). PlanTAPDB, a phylogeny-based resource of plant transcription-associated proteins. *Plant Physiology* **143**, 1452-1466.

Robles, P., and Pelaz, S. (2005). Flower and fruit development in *Arabidopsis thaliana*. *International Journal of Developmental Biology* **49**, 633-643.

Rowe, N., and Speck, T. (2005). Plant growth forms: an ecological and evolutionary perspective. *New Phytologist* **166**, 61-72.

Rubinstein, C.V., Gerrienne, P., de la Puente, G.S., Astini, R.A., and Steemans, P. (2010). Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). *New Phytologist* **188**, 365-369.

Sakakibara, K., Nishiyama, T., Deguchi, H., and Hasebe, M. (2008). Class 1 KNOX genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development. *Evolution and Development* **10**, 555-566.

Sakakibara, K., Ando, S., Yip, H.K., Tamada, Y., Hiwatashi, Y., Murata, T., Deguchi, H., Hasebe, M., and Bowman, J.L. (2013). KNOX2 genes regulate the haploid-to-diploid morphological transition in land plants. *Science* **339**, 1067-1070.

Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahori, S., and Matsuoka, M. (2001). KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. *Genes & Development* **15**, 581-590.

Sano, R., Juarez, C.M., Hass, B., Sakakibara, K., Ito, M., Banks, J.A., and Hasebe, M. (2005). KNOX homeobox genes potentially have similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems. *Evolution and Development* **7**, 69-78.

Sato, Y., Sentoku, N., Nagato, Y., and Matsuoka, M. (1998). Isolation and characterization of a rice homeobox gene, OSH15. *Plant Molecular Biology* **38**, 983-998.

Sato, Y., Sentoku, N., Miura, Y., Hirochika, H., Kitano, H., and Matsuoka, M. (1999). Loss-of-function mutations in the rice homeobox gene OSH15 affect the architecture of internodes resulting in dwarf plants. *EMBO Journal* **18**, 992-1002.

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

Scofield, S., Dewitte, W., and Murray, J.A. (2007). The KNOX gene SHOOT MERISTEMLESS is required for the development of reproductive meristematic tissues in Arabidopsis. *Plant Journal* **50**, 767-781.

Smith, H. M., Boschke, I., and Hake, S. (2002). Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 9579-9584.

Smith, H.M., and Hake, S. (2003). The interaction of two homeobox genes, BREVIPEDICELLUS and PENNYWISE, regulates internode patterning in the Arabidopsis inflorescence. *Plant Cell* **15**, 1717-1727.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.

Stewart, W.N., and Rothwell, G.W. (1993). *Paleobotany and the Evolution of Plants*, 2nd edition. Cambridge, Cambridge University Press.

Sundas-Larsson, A., Svenson, M., Liao, H., and Engstrom, P. (1998). A homeobox gene with potential developmental control function in the meristem of the conifer *Picea abies*. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 15118-15122.

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

Tanahashi, T., Sumikawa, N., Kato, M., and Hasebe, M. (2005). Diversification of gene function: homologs of the floral regulator FLO/LFY control the first zygotic cell division in the moss *Physcomitrella patens*. *Development* **132**, 1727-1736.

Tattersall, A.D., Turner, L., Knox, M.R., Ambrose, M.J., Ellis, T.H., and Hofer, J.M. (2005). The mutant *crispa* reveals multiple roles for PHANTASTICA in pea compound leaf development. *Plant Cell* **17**, 1046-1060.

Timmermans, M.C., Hudson, A., Becraft, P.W., and Nelson, T. (1999). ROUGH SHEATH2: a Myb protein that represses knox homeobox genes in maize lateral organ primordia. *Science* **284**, 151-153.

Tsiantis, M., Schneeberger, R., Golz, J.F., Freeling, M., and Langdale, J.A. (1999). The maize rough sheath2 gene and leaf development programs in monocot and dicot plants. *Science* **284**, 154-156.

Ueda, M., and Laux, T. (2012). The origin of the plant body axis. *Current Opinion in Plant Biology* **15**, 578-584.

Venglat, S.P., Dumonceaux, T., Rozwadowski, K., Parnell, L., Babic, V., Keller, W., Martienssen, R., Selvaraj, G., and Datla, R. (2002). The homeobox gene BREVIPEDICELLUS is a key regulator of inflorescence architecture in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 4730-4735.

Villarreal, J.C., Cargill, D.C., Hagborg, A., Soderstrom, L., and Renzaglia, K.S. (2010). A synthesis of hornwort diversity: Patterns, causes and future work. *Phytotaxa* **9**, 150-166.

Vollbrecht, E., Veit, B., Sinha, N., and Hake, S. (1991). The developmental gene Knotted-1 is a member of a maize homeobox gene family. *Nature* **350**, 241-243.

Vroemen, C.W., Mordhorst, A.P., Albrecht, C., Kwaaitaal, M.A., and de Vries, S.C. (2003). The CUP-SHAPED COTYLEDON3 gene is required for boundary and shoot meristem formation in Arabidopsis. *Plant Cell* **15**, 1563-1577.

Wagner, A. (2001). Birth and death of duplicated genes in completely sequenced eukaryotes. *Trends in Genetics* **17**, 237-239.

Weigel, D., and Clark, S.E. (1996). Sizing Up the Floral Meristem. *Plant Physiology* **112**, 5-10.

Wellman, C.H., Osterloff, P.L., and Mohiuddin, U. (2003). Fragments of the earliest land plants. *Nature* **425**, 282-285.

Yanai, O., Shani, E., Dolezal, K., Tarkowski, P., Sablowski, R., Sandberg, G., Samach, A., and Ori, N. (2005). Arabidopsis KNOXI proteins activate cytokinin biosynthesis. *Current Biology* **15**, 1566-1571.

Yasumura, Y., Crumpton-Taylor, M., Fuentes, S., and Harberd, N.P. (2007). Step-by-step acquisition of the gibberellin-DELLA growth-regulatory mechanism during land-plant evolution. *Current Biology* **17**, 1225-1230.

Appendix 1

List of *KNOX* genes used for *KNOX* phylogeny

Gene (Class 1 KNOX)	GenBank accession/reference	Species	Taxon
<i>STM</i>	AAC49148.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>BP/KNAT1</i>	NP_192555.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT2</i>	NP_177208.2	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT6</i>	BAB69679.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>EcKNAT1</i>	AEO14149.1	<i>Eschscholzia californica</i>	Angiosperm
<i>EcKNAT6</i>	AAZ93630.1	<i>Eschscholzia californica</i>	Angiosperm
<i>EcSTM</i>	AAZ93630.1	<i>Eschscholzia californica</i>	Angiosperm
<i>ChBP</i>	ABG66654.1	<i>Cardamine hirsuta</i>	Angiosperm
<i>ChSTM</i>	ABF59514.1	<i>Cardamine hirsuta</i>	Angiosperm
<i>OSH1</i>	JQ2379	<i>Oryza sativa</i>	Angiosperm
<i>OSH15</i>	NP_911093	<i>Oryza sativa</i>	Angiosperm
<i>Os_rough_sheath1</i>	BAA79225	<i>Oryza sativa</i>	Angiosperm
<i>OSH10</i>	XP_469241	<i>Oryza sativa</i>	Angiosperm
<i>OSH3</i>	BAA79223	<i>Oryza sativa</i>	Angiosperm
<i>OSH6</i>	NP_914102	<i>Oryza sativa</i>	Angiosperm
<i>OSJNBa0035124.10</i>	AAR87205	<i>Oryza sativa</i>	Angiosperm
<i>OSH71</i>	BAA79226	<i>Oryza sativa</i>	Angiosperm
<i>lg3</i>	AAD13611	<i>Zea mays</i>	Angiosperm
<i>RS1</i>	Q41853	<i>Zea mays</i>	Angiosperm
<i>GN1</i>	AAP76320	<i>Zea mays</i>	Angiosperm
<i>PUEIN13TD</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>kn1</i>	AAP76321	<i>Zea mays</i>	Angiosperm
<i>lg4a</i>	AAP31409	<i>Zea mays</i>	Angiosperm
<i>AZM432736</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM4113483</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM4113484</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>LET6</i>	NP_001233941.1	<i>Solanum lycopersicum</i>	Angiosperm
<i>TKn1</i>	AAC49251.1	<i>Solanum lycopersicum</i>	Angiosperm
<i>Tkn2</i>	AAD00251.1	<i>Solanum lycopersicum</i>	Angiosperm
<i>Tkn3</i>	NP_001233927.1	<i>Solanum lycopersicum</i>	Angiosperm
<i>Tkn4</i>	AAO33774.1	<i>Solanum lycopersicum</i>	Angiosperm
<i>KNAP1</i>	O04134.1	<i>Malus domestica</i>	Angiosperm
<i>KNAP2</i>	O04135.1	<i>Malus domestica</i>	Angiosperm
<i>invaginata</i>	AAL67665.1	<i>Antirrhinum majus</i>	Angiosperm
<i>hirzina</i>	AAL67666.1	<i>Antirrhinum majus</i>	Angiosperm
<i>BOSTM-1</i>	AAM89270.1	<i>Brassica oleracea</i>	Angiosperm

<i>CIKN1</i>	ABC71529.1	<i>Chasmanthium latifolium</i>	Angiosperm
<i>DOH1</i>	CAB88029.1	<i>Dendrobium grex Madame Thong-In</i>	Angiosperm
<i>SBH1</i>	AAA20882.1	<i>Glycine max</i>	Angiosperm
<i>Hakn1</i>	AAM28231.1	<i>Helianthus annuus</i>	Angiosperm
<i>Hakn2</i>	AAM28232.1	<i>Helianthus annuus</i>	Angiosperm
<i>HvKn1</i>	AAQ11882.1	<i>Hordeum vulgare</i>	Angiosperm
<i>Hvh21</i>	AAB81079.1	<i>Hordeum vulgare</i>	Angiosperm
<i>HvKnox3</i>	CAA58503.1	<i>Hordeum vulgare</i>	Angiosperm
<i>InPKn1</i>	BAA31698.1	<i>Ipomoea nil</i>	Angiosperm
<i>InPKn2</i>	BAA31699.1	<i>Ipomoea nil</i>	Angiosperm
<i>InPKn3</i>	BAA31701.1	<i>Ipomoea nil</i>	Angiosperm
<i>Hop1</i>	AAC32262.1	<i>Pisum sativum</i>	Angiosperm
<i>LcKN2</i>	AAX21346.1	<i>Lotus corniculatus var. japonicus</i>	Angiosperm
<i>LcKN3</i>	AAX21347.1	<i>Lotus corniculatus var. japonicus</i>	Angiosperm
<i>MtKN1</i>	AAG27464.1	<i>Medicago truncatula</i>	Angiosperm
<i>PmKN1</i>	ABC71526.1	<i>Panicum miliaceum</i>	Angiosperm
<i>LvHirz</i>	JN005930.1	<i>Linaria vulgaris</i>	Angiosperm
<i>LvIna</i>	JN005931.1	<i>Linaria vulgaris</i>	Angiosperm
<i>FaKNOX1</i>	ACV88386.1	<i>Fragaria x ananassa</i>	Angiosperm
<i>Pt581961</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt570713</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt61066</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt88373</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt662050</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PtKNAT6</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt568092</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PtKN1</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt658310</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>POTH1</i>	AAB41849.1	<i>Solanum tuberosum</i>	Angiosperm
<i>KNOPE1</i>	ABC84491.1	<i>Prunus persica</i>	Angiosperm
<i>KNOPE2</i>	ADD14044.1	<i>Prunus persica</i>	Angiosperm
<i>KNOPE2.1</i>	AFF27502.1	<i>Prunus persica</i>	Angiosperm
<i>KNOPE6</i>	ADC35598.1	<i>Prunus persica</i>	Angiosperm
<i>PpSTMlike 1</i>	ADD14041.1	<i>Prunus persica</i>	Angiosperm
<i>ppSTMlike 2</i>	ADD14042.1	<i>Prunus persica</i>	Angiosperm
<i>PhSTM</i>	AAM47027.1	<i>Petunia x hybrida</i>	Angiosperm
<i>NTH1</i>	BAA76750.1	<i>Nicotiana tabacum</i>	Angiosperm
<i>NTH9</i>	BAA76903.1	<i>Nicotiana tabacum</i>	Angiosperm
<i>NTH15</i>	BAA25546.1	<i>Nicotiana tabacum</i>	Angiosperm
<i>NTH20</i>	BAA76904.1	<i>Nicotiana tabacum</i>	Angiosperm
<i>NTH22</i>	BAA76905.1	<i>Nicotiana tabacum</i>	Angiosperm
<i>PtKN1</i>	AAV54618	<i>Pinus taeda</i>	Gymnosperm

<i>PtKn2</i>	AY680403	<i>Pinus taeda</i>	Gymnosperm
<i>PtKN3</i>	AAV54620.1	<i>Pinus taeda</i>	Gymnosperm
<i>PtKN4</i>	AAV63998.1	<i>Pinus taeda</i>	Gymnosperm
<i>SKN1</i>	AAD00691.1	<i>Picea mariana</i>	Gymnosperm
<i>PmSKn2</i>	AAD00692.1	<i>Picea mariana</i>	Gymnosperm
<i>PmKN4</i>	AAV54621.1	<i>Picea mariana</i>	Gymnosperm
<i>HBK1</i>	AAC84001.1	<i>Picea abies</i>	Gymnosperm
<i>HBK2</i>	AAN77690.1	<i>Picea abies</i>	Gymnosperm
<i>HBK3</i>	AAN77691.1	<i>Picea abies</i>	Gymnosperm
<i>CrKNOX1</i>	BAB18582.1	<i>Ceratopteris richardii</i>	Monilophyte
<i>CrKNOX2</i>	BAB18584.1	<i>Ceratopteris richardii</i>	Monilophyte
<i>Sm159366</i>	Mukherjee et al., 2009	<i>Selaginella moellendorffii</i>	Lycophyte
<i>Sm415291</i>	Mukherjee et al., 2009	<i>Selaginella moellendorffii</i>	Lycophyte
<i>Sm39529</i>	Mukherjee et al., 2009	<i>Selaginella moellendorffii</i>	Lycophyte
<i>SkKNOX1</i>	AAW62517.1	<i>Selaginella kraussiana</i>	Lycophyte
<i>SkKNOX2</i>	AAW62518.1	<i>Selaginella kraussiana</i>	Lycophyte
<i>MKN2</i>	XP_001758592.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MKN4</i>	XP_001781477.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MKN5</i>	XP_001778265.1	<i>Physcomitrella patens</i>	Bryophyte - Moss

Gene (Class 2 KNOX)	GenBank accession/reference	Species	Taxon
<i>KNAT3</i>	AAM53320.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT4</i>	CAA63131.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT5</i>	AAL37042.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>KNAT7</i>	AAG40858.1	<i>Arabidopsis thaliana</i>	Angiosperm
<i>OSH45</i>	Q0J6N4.2	<i>Oryza sativa</i>	Angiosperm
<i>HOS59</i>	BAA77822.2	<i>Oryza sativa</i>	Angiosperm
<i>HOS66</i>	BAB55660	<i>Oryza sativa</i>	Angiosperm
<i>HOS58</i>	BAB55658	<i>Oryza sativa</i>	Angiosperm
<i>AZM440326</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM451939</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM459391</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>AZM448186</i>	Mukherjee et al., 2009	<i>Zea mays</i>	Angiosperm
<i>PtKNAT4</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PtKNAT3</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt672778</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>Pt738080</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PtKNAT3b</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>PpKNAT7</i>	Mukherjee et al., 2009	<i>Populus trichocarpa</i>	Angiosperm
<i>NTH23</i>	BAA25921.1	<i>Nicotiana tabacum</i>	Angiosperm
<i>Let12</i>	NP_001233945.1	<i>Solanum lycopersicum</i>	Angiosperm
<i>StKNAT7</i>	Mukherjee et al., 2009	<i>Solanum tuberosum</i>	Angiosperm
<i>KNOPE3</i>	ABH03531.1	<i>Prunus persica</i>	Angiosperm
<i>KNOPE4</i>	ABO26062.1	<i>Prunus persica</i>	Angiosperm
<i>KNOPE7</i>	AFF27503.1	<i>Prunus persica</i>	Angiosperm
<i>KNAP3</i>	O04136.1	<i>Malus domestica</i>	Angiosperm
<i>CrKNOX3</i>	BAB18585.1	<i>Ceratopteris richardii</i>	Monilophyte
<i>SkKNOX3</i>	AAW62519.1	<i>Selaginella kraussiana</i>	Lycophyte
<i>Sm135843</i>	Mukherjee et al., 2009	<i>Selaginella moellendorffii</i>	Lycophyte
<i>Sm90744</i>	Mukherjee et al., 2009	<i>Selaginella moellendorffii</i>	Lycophyte
<i>MKN1</i>	AAK61309.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MKN6</i>	XP_001765575.1	<i>Physcomitrella patens</i>	Bryophyte - Moss
<i>MpKNOX</i>	--	<i>Marchantia polymorpha</i>	Bryophyte - Liverwort
<i>GSM1</i>	ABS71849.1	<i>Chlamydomonas reinhardtii</i>	Chlorophyte

Appendix 2 - List of primers

Primer name	Primer sequence (5'-3')	Primer description
promSTM_F	GTCGACCCATTATAACCCAACCTAGCTGTC	Forward primer to amplify STM promoter
promSTM_R	GGATCCACTCTCGATCTTCTCTTTCTCTCAC	Reverse primer to amplify STM promoter
pKNAT_Fc	GCGGCCGCATCGAGTGGAGAAAGCCAAA	Internal forward primer to amplify BP promoter
pKNAT_Rb	GTCGACACCCAGATGAGTAAAGATTTGAGAGA	Internal reverse primer to amplify BP promoter
pKNAT1_nest_F	CTTGTAAGTTCAACTTTAGTCCTAGTGAATGC	External forward primer for nested PCR to amplify BP promoter
BPseqR	GATCTAACTCCGGGTCTCTAG	External reverse primer for nested PCR to amplify BP promoter
pKNAT1.ocs3_F1b	GACGGCCCATGGCTCGAG-GTCCTGCTTTAATGAGATATGCG	Forward primer to amplify ocs3' for pBP::KNOX constructs
pKNAT1.ocs3_F2b	GACAGCCCATGGGGTACC-GTCCTGCTTTAATGAGATATGCG	Forward primer to amplify ocs3' for pBP::KNOX constructs
pKNAT1.ocs3_R	GTGTCACCTAAATCTGCGGCCGCCATGGGACAGC	Reverse primer to amplify ocs3' for pBP::KNOX constructs
pSTM.ocs3_F1	GACACCTCTAGAGTCCTGCTTTAATGAGATATG	Forward primer to amplify ocs3' for pSTM::KNOX constructs
pSTM.ocs3_F2	GACACATCTAGAAAGCTTGTCTGCTTTAATGAGATATGCG	Forward primer to amplify ocs3' for pSTM::KNOX constructs
pSTM.ocs3_R	GTGTCACCTAAATCTGCGGCCGCATCGATGACAGC	Reverse primer to amplify ocs3' for pSTM::KNOX constructs
KpnI.Sk1F	GCACGGGTACCATGGAGTTACGGGGTTCG	Forward primer to amplify SkKNOX1 cds for pBP::SkKNOX1 construct
KpnI.Sk1R	CACGCGGTACCTCAGCTTCCACTGCTGGC	Reverse primer to amplify SkKNOX1 cds for pBP::SkKNOX1 construct
XhoI.Sk2F	CGCGACTCGAGATGACAAGCGCTGCTGC	Forward primer to amplify SkKNOX2 cds for pBP::SkKNOX2 construct
XhoI.Sk2R	CAGCGCTCGAGTCAATCACCATCCTG	Reverse primer to amplify SkKNOX2 cds for pBP::SkKNOX2 construct
M_pCr2F.KpnI	CGCACG-GGTACC -ATGGCTACGAACGCTGTG	Forward primer to amplify CrKNOX2 cds for pBP::CrKNOX2 construct
M_pCr2R.KpnI	CGTGAG-GGTACC-TCAATGAAGAGATGACAT	Reverse primer to amplify CrKNOX2 cds for pBP::CrKNOX2 construct
HindIII.Sk2F	CGCAGAAGCTTATGACAAGCGCTGCTGC	Forward primer to amplify SkKNOX2 cds for pSTM::SkKNOX2
HindIII.Sk2R	CAGCGAAGCTTTCAATCACCATCCTG	Reverse primer to amplify SkKNOX2 cds for pSTM::SkKNOX2
pSk1F	GCGCGGGATCCATGGAGTTACGGGGTTCG	Forward primer to amplify SkKNOX1 cds
pSk1FRnew	CGCGCTCTAGATCAGCTTCCACTGCTGGC	Reverse primer to amplify SkKNOX1 cds
pSk2F	CGCGGCCCGGGATGACAAGCGCTGCTGC	Forward primer to amplify SkKNOX2 cds
pSk2R	CCGCGCCCGGGTCAATCACCATCCTG	Reverse primer to amplify SkKNOX2 cds
pCr1F	GCGCGGGATCCATGGACTACGACGGTC	Forward primer to amplify CrKNOX1 cds
pCr1R	CGCGCTCTAGATTACTGCAGCGATGGCAG	Reverse primer to amplify CrKNOX1 cds

pCr2F	GCGCGGGATCCATGGCTACGAACGCTGTG	<i>Forward primer to amplify CrKNOX2 cds</i>
pCr2R	CGCGCGGATCCTCAATGAAGAGATGACAT	<i>Reverse primer to amplify CrKNOX2 cds</i>
M_pBPF	CGCGGCTCGAGATGGAAGAATACCAGCATGAC	<i>Forward primer to amplify BP cds</i>
M_pBPR	CCGCGCTCGAGTTATGGACCGAGACGATAAGG	<i>Reverse primer to amplify BP cds</i>
M_pSTMF	CGCGGCTCGAGATGGAGAGTGGTTCCAACAGC	<i>Forward primer to amplify STM cds</i>
M_pSTMR	CCGCGCTCGAGTCAAAGCATGGTGGAGGAGAT	<i>Reverse primer to amplify STM cds</i>
pSk1F.Mlul	GCGCGACGCGTATGGAGTTACGGGGTCG	<i>Forward primer to amplify SkKNOX1 cds for pMKN2::SkKNOX1 construct</i>
pSk1R.Mlul	CGCGCACGCGTTCAGCTTCCACTGCTGGC	<i>Reverse primer to amplify SkKNOX1 cds for pMKN2::SkKNOX1 construct</i>
M_pSk2F	CGCGGCTCGAGATGACAAGCGCTGCTGC	<i>Forward primer to amplify SkKNOX2 cds for pMKN2::SkKNOX2 construct</i>
M_pSk2R	CCGCGAAGCTTTCAATCACCATCCTG	<i>Reverse primer to amplify SkKNOX2 cds for pMKN2::SkKNOX2 construct</i>
M_pCr1F	GCGCGCTCGAGATGGACTACGACGGTC	<i>Forward primer to amplify CrKNOX1 cds for pMKN2::CrKNOX1 construct</i>
M_pCr1R	CGCGCCTCGAGTTACTGCAGCGATGGCAG	<i>Reverse primer to amplify CrKNOX1 cds for pMKN2::CrKNOX1 construct</i>
M_pCr2Fb	CATGCTCGAGATGGCTACGAACGCTGTGTCTTC	<i>Forward primer to amplify CrKNOX2 cds for pMKN2::CrKNOX2 construct</i>
M_pCr2Rb	CTCGAGTCAATGAAGAGATGACATCATTGTAG	<i>Reverse primer to amplify CrKNOX2 cds for pMKN2::CrKNOX2 construct</i>
M_pMKN2_F	CTCGAGATGGAGCAGCAAACCCGTCGGC	<i>Forward primer to amplify MKN2 cds for pMKN2::MKN2 construct</i>
M_pMKN2_R	CAATGCAGAGGTCATCTGGAGTAGAAGCTT	<i>Reverse primer to amplify MKN2 cds for pMKN2::MKN2 construct</i>
MKN2_3_F	GCGACCCGGGGAACAATCGAATCGTCTTCAGTGATCGT	<i>Forward primer to amplify 3' MKN2 flanking region</i>
MKN2_3_R	GCGAGCGGCCGCAAGCTCTAAATCCTTAATCTTGAAAGACA	<i>Reverse primer to amplify 3' MKN2 flanking region</i>
MKN2_5'_F2	GCGGGGTACCCTTCAGCCTAATTATAAAC	<i>Forward primer to amplify 5' MKN2 flanking region</i>
MKN2_5'_R	CCGCCTCGAGGAACTATAGGGCTGGAAGG	<i>Reverse primer to amplify 5' MKN2 flanking region</i>
NOS_F	GCGCGAAGCTTAGTCAAGCAGGATCGTTC	<i>Forward primer to amplify NOS terminator for pMKN2::KNOX constructs</i>
NOS_R	CGCGCGATATC ATCAGCTTGCATGCCGGTCGATC	<i>Reverse primer to amplify NOS terminator for pMKN2::KNOX constructs</i>
NOS_F.Mlul	GCGCGAAGCTT ACGCGTAGTCAAGCAGGATCGTTC	<i>Forward primer to amplify NOS terminator for pMKN2::SkKNOX1 constructs</i>
NOS_R.Mlul	CGCGC AAGCTT ATCAGCTTGCATGCCGGTCGATC	<i>Reverse primer to amplify NOS terminator for pMKN2::SkKNOX1 constructs</i>
MKN2_5genF2	GACACAATTGGCTCATATGAAG	<i>Forward primer for pMKN2::SkKNOX1 lines genotyping</i>
npII-F2	AGGAGGAAGACAAGGAAGGATAAGGTTGCA	<i>Forward primer for pMKN2::SkKNOX1 lines genotyping</i>

MKN2_3genR1	CTCCAAAACACAGCAACACAG	<i>Reverse primer for pMKN2::SkKNOX1 lines genotyping</i>
PSSk1_R	GAGACGACAGATTCCTGCACAAG	<i>Reverse primer for pMKN2::SkKNOX1 lines genotyping</i>
PSSk2_R	CAATGTTTGTATGCACTGAAGATG	<i>Reverse primer for pMKN2::SkKNOX2 lines genotyping</i>
PSCr1_R	AGGTCTGATTCCAATGTGAGG	<i>Reverse primer for pMKN2::CrKNOX1 lines genotyping</i>
PRCr2_R	GCCATCACAAGCCGAGGGTAAGATG	<i>Reverse primer for pMKN2::CrKNOX2 lines genotyping</i>
AtACT2GSPF	GCCATCCAAGCTGTTCTCTC	<i>Forward primer to amplify ACT2 for RT-PCR</i>
AtACT2GSPR	GAACCACCGATCCAGACACT	<i>Reverse primer to amplify ACT2 for RT-PCR</i>
PpGapC1F3	TCCGTGTCCCCACTACAGAT	<i>Forward primer to amplify GAPC1 for RT-PCR</i>
PpGapC1R3	CTGCACAGGAGATAGGAGCA	<i>Reverse primer to amplify GAPC1 for RT-PCR</i>
Sk1RT_F2	TGGAGTTACGGGGTCGTAGA	<i>Forward primer to amplify SkKNOX1 N-terminus domain (for RT-PCR)</i>
Sk1RT_R1	TGCTACTACCGGATGCACTG	<i>Reverse primer to amplify SkKNOX1 N-terminus domain (for RT-PCR)</i>
Sk2RT_F1	GCTGCTGCATCATCTTCCTC	<i>Forward primer to amplify SkKNOX2 N-terminus domain (for RT-PCR)</i>
Sk2RT_R2	TAAGTGCGTCAATGGGTTCA	<i>Reverse primer to amplify SkKNOX2 N-terminus domain (for RT-PCR)</i>
Cr1RT_F1	ACGGTCATATGCGAGCTTCT	<i>Forward primer to amplify CrKNOX1 N-terminus domain (for RT-PCR)</i>
Cr1RT_R1	CAGACTGTCCACTGCTGGAA	<i>Reverse primer to amplify CrKNOX1 N-terminus domain (for RT-PCR)</i>
Cr2RT_F1	ACGCACTCTGCGTCTTCTTT	<i>Forward primer to amplify CrKNOX2 N-terminus domain (for RT-PCR)</i>
Cr2RT_R1	CGCTCCGATAACTTCTGAGG	<i>Reverse primer to amplify CrKNOX2 N-terminus domain (for RT-PCR)</i>
BPRT_F4	GAGCTCCACCTGATGTGGTT	<i>Forward primer to amplify BP N-terminus domain (for RT-PCR)</i>
BPRT_R4	TTCTGCCGTGCTTCTTTAGG	<i>Reverse primer to amplify BP N-terminus domain (for RT-PCR)</i>
MKN2_RT_F	TGACCAGCTCCCACAATGTA	<i>Forward primer to amplify MKN2 N-terminus domain (for RT-PCR)</i>
MKN2_RT_R	TAACGTGGGCTAGGACCATC	<i>Reverse primer to amplify MKN2 N-terminus domain (for RT-PCR)</i>
ArabiSF3	GCGTTCAAAAGTCGCCTAAG	<i>Forward primer to amplify probe for pBP::KNOX lines Southern blot analysis</i>
ArabiSR3	ATACTTTCTCGGCAGGAGCA	<i>Reverse primer to amplify probe for pBP::KNOX lines Southern blot analysis</i>
PpSouthF1	ATCAACAGCATCCCCATCTC	<i>Forward primer to amplify probe for pMKN2::KNOX lines Southern blot analysis</i>

PpSouthR1	CTACGGGGGAGTCAACGTAA	<i>Reverse primer to amplify probe for pMKN2::KNOX lines Southern blot analysis</i>
------------------	----------------------	---