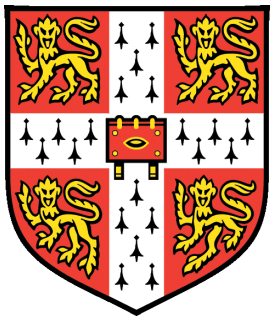
Role of *KNOX* genes in the evolution and development of floral nectar spurs



Mathew Box

Department of Plant Sciences

University of Cambridge

A thesis submitted for the degree of

Doctor of Philosophy

February 2010

Research is the act of going up alleys to see if they are blind

- Plutarch.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration unless stated otherwise in the text. The morphological results presented in Chapter 3 have been published in the *Botanical Journal of the Linnean Society* in collaboration with Richard Bateman, Beverley Glover and Paula Rudall (Box *et al.*, 2008).

- Mathew Box, April 1, 2010

Acknowledgements

Even though this PhD dissertation can yield at most one doctor's title for at most one person, it has involved the help, interest, and support of many, some of whom already have a PhD of their own, and others who do not, but most certainly deserve one.

First and foremost thanks to my supervisor Beverley Glover, who has contributed enormously to this work both in terms of intellectual input and of support and encouragement. Despite the extent of her teaching, research, administrative and family responsibilities, she has always given generously of her time and insights. I feel deeply grateful and privileged to have been her student.

I would also like to thank my co-supervisors Paula Rudall and Richard Bateman at the Royal Botanic Gardens, Kew. I have known Paula and Richard for some years now; their continued support, advice and expertise have been integral to the development of my career thus far and have continued to be a crucial factor in the success of this project. Particular credit is owed to Steven Dodsworth with whom I have enjoyed a close collaboration throughout the duration of this project. Whether in the guise of an undergraduate project student, or summer placement student, Steve has shown a particular interest in this work and has helped me to advance the project to the point it is at today. I would have been unable to make such significant progress without his help and enthusiasm. Special thanks are also owed to Tim Upson and Alex Goodall at the Cambridge University Botanic Gardens for organising the care of my wild collected plant material and John Haggar for advice and assistance with caring for my orchids.

Last, but certainly not least, thanks to everyone in the lab for their day-to-day advice and encouragement, particularly to Kit Wilkins for teaching me all of the fundamental techniques at the start of my PhD. Thanks also to Elinor Thompson, Heather Whitney, Matthew Dorling, Murphy Thomas, Keith Macaulay, Rubén Álvarez Fernández, Rachel Walker and Samuel Brockington for helping me with pointless banter and various experimental techniques. Finally, thanks to Helen for listening to endless science speak and for her love and encouragement.

For financial support I acknowledge the BBSRC, Corpus Christi College and the Campbell Trust Fund.

Abstract

A key question in biology is how changes in gene function or regulation produce new morphologies during evolution. The nectar spur is an evolutionarily labile structure known to influence speciation in a broad range of angiosperm taxa. Here, the genetic basis of nectar spur development, and the evolution of differences in nectar spur morphology, is investigated in *Linaria vulgaris* and two closely related species of orchid, the primitively longer-spurred *Dactylorhiza fuchsii*, and more derived short-spurred *D. viridis* (Orchidinae, Orchidaceae).

Despite considerable morphological and phylogenetic differences, nectar spur ontogeny is fundamentally similar in each of the study species, proceeding from an abaxial bulge formed on the ventral petal relatively late in petal morphogenesis. However, spur development is progenetically curtailed in the short-spurred orchid *D. viridis*. In each case spur development involves class 1 KNOTTED1-like homeobox (KNOX) proteins. *KNOX* gene expression is not restricted to the spur-bearing petal, indicating that additional components are required to define nectar spur position, e.g. canonical ABC genes, determinants of floral zygomorphy, and additional (currently unknown) factors. However, constitutive expression of class 1 KNOX proteins in transgenic tobacco produces flowers with ectopic outgrowths on the petals, indicating that KNOX proteins alone are, to some degree, capable of inducing structures similar to nectar spurs in a heterologous host. Interestingly, *KNOX* gene expression is high in the ovary of all study taxa, suggesting that KNOX proteins may also have been involved in the evolution of this key angiosperm feature.

Although principally involved in maintaining indeterminacy in the shoot apical meristem (SAM), members of the *KNOX* gene family have been co-opted in the evolution and development of compound leaves where they suppress differentiation and extend the morphogenetic potential of the leaf. A similar model is presented here to explain the role of KNOX proteins in nectar spur development. Co-option of *KNOX* gene expression to the maturing perianth delays cellular differentiation, *facilitating* the development of the nectar spur but requiring additional, unknown factors, to determine nectar spur fate. As facilitators of nectar spur development, changes in the spatio-temporal patterns of *KNOX* gene expression may alter the potential for nectar spur development and explain the critical length differences observed between the orchids *D. fuchsii* and *D. viridis* (and among other angiosperm taxa). Taken together, the available data indicate that *KNOX* genes confer a meristematic state upon plant tissues in a variety of morphogenetic contexts, making the gene family a potentially versatile tool to mediate a wide variety of evolutionary transformations.

Contents

List of Figures

List of Tables

Abbreviations

BAP	6-Benzylamino purine (synthetic cytokinin)
BCIP	5-bromo-4-chloro-indolylphosphate-4 toluidine salt
BI	Bayesian inference
BLAST	basic local alignment search tool
bp	nucleotide base pairs
BSA	bovine serum albumen
CIP	calf intestinal phosphatase
CaMV 35S	cauliflower mosaic virus 35S constitutive promoter
cDNA	complementary DNA
cfu	colony forming units
CODEHOP	consensus degenerate hybrid oligonucleotide primers
cPCR	colony PCR
CTD	C-terminal domain of a protein
DEPC	diethyl pyrocarbonate
FAA	formalin-acetic-alcohol
dH_2O	deionised water
DIG-11-UTP	Digoxigenin-11-uridine-5-triphosphate
DMSO	dimethyl sulphoxide
dNTP	deoxynucleotide triphosphate
dPCR	degenerate PCR
EDTA	ethylene diaminetetraacetic acid
EtBr	ethidium bromide
EtOH	ethanol
GA	gibberellic acid
gDNA	genomic DNA
GoF	gain of function mutant
GSP	gene specific primer
HD	homeodomain

IAA	indole-3-acetic acid (auxin)
IMS	industrial methylated spirit
IPTG	isopropyl thiogalactoside
ISH	*in situ* hybridisation
IVT	*in vitro* transcription
Kb	thousand nucleotide basepairs
KNOX	*KNOTTED1*-like homeobox transcription factors
LB	Luria-Bertani growth medium
LoF	loss of function mutant
MD	meinox domain
ME	minimum evolution
ML	maximum likelihood
MOPS	3-[N-morpholino] propanesulphonic acid
MP	maximum parsimony
mRNA	messenger RNA
MS	Murashige-Skoog medium
NBT	4-nitro-blue-tetrazolium chloride
nt	nucleotide
NTD	N-terminal domain of a protein
NTE	sodium chloride, tris EDTA buffer
NJ	neighbour joining
OD_{xxx}	optical density at XXXnm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PTGS	post transcriptional gene silencing
PVP	polyvinylpyrrolidone
QRT-PCR	quantitative real time PCR
RACE	rapid amplification of cDNA ends
RLM-RACE	RNA ligase mediated rapid amplification of cDNA ends
RT-PCR	reverse transcription PCR
SDS	sodium dodecyl sulphate
SEM	scanning electron microscopy/microscope
SOC	super optimal broth (catabolite repression)
SSC	sodium chloride/sodium citrate buffer

TBE	tris-borate-EDTA buffer
T-DNA	transferred DNA (transgene)
TFB	transformation buffer
TFs	transcription factors
Ti	tumor inducing plasmid
TRV-VIGS	tobacco rattle virus viral induced gene silencing
UTR	untranslated region
U	enzymatic units (may vary with supplier)
X-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside
XRT	X-ray computed tomography

Chapter 1

Introduction

Of the five major lineages of seed plants the angiosperms are by far the most speciose, consisting of between 250,000 and 400,000 extant species. Other extant seed plant lineages are considerably less diverse, consisting of approximately 600 species of Conifer, 120 species of Cycad, 70 species of Gnetales and one species of *Ginkgo*. Fossil evidence suggests that the seed plants originated approximately 370 MYA (Kenrick & Crane, 1997), towards the end of the Devonian. By the end of the Late Carboniferous/Early Permian (290-320 MYA), at least three of the five extant lineages of seed plants had already diverged (Donoghue & Doyle, 2000; Kenrick & Crane, 1997). By contrast, the first unambiguous evidence of angiosperms in the fossil record is from the early Cretaceous, only 130 MYA (Friis *et al.*, 1999) and by 90 MYA most of the extant angiosperm lineages were established (Crepet & Niklas, 2009). In just 40 million years the angiosperms had spread from their point of origin in the tropics, and diversified dramatically, generating the vastly diverse and dominant flowering plant communities we can observe on almost every continent and in almost every available biome.

Molecular data, however, suggest that the angiosperms may have originated considerably earlier than evidenced from the fossil record, c.290 MYA (Qiu *et al.*, 1999). Despite this, it is clear that angiosperms have diversified considerably by comparison to other seed plant lineages. On two occasions in letters to Oswald Heer (1875) and J.D. Hooker (1879), Charles Darwin referred to the rapid diversification of the angiosperms as both an '*abominable mystery*' and an '*abominably perplexing phenomenon*', see Darwin & Seward (1903). Darwin's confusion as to the sheer number and diversity of angiosperm species was derived from the fact that evolution was thought to proceed by a slow, gradual accumulation of change, over significant periods of geological time.

Precisely how the angiosperms have become so diverse in such a short period of time has remained a prevailing mystery (Crepet & Niklas, 2009). However, in recent years

considerable progress has been made in our understanding of this phenomenon. The answer, somewhat predictably, is a complex one. For many authors, however the flower is considered to be the main, although not the sole, angiosperm innovation that led to the rapid diversification of flowering plants (Crepet, 2008; Crepet & Niklas, 2009; Dilcher, 2000). Understanding how the possession of flowers has contributed to the evolution of such astounding levels of biological diversity in angiosperms remains one of the principal aims in the new field of plant evolutionary development (Evo-Devo).

1.1 Solving Darwin's *abominable mystery...*

The angiosperm flower is a remarkable feat of evolutionary tinkering. Although the flower in itself is of vast ecological importance to the success of the angiosperms, specialised floral structures are arguably more important to angiosperm diversity than other general attributes of flowers.

The key to understanding the astounding diversity of flowering plants is in the plethora of floral specialisations for an even greater diversity of insect pollinators (Burger, 1981; Dilcher, 2000; Grant, 1949; Stebbins, 1970). This has permitted the co-evolution of mutually beneficial animal-plant relationships that are nowhere better developed, nor more widespread, than among angiosperm species and their biotic vectors for pollination and dispersal (Crepet & Niklas, 2009).

The tale of Darwin's orchid, *Angraecum sesquipedale* (Figure 1.1, 1.2), with an exceptionally long floral nectar spur up to 35cm in length, provides us with an example of just how plant-pollinator interactions can co-evolve. Darwin (1862) predicted that the flower was pollinated by a hawkmoth with an equally long proboscis, long enough to reach the nectar reward at the base of the floral nectar spur. Darwin was ridiculed for this idea until the later discovery of such a moth, *Xanthopan morganii* ssp. *praedicta*, in 1903 with a tongue length of 22cm. Whilst co-evolutionary theories such as these have led to somewhat misleading generalisations about the convergent evolution of specific plant-pollinator traits in a diverse range of angiosperm taxa (Ollerton *et al.*, 2009), undoubtedly it is the possession of evolutionarily labile floral features, such as nectar spurs, that have helped make angiosperms the most diverse group of land plants we see today.

1.2 What are nectar spurs and why are they important?

Nectar spurs are often described as a *key innovation* in the botanical literature and are linked to the high species diversity of many angiosperm groups (Figure 1.1). They are cylindrical or conical outgrowths of petals, less commonly sepals, which increase the distance between the floral reward (nectar) and the reproductive parts of the flower (Hodges *et al.*, 2004).

Spurs are widely considered to have originated as adaptations to accommodate large volumes of nectar and attract long-proboscid pollinators. In many angiosperms the spur is a non-secretory structure that merely acts as a reservoir in which nectar, secreted from elsewhere in the flower, accumulates e.g. the spurs of most Antirrhineae and the Butterfly orchids (*Platanthera*) (Stpiczynska, 2003; Sutton, 1988). However, in other angiosperm species, the spur is not simply a petal-derived tube that has evolved to store nectar but also to secrete it. This has been particularly well documented for the fragrant orchid *Gymnadenia conopsea* and its close relatives in which the spur has micromorphological features associated with nectar secretion and re-absorption (Bell *et al.*, 2009; Box *et al.*, 2008; Stpiczynska & Matusiewicz, 2001).

1.2.1 Nectar spurs have evolved repeatedly and can be associated with high species diversity

Nectar spurs occur in a large variety of different flowering plant groups (Figure 1.1) in both monocots (e.g. Orchidaceae and Liliaceae) and eudicots (e.g. Ranunculaceae, Plantaginaceae and Lentibulariaceae) (Endress & Matthews, 2006), yet are thought to be derived features that are absent from the basal angiosperms (Donoghue & Mathews, 1998; Takhtajan, 1969, 1991). Spurs have therefore evolved repeatedly on multiple occasions in angiosperms (Hodges, 1997), especially among eudicots and within Orchidaceae, though more rarely among non-orchid monocots, e.g. *Tricyrtis* (Rudall *et al.*, 2003).

Spur morphology is intimately tied to reproduction. Characters that affect reproductive isolation are likely drivers of speciation. Simple differences in the length, shape, orientation, colouration and even curvature of spurs are thought to be associated with different pollinators and affect reproductive isolation (Hodges & Arnold, 1995), an underlying requirement of most speciation concepts (Dobzhansky, 1937; Grant, 1963; Mayr, 1942). Where present, nectar spurs are associated with high species diversity (Hodges, 1997; Hodges & Arnold, 1995; Hodges *et al.*, 2002). Hodges & Arnold (1995) used a

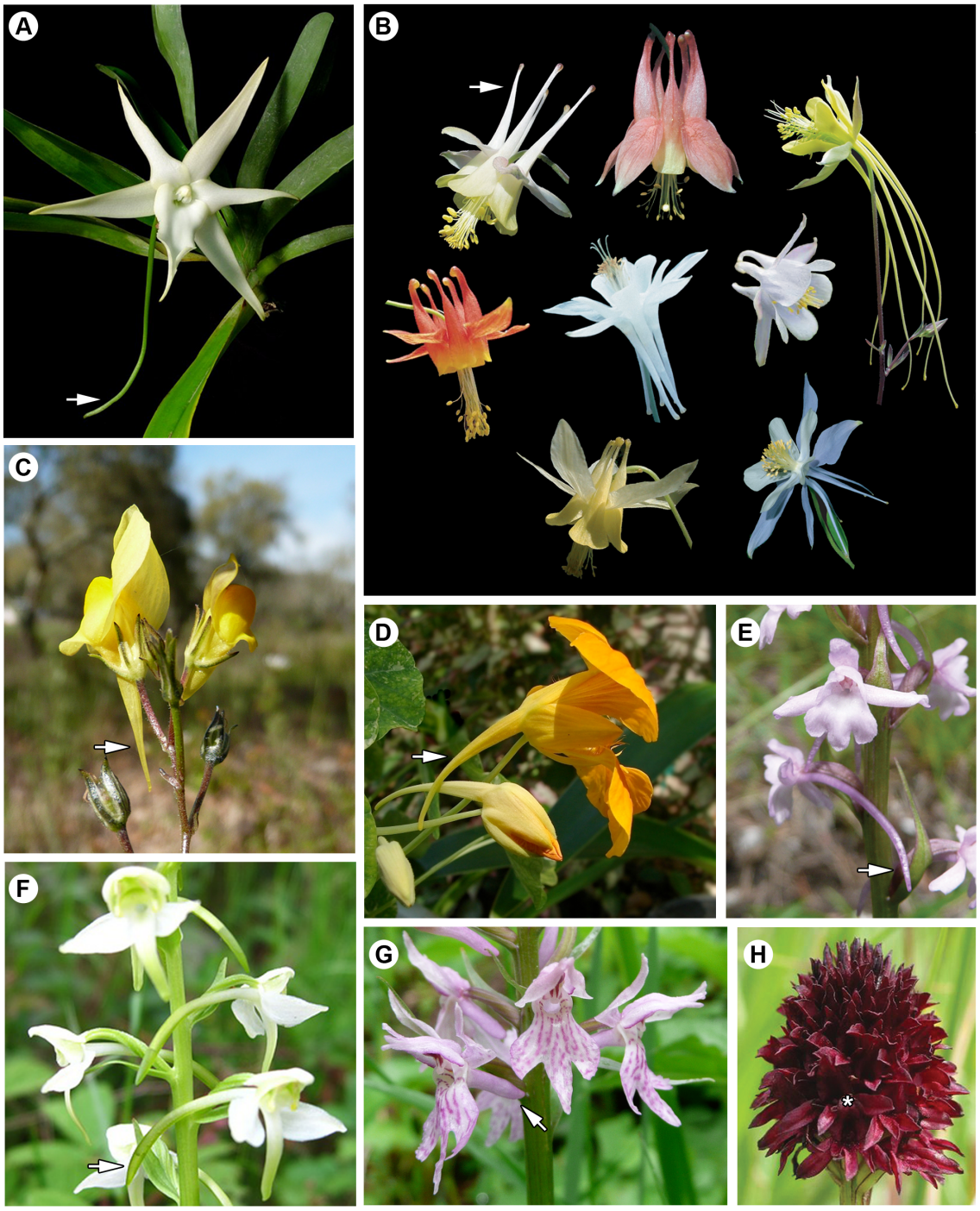


Figure 1.1: Nectar spurs are present in diverse angiosperm taxa
A. The exceptionally long nectar spur of *Angraecum sesquipedale*. B. Spur morphology varies considerably among *Aquilegia spp.* (Ranunculaceae), photo: S.A. Hodges. C. *Linaria vulgaris* (Plantaginaceae). D. The axial spur of *Tropaeolum majus* (Tropaeolaceae) is derived from the receptacle. E-H. The variety of nectar spurs in orchids (Orchidaceae), nectar spurs are long and narrow in (E) *Gymnadenia conopsea* and long and broad in (F) *Platanthera bifolia* and (G) *Dactylorhiza fuchsii*, but much reduced in length in (H) *G. rhellicani*. Arrows indicate nectar spurs, the asterisk indicates the short obscured spurs of *G. rhellicani*.

molecular phylogenetic approach to relate the evolution of spurs in the columbine genus *Aquilegia* (Ranunculaceae) to an increase in the diversification rate of the columbines (Figure 1.1). In addition to phylogenetic assessments, the link between spurs and reproductive isolation has been clearly demonstrated by field experiments.

Some of the best supporting evidence for this relationship was provided by Hodges & Arnold (1994) who measured the shapes of clines for floral features and molecular markers between *Aquilegia formosa* and *A. pubescens* across a hybrid-zone. Cline shape describes the rate of change in a given character state across a physical distance. The steepness of the cline depends on the strength of natural selection and gene flow, and can be used to infer how selection is operating on a given character state across a hybrid zone and how important that character is in promoting reproductive isolation. Hodges & Arnold (1994) found that most floral characters, including spur length and orientation, formed steep clines, whilst Randomly Amplified Polymorphic DNA (RAPD) markers, presumably neutral with respect to plant fitness, formed broader clines. These data firmly indicate that, between *A. formosa* and *A. pubescens* at least, within species variation in floral spurs is an important reproductively isolating character possibly acting through differences in pollinator visitation.

Other taxa with nectar spurs also show this relationship. Experimentally reducing the length of the spur in several disparate groups of taxa has demonstrated the close link between nectar spurs and reproductive success. In the moth-pollinated orchid genus *Platanthera*, experimentally reducing nectar spur length demonstrated a pronounced effect on both the insertion and removal of pollinia (pollen), and subsequent fruit set (Little *et al.*, 2005; Nilsson, 1988, 1983), although fruit set may not be an important indicator of ecological success in orchids, which often produce a surfeit of dust-like seed. However, it is clear that the length of an individual plant's spur can strongly influences reproductive success. However, this relationship is by no means perfect. A number of abiotic factors such as latitude, soil characteristics (pH and moisture content) and even the degree of shade experienced by the orchids, have also been shown to play their part in determining spur length in *Platanthera* (Bateman & Sexton, 2008). Factors such as these should also be borne in mind when considering changes in spur length.

Additional evidence has been obtained from a variety of natural populations. Members of the *Disa draconis* complex (Orchidaceae) have a median sepal-derived spur (Kurzweil, 1998) that varies widely in length. Variation in spur length has been correlated with the proboscis length of the flies that visit them (Johnson & Steiner, 1997; Johnson *et al.*, 1998) (Figure 1.2). Johnson & Steiner (1997) observed pollinator visitation in multiple species of *Disa* along transects and identified several long-proboscid fly pollinators. The

insects carrying pollinia of *Disa spp.* were captured, identified and the proboscis length measured. The length of the spur and the proboscis of the pollinator are highly correlated. In addition to this correlative data, experimentally reducing the length of the spur in wild populations, by constricting the spur with yarn, reduced both the frequency of pollen deposition and subsequent fruit set compared to that of unmanipulated control plants.

These studies strongly suggest that pollinators can influence spur morphology, and in doing so may provide a mechanism for reproductive isolation. At least in some taxa, the evolution of spurs may, in some part, explain how so many angiosperm species have evolved in such a short time.

1.2.2 Punctuated shifts in spur morphology are required for rapid speciation

How are spurs so influential in terms of morphological evolution and speciation?

1.2.2.1 The co-evolutionary race model

To return to the example of Darwin's orchid, *Angraecum sesquipedale*, the relationship between the exceptionally long spur and long-proboscid pollinator was explained by Darwin (1862), and later Wallace (1867), to result from a co-evolutionary race (Figure 1.2). According to this hypothesis pollinators with the longest tongues are selected for because they can obtain the largest food reward. Plants with the longest nectar spurs are therefore selected because they are able to provide larger food rewards and also, to ensure that the pollinators optimally contact the reproductive organs of the flower, thereby allowing the plant to achieve the greatest reproduction. The plant and pollinator subsequently enter a co-evolutionary race in which nectar spur and tongue length continually and gradually increase in length until halted by other selectable constraints. Whilst this model makes logical sense, it implies that changes in spur morphology are gradual. Such a gradual change is not consistent with the rapid and extensive diversifications demonstrated for many taxa with nectar spurs.

1.2.2.2 The pollinator-shift model

The rapid and extensive diversification of the North American *Aquilegia spp.* is the result of a recent and rapid radiation related to the increasing length of floral nectar spurs (Hodges & Arnold, 1994, 1995). The traditional co-evolutionary race model does not explain how such a rapid radiation would occur. However, recently a more satisfactory model has been proposed (Whittall & Hodges, 2007).

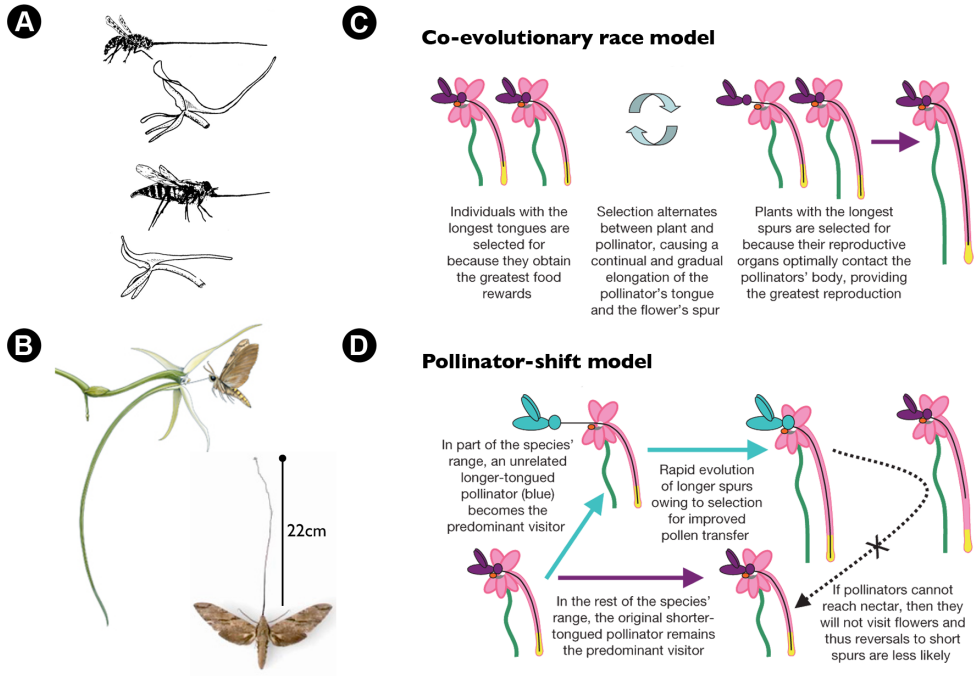


Figure 1.2: Nectar spurs promote speciation through reproductive isolation
A. Spur length in the *Disa* complex (Orchidaceae) is strongly related to pollinator tongue length, (Top) *D. draconis* long-spurred form and the tanglewing fly *Moegistorynchus longirostris*, (Bottom) short-spurred form and the horsefly *Piloliche rostrata*, modified from Johnson & Steiner (1997). B. The exceptionally long nectar spur of *Angraecum sesquipedale* and equally long tongue of its pollinator *Xanthopan morganii* ssp. *praedicta*, image: E. Damstra. C. The traditional co-evolutionary race model. D. The pollinator-shift model. (C & D from Whittall & Hodges (2007)).

According to this model spur lengths have been increasing, not because of co-evolution with the tongue length of an associated pollinator, but as a result of shifts in the pollinator demographic as a consequence of colonisation by plants of new habitats, or changes in pollinator abundance (Figure 1.2). This model assumes that tongue length is less evolutionarily labile than spur length, which changes in accordance with the tongue lengths of available pollinators. As the plants colonise a new area or experience a difference in pollinator abundance, perhaps in relation to climate change, they become exposed to a variety of new pollinators that they may not have encountered before. Some of these pollinators may have longer proboscis lengths than is optimal for efficient transfer of pollen. Equally, pollinators with incorrectly sized mouthparts will avoid plants with sub-optimal spur-length (Hodges, 1995). As a result, selection favours plants that are able to increase their spur length to accommodate the new pollinators that are available to them (Whittall & Hodges, 2007), rapidly promoting reproductive isolation and subsequent speciation.

The Whittall & Hodges (2007) model considers the increasing spur-length of North American *Aquilegia spp.* to be the result of **pollinator-shifts** rather than traditional co-evolution. Pollinator shifts generally result in reproductive isolation (Grant, 1949), such a pollinator-shift hypothesis thereby concentrates change in spur-length at speciation events and permits punctuated morphological evolution, consistent with rapid and extensive diversification. This hypothesis goes a long way to explaining how angiosperm lineages with spurs are highly speciose, however the traditional co-evolution model may also be important under different circumstances, possibly acting in concert with pollinator-shifts to generate changes in spur length over longer evolutionary timescales.

The pollinator-shift model predicts only that spur-length will increase and that reductions in spur length are unlikely to occur as pollinators avoid flowers from which they cannot gain a reward (Hodges, 1995; Whittall & Hodges, 2007). Reduction in spur length, or even spur loss, have also been documented in a number of angiosperm taxa. A particularly good example of spur loss/reduction is among members of the orchid subtribe Orchidinae, in which nectar spur losses or reductions have also been tied to speciation events (Bateman & DiMichele, 2002; Bateman & Rudall, 2006).

Despite the hypothetical limitations postulated by the Whittall & Hodges (2007) pollinator-shift model, speciation associated with spur reductions may also be related to pollinator-shifts or changes in reproductive strategy. A significant reduction in the abundance of long-proboscid pollinators, for example, could promote rapid diversification of short-spurred species under certain conditions, perhaps by selectively favouring short-spurred species that may be able to invest less in expensive floral rewards and elaborations. Whilst it is unclear precisely how spurs are associated with high species diversity, what

is clear is that the possession of spurs facilitates effective pre-zygotic isolation, which is a necssary first step towards reproductive isolation and the evolution of a new species.

1.2.3 Nectar spurs: a model system for studies of speciation and character evolution

The nectar spur is a structure of acknowledged function(s) that is known to influence speciation, is unusually evolutionarily labile (in terms of both acquisition and loss), commonly subjected to strong directional and/or disruptive selection, and is frequently discussed in the literature as a key innovation and an example of parallel and/or convergent evolution that has played a critical role in determining close co-evolutionary relationships with specific plant pollinators. These properties make nectar spurs an excellent system to address a number of fundamental evolutionary questions (see section 1.6.4).

1.3 Understanding the evolution and development of nectar spurs

A central question in biology concerns how differences in form arise and are established over evolutionary time. Advances in developmental genetics during the 1980s provided the foundation for researchers to answer this question (Nusslein-Volhard, 1994) and bridge the gap between evolutionary and developmental science; two disciplines which had been pursuing different goals since the separation of phylogeny from ontogeny and the collapse of Haeckel's 'Biogenetic law' (Garstang, 1922).

Garstang (1922), was one of the principal opponents to the biogenetic law, and provided a more sophisticated analysis of the relationship between evolution and development, showing that alterations in development could produce evolutionary changes, a concept I will introduce in Chapter 3. The work of Garstang (1922) and his contemporaries, in combination with more recent advances in developmental genetics, now makes it possible to systematically assess questions pertaining to the evolution of form under the new discipline of 'Evo-Devo'. This new field was principally developed in animal systems, making use of the considerable advances in animal developmental genetics that followed the great impetus in molecular biology during the 1980s and 1990s. Presently a variety of molecular-genetic, developmental and systematic tools are available and can be employed in the study of the evolution of form. Recent progress in plant developmental genetics now makes it possible to explore exactly how evolutionary processes have shaped the diversity of flowering plants.

In the last 15 years considerable progress has been made in our understanding of plant development and morphological evolution using such Evo-Devo approaches, particularly those concerned with the evolution of floral developmental programmes (Coen & Meyerowitz, 1991; Kramer & Irish, 1999; Soltis *et al.*, 2002; Theißen & Saedler, 1999). In order to take advantage of the power of this new discipline a good understanding of phylogenetic relatedness in the study taxa, the ontogeny of the morphological features under test, and some idea of the developmental genetics of the trait of interest are required. As such Evo-Devo is a powerful, but demanding, field of scientific study.

1.3.1 Morphological development of nectar spurs

In the absence of a genetic model species with floral nectar spurs, developmental studies of spur formation are few and far between. Those that have been conducted have focused on members of the Ranunculaceae and Orchidaceae, the traditional main stays of nectar spur research. Tepfer (1953) and Gottlieb (1984) studied patterns of spur development in *Aquilegia* and noted that the spur was among the last of the floral organs to develop in the ontogeny of the flower, almost as if it had been added as an innovation to the end of floral ontogeny.

More recently, nectar spur development has been more fully characterised in *Aquilegia* (Tucker & Hodges, 2005) and spur-bearing members of the orchid subtribe Orchidinae (Box *et al.*, 2008). In *Aquilegia* the nectar spur starts development as an abaxial bulge late in petal ontogeny (Tucker & Hodges, 2005), this is also true of nectar spur development in Orchidinae in which the spur is initiated at the base of the differentiated labellum (lip) petal (Box *et al.*, 2008). Examination of nectar spur ontogeny in each of these plant groups demonstrates that despite the significant morphological differences between the flowers of *Aquilegia* and spur bearing Orchidinae, nectar spurs are among the last floral organs to develop and are derived from a late forming abaxial bulge at the base of one or more petals that are themselves in a relatively advanced stage of ontogeny.

These observations could indicate that nectar spurs develop by the same underlying means in diverse angiosperm taxa, despite their multiple independent origins (Hodges, 1997; Hodges & Arnold, 1995). Furthermore, comparing floral nectar spur ontogeny in ancestrally longer-spurred orchid species to those with shorter spurs demonstrates that differences in spur length between closely related species are driven principally by shifts in the timing of organ development (heterochrony; section 3.4.2) (Box *et al.*, 2008). The same may also be said of the increasing length of nectar spurs documented among North American species of *Aquilegia*. These morphological shifts may be related to changes in the timing of expression of genes related to spur initiation, growth and differentiation.

1.3.2 The current state of nectar spur developmental genetics

What we currently know of the genetic basis of nectar spur development is very limited, but we have been able to glean quite a lot of information about the relationship of the spur development pathway to other key processes in floral development. Virus Induced Gene Silencing (VIGS) of the *Aquilegia vulgaris* petal-determining *PISTILLATA* B-function gene (*AqvPI*), generates a typical B-function silencing phenotype, in which the flowers are lacking in petalloid characters (Figure 1.3). Strikingly, the subsequent apetalloid organs produced are devoid of floral nectar spurs (Kramer *et al.*, 2007). Whilst this experiment by no means indicates a genetic basis for spurs it does show that genes involved in nectar spur development operate downstream of canonical 'ABC' genes. This is not surprising as nectar spurs are predominantly petal-derived structures.

According to the ABC model of floral development A, B and C- function transcription factors act in a combinatorial manner to determine floral organ identity (Coen & Meyerowitz, 1991). Class A proteins alone are responsible for the development of sepals, but act in combination with class B proteins to effect petal development. C-function proteins determine carpel identity, but act together with class B proteins to determine the development of stamens. In recent years the ABC model has been expanded to include D and E- function genes, see Krizek & Fletcher (2005) for a detailed review of the ABC model.

It is likely that genes involved in nectar spur development may be responsive to canonical A and B-function MADS-box transcription factors, orthologous to the A and B-function proteins APETALA1 (AP1) and PISTILLATA (PI) of *Arabidopsis*, SQUAMOSA (SQA) and GLOBOSA (GLO) of snapdragon. Although not all A-function proteins encode MADS-box transcription factors, most ABC genes belong to this family of transcription factors, as such genes involved in nectar spur development may carry MADS-box transcription factor binding sites (CArG boxes) in their promoters.

In *Aquilegia* nectar spurs form on each petal, however in many other angiosperms there is a close correlation between the possession of a spur and floral zygomorphy (Neal *et al.*, 1998; Sargent, 2004). Many angiosperm taxa only possess a single nectar spur which forms on a precisely defined petal. Much of what we know about floral zygomorphy at the genetic level has been derived from genetic and molecular analyses of snapdragon, *Antirrhinum majus*, and common toadflax, *Linaria vulgaris* (both Veronicaceae, Lamiales; Olmstead *et al.* (2001)). *Linaria* is a close relative of snapdragon with a single spur located in a ventral position at the base of the corolla tube.

It is outside the scope of this work to discuss floral zygomorphy in detail (for a review see (Cubas, 2004)), however the isolation and characterisation of orthologous

symmetry-breaking genes in a well-characterised radially-symmetrical mutant of common toadflax (Cubas *et al.*, 1999), with five-spurred actinomorphic flowers, clearly demonstrates that the genes governing nectar spur development are also downstream of floral symmetry breaking genes (Figure 1.3). In snapdragon and *Linaria*, these genes include *DIVARICARATA* (*DIV*) and *RADIALIS* (*RAD*) which determine ventral and dorsal petal identity respectively; both of which act in combination with two additional dorsalising factors, CYCLOIDEA (CYC) and DICHOTOMA (DICH), that retard the growth at the dorsal part of the floral meristem and control the number of organ primordia that form in this region, later affecting the growth patterns of dorsal petals and arresting development of the dorsal most stamen.

1.3.3 How many genes regulate spur development?

Ontogenetic analyses of nectar spurs have given us a few indications as to the complexity of genetic pathways that determine nectar spurs at the molecular level. The observation that spur development occurs late in floral ontogeny in a number of unrelated angiosperms led Gottlieb (1984) to suggest that just one major neomorphic allele could be involved in the evolution and development of floral nectar spurs.

More convincing demonstrations of the involvement of only one or few key genes comes from a set of now classic genetic experiments using *Aquilegia* (Prazmo, 1965). In these experiments non-spurred *A. ecalcarata* was crossed with spurred species of *Aquilegia*. The resulting F2 generation (following selfing of the F1) had spurs in an approximately 3:1 Mendelian ratio, indicating not only that the possession of spurs was a dominant trait but that spur inheritance is governed by only one or two genes.

A number of equally convincing studies, however, have shown that spur development may be highly complex at the genetic level. Naturally occuring intra-generic orchid hybrids have been observed in between medium and short-spurred species such as *Gymnadenia odoratissima* and *G. rhellicani*, respectively. The progeny of such hybridisation events are often intermediate in many floral traits including nectar spur length (Box & Bateman, personal observation; Figure 1.3). Whilst these observations suggest that there are likely to be a a small number of genes of large effect, the intermediacy of the phenotype indicates that additional genes may also influence nectar spur length. Morphometric analyses of spur dimensions, mostly conducted to explore the details of inferred selection pressure, generally reveal bell-shaped curves characteristic of polygenic control (Nilsson, 1983). Even the neat crossing experiments of Prazmo (1965) using *Aquilegia* have recently come under fire as a result of a phylogenetic re-assessment of the group. Recent

phylogenetic analyses indicate that the lack of floral spurs in *A. ecalcarata* can be interpreted as a loss (Hodges, 1997). Rather than indicating that one or two key genes are necessary for spur development, it now appears more likely that *A. ecalcarata* has lost its spurs, possibly as a result of a mutation in a gene early in the spur development pathway, thereby appearing as though a single gene was required for the trait. The evolution of spurs in *Aquilegia* may therefore have required a number of mutations generating an entire pathway necessary for the development of spurs (Tucker & Hodges, 2005).

Experimental evidence for and against simple genetic regulation of spurs is contradictory. Despite this background knowledge, the identity and number of spur determining gene(s) remain uncertain.

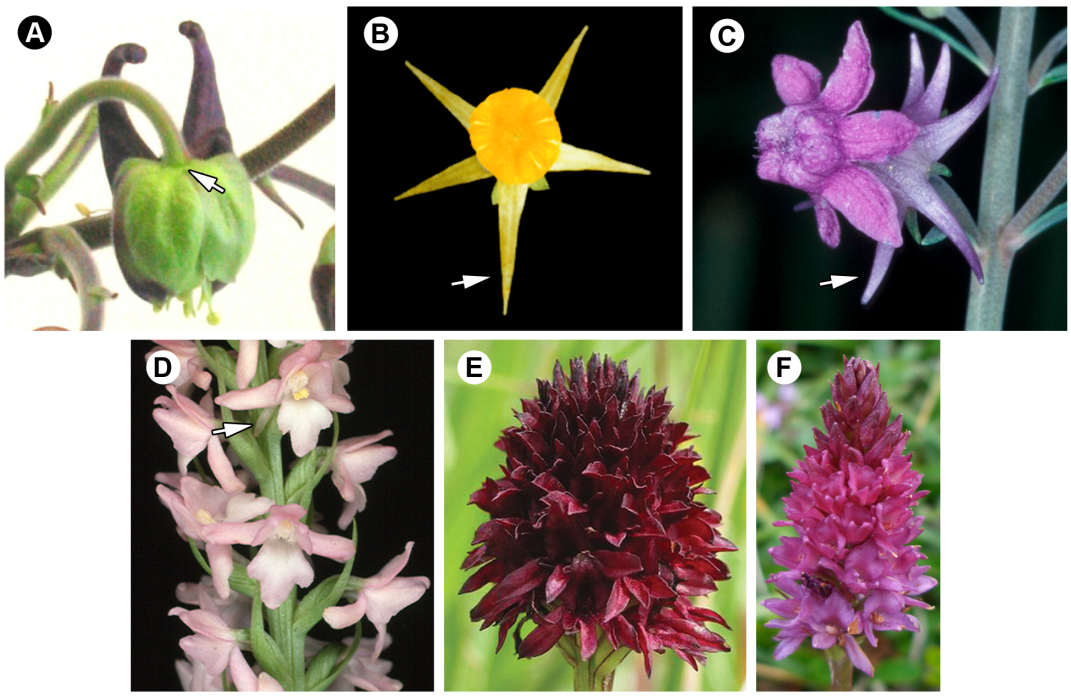


Figure 1.3: Nectar spur development is influenced by ABC and floral symmetry genes
A. Combined Viral Induced Gene Silencing (VIGS) of the *Aquilegia vulgaris* B-function gene *AqvPistillata* and the pigment gene *AqvAnthocyanidin Synthase* (used as a visual indicator of silencing) generates apetaloid flowers lacking nectar spurs (Kramer *et al.* (2007); arrow indicates the expected position of the nectar spur). B, C. Peloric mutants of (B) *Linaria vulgaris* and (C) *L. purpurea* with five-spurred actinomorphic flowers, photos: (B) Cubas *et al.* (1999), (C) P. J. Rudall (unpub). D-F. Hybridisation between the medium spur-length orchid *G. odratissima* (D) and its closely related short-spurred relative *G. rhellicanii* (E) generate hybrid progeny (F) with intermediate floral characteristics including spur length. Arrows indicate spurs, unless otherwise stated.

1.3.4 Snapdragon mutants indicate a role for homeobox genes in nectar spur development

Two independent snapdragon mutants; *Hirzina-d153* (*Hirz-d153*) and *Invaginata-d1* (*Ina-d1*), have been identified from *A. majus* with novel outgrowths on the petals resembling the floral nectar spurs of closely related species such as *Linaria vulgaris* (Golz *et al.*, 2002). All relatives of *A. majus* in the tribe Antirrhineae have tubular corollas and most have a high degree of floral zygomorphy, indicated most notably by the presence of a floral nectar spur on the ventral part of the corolla tube (Sutton, 1988). The garden snapdragon lacks the floral nectar spurs present amongst some of its close relatives, e.g. *L. vulgaris* (Almeida *et al.*, 1997). Therefore, the appearance of spur-like structures in mutants of *A. majus* (Figure 1.4) provides an opportunity to investigate the evolution and development of floral nectar spurs among its close relatives.

Homozygous *Hirz-d153* plants have several floral defects. The united part of the petal tube is reduced in length, shifting the positions at which the petal lobes diverge from the proximal region of the petal tube. In addition, tubular outgrowths form late in floral ontogeny from the ventral part of the corolla tube (Figure 1.4). The ectopic outgrowth has similar micromorphology to the wild type corolla tube but with a reversed polarity of tissues, suggesting that the tubular outgrowths are ectopic petal tubes (Golz *et al.*, 2002). Flowers of the *Ina-d1* mutant plants are much the same but the severity of the phenotype is more variable, less frequently forming ectopic petal tubes (Golz *et al.*, 2002).

In addition to this unique floral phenotype, homozygous *Hirz-d153* mutant plants have altered leaf morphology (Figure 1.4). The leaves are smaller and rounder than wild type and frequently develop ectopic trichomes and midribs. The lamina of the leaf is also often buckled as a result of excessive cell proliferation in the central and proximal regions. Golz *et al.* (2002) recognised that this leaf phenotype is reminiscent of constitutive expression of class 1 KNOX homeodomain transcription factors in transgenic tobacco plants (Sinha *et al.*, 1993).

Subsequent genetic analysis has revealed that the *Hirz-d153* and *Ina-d1* mutant phenotypes are due to transposon insertions in two genes encoding highly similar class 1 KNOX homeobox transcription factors, with high sequence similarity to the *SHOOT MERISTEMLESS* (*STM*) gene of *Arabidopsis thaliana*. They were subsequently called *AmSTM1* and *AmSTM2*. The *AmSTM1* allele of *Hirz-d153* plants was found to contain a Tam1 transposon insertion in the first intron of the gene, whilst a Tam3 transposon insertion was found in the 5′UTR of *AmSTM2* in *Ina-d1* mutants. Removing the transposon in each case by crossing the *Hirz-d153* and *Ina-d1* mutants into transposon active

backgrounds generated progeny with wild type flowers. As such *AmSTM1* was established as the causative allele for the *Hirz-d153* phenotype and *AmSTM2* for that of the *Ina-d1* mutants (subsequently referred to as *HIRZINA* (*HIRZ*) and *INVAGINATA* (*INA*) respectively).

Whilst the presence of the Tam1 and Tam3 transposons did not affect the coding sequence of either gene, or the mRNA products generated, their presence altered the spatio-temporal patterns of *HIRZ* and *INA* expression, which is normally confined to the shoot apical meristem (SAM; section 1.5.1) (Golz *et al.*, 2002). The mutant phenotypes of *Hirz-d153* and *Ina-d1* plants were interpreted as a result of neomorphic alterations of *HIRZ* and *INA* expression in developing organs (Golz *et al.*, 2002) to promote a novel axis of growth. Further genetic analysis of the *Hirz-d153* mutant also showed that the novel axis of growth generated by ectopic expression of *HIRZ* was strongly dependent on genetic factors influencing floral symmetry. In *rad, Hirz-d153* double mutants, multiple ectopic petal tubes were generated (Golz *et al.*, 2002), producing a flower resembling the peloric mutants of *L. vulgaris* (Figures 1.3 1.4). Although many of the phenotypic features described in the *Hirz-d153* and *Ina-d1* mutants are not uncommon in relation to KNOX, these mutants are the first identified that implicate class 1 *KNOX* genes in the development of novel axes of growth.

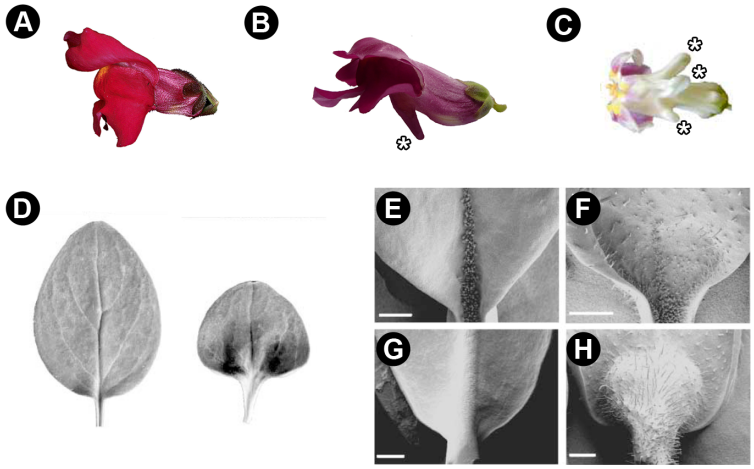


Figure 1.4: Ectopic petal tubes in *Antirrhinum majus* mutants indicate a role for homeobox genes in nectar spur development
A. Wild type *A. majus* flower. B. The *Hirz-d153* mutant of *A. majus* with an ectopic petal tube on the ventral part of the corolla tube. C. In a *rad, Hirz-d153* double mutant additional ectopic petal tubes are generated similar to those in peloric mutants of *L. vulgaris* (Figure 1.3). D. Mature wild type *A. majus* (left) versus *Hirz-d153* mutant (right) leaves. Leaves from the *Hirz-d153* mutant are smaller and rounder than wild type often with a buckled lamina. E-H. Compared to wild-type (E, G), *Hirz-d153* mutant leaves develop ectopic trichomes on both the dorsal (F) and ventral (H) surfaces. Images: B-H, from Golz *et al.* (2002). Asterisks indicate ectopic petal tubes.

1.4 Plant homeobox genes

KNOX genes in plants are members of the homeobox transcription factor family, all of which encode proteins with a highly evolutionarily conserved homeodomain (HD). The HD is encoded by a 180bp consensus DNA sequence, the homeobox, present in a number of genes initially discovered by screens of homeotic mutants in *Drosophila melanogaster* (McGinnis *et al.*, 1984; Scott & Weiner, 1984) and subsequently shown to be present in evolutionarily distant organisms including plants, animals and fungi (Derelle *et al.*, 2007). The homeodomain is a short 60 amino acid sequence with a helix-loop-helix-turn-helix structure generating a distinct amphipathic α-helical secondary structure that forms a tight globular conformation with a hydrophobic core (Billeter *et al.*, 1993; Otting *et al.*, 1990; Qian *et al.*, 1989). A conserved amino acid motif, WF**X**N, in helix 3 of the HD is responsible for associating with the DNA of downstream transcriptional targets, in particular the amino acid residue X is critical to the DNA binding specificity of helix 3 (Treisman *et al.*, 1989). As a result helix 3 is often referred to as the 'recognition helix' (Gehring *et al.*, 1994; Scott *et al.*, 1989) and is required for recognition of a degenerate target DNA-binding motif (TGACAGG/CT) by fitting into the major groove of DNA (Figure 1.5), thereby affecting the transcription of downstream targets. Therefore, in order to promote specific activation of downstream targets additional interacting factors are required (see section 1.4.2).

Homeobox genes play vital roles in the control of a diverse array of cell and developmental processes in animals. Famous examples from insects include the *Antennapedia* (Figure 1.5) and *Bithorax* homeobox genes that determine segment identity by promoting the development of appropriate anatomical features within particular segmental or parasegmental domains of the body (Lawrence & Morata, 1994). In fact, changes in the expression domains of insect homeobox genes over evolutionary timescales have been shown to be integral to the evolution of arthropod diversity (Hughes & Kaufman, 2002). Vertebrate homeobox genes (HOX) are also fundamental to a number of key developmental patterning processes such as the development of the central nervous system and axial skeleton (Manak & Scott, 1994) among others. For an excellent review of Homeobox genes in general see Gehring (1998).

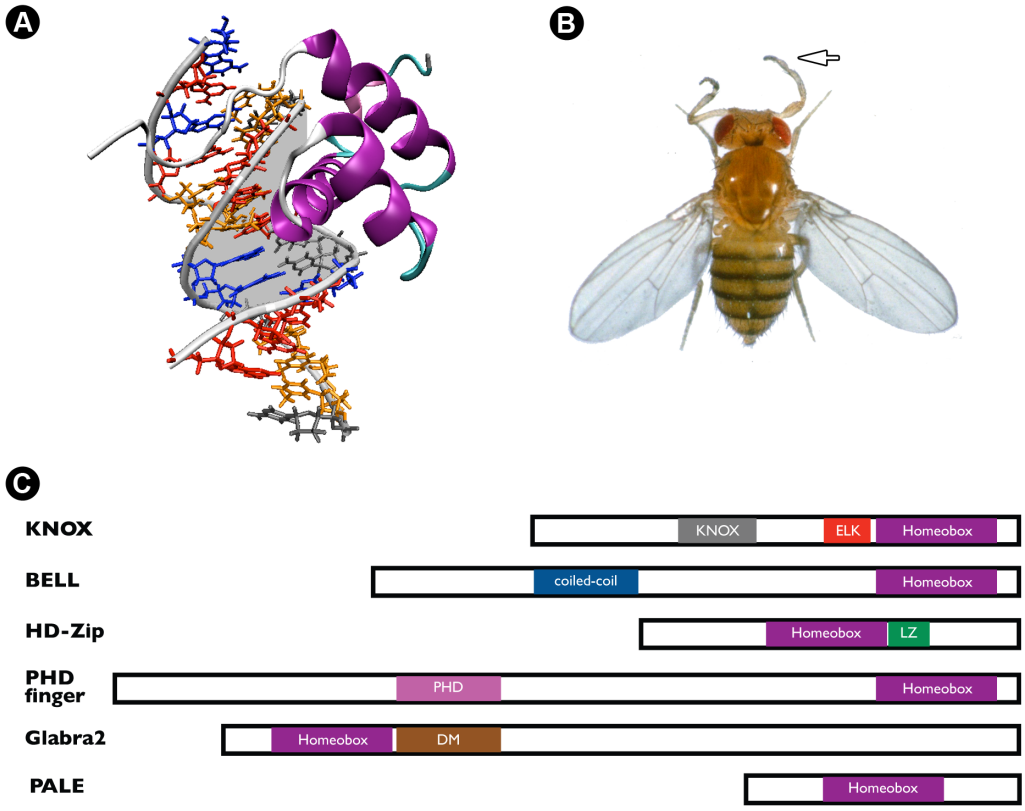


Figure 1.5: Homeobox genes in plants and animals
A. The Antennapedia homeodomain protein from *Drosophila melanogaster* bound to a DNA fragment, illustrating the binding interactions of the recognition helix (central, purple) and unstructured N-terminus with the major (grey) and minor grooves of DNA. Created from PDB entry 1AHD using the freely available visualization and analysis package Visual Molecular Dynamics (VMD). B. The *Antennapedia* mutant of *Drosophila melanogaster* in which the antenna are homeotically converted to legs (arrow), photo: M. Scott. C. Schematic representation of members from each family of plant homeodomain proteins. Conserved motifs are shown: homeodomain, purple; KNOX domain, grey; ELK domain, red; Coiled-coil domain, blue; Leucine-Zipper domain, green; PHD-finger domain, pink; Dimerization motif (DM), brown. Adapted from Chan *et al.* (1998).

1.4.1 *KNOX* genes: a unique family of plant homeobox genes

Plant homeobox genes have only recently been discovered. The first plant homeobox gene was isolated from a gain-of-function mutant of maize by transposon tagging (Vollbrecht *et al.*, 1991), and named *Knotted1* due to the presence of 'knots' along the lateral veins of the leaf blades of maize *Knotted1* mutants. Although considerably less well characterised, plant homeobox genes are also involved in orchestrating a variety of key developmental processes.

Plant homeodomain proteins fall into seven families based on differences in their DNA coding sequence, domain structures and expression patterns (Figure 1.5); these include KNOX and BEL, belonging to the TALE superclass (Bürglin, 1997), ZM-HOX, HAT1, HAT2, ATHB8, and GL2. HAT1, HAT2, ATHB8, and GL2 are all characterised by a leucine zipper motif (Ruberti *et al.*, 1991) and have been successively renamed HD-ZIP I, HD-ZIP II, HD-ZIP III and HD-ZIP IV respectively. Chan *et al.* (1998) proposed an alternative classification into five groups (HD-ZIP, GLABRA, KNOTTED, PHD, and BEL). More extensive genome-wide analyses from flowering plants, moss, *Selaginella*, unicellular green algae, and red algae (Mukherjee *et al.*, 2009) currently recognise 14 distinct classes including, HD-ZIP I to IV, BEL, KNOX, PLINC, WOX, PHD, DDT, NDX, LD, SAWADEE and PINTOX genes. Irrespective of their precise classification, members from each family are present in angiosperms, *Selaginella* and moss, suggesting that the plant homeobox gene families diverged before the last common ancestor of moss and vascular plants.

Over the last 15 years numerous members of each plant homeobox gene family have been identified and shown to play key roles in a diverse array of plant developmental processes such as photomorphogenesis, vascular development, defense gene regulation and trichome formation. Here, however we are concerned only with the *KNOX* gene family (*Knotted1-like homeobox*), named after its founding member, *Knotted1* from maize.

1.4.2 Domain structure of KNOX proteins

All KNOX proteins possess a C-terminal homeodomain carrying the WFXN motif in helix 3 (Treisman *et al.*, 1989). They are atypical homeodomain proteins in that they have three extra amino acids between helix 1 and helix 2, as such they belong to the TALE (Three Amino acid Loop Extension) superclass of homeodomain transcription factors, along with members of the BEL family of plant homeodomain proteins plus several homeodomain families from animals (Bertolino *et al.*, 1995; Bürglin, 1997; Mukherjee & Bürglin, 2007;

Mukherjee *et al.*, 2009). Both atypical TALE, and typical homeodomain genes were present in the eukaryote ancestor of plants, animals and fungi (Derelle *et al.*, 2007).

Plant KNOX proteins are characterised by an N-terminal KNOX domain which is composed of two neighbouring KNOX subdomains, the N-terminal most KNOX1 and adjoining KNOX2 domains respectively (Figure 1.6). The KNOX domain is closely related to the myeloid ecotropic viral integration site (MEIS domain) of TALE homeodomain transcription factors in humans. The KNOX and MEIS domains of both protein families are highly similar, as such it has been suggested that the KNOX and MEIS domains have been derived from the same common ancestral domain, the MEINOX domain, prior to the divergence of plants from fungi and animals (Bürglin, 1997, 1998; Mukherjee & Bürglin, 2007). As such the KNOX and MEIS domains are often referred to as the MEINOX domain. In plants, the KNOX domain is necessary and sufficient for interaction with the BEL family of TALE homeodomain proteins in a DNA-independent manner (Figure 1.6), forming KNOX-BEL heterodimer transcriptional units (Bellaoui *et al.*, 2001; Chen *et al.*, 2003; Müller *et al.*, 2001; Smith *et al.*, 2002), although alternative roles have also been proposed for these complexes (see section 1.5.3.4).

Each *KNOX* gene product is believed to interact with only a small subset of BEL proteins in specifying targeted genes (Smith & Hake, 2003; Smith *et al.*, 2002). Both BEL and KNOX proteins possess the DNA-binding WF*I*N amino acid motif (Figure 1.6) and recognise the TGACAGG/CT DNA-binding motif in downstream transcriptional targets. The formation of BEL-KNOX heterodimers considerably increases the affinity of KNOX-BEL transcriptional complexes to target DNA sequences (Viola & Gonzalez, 2006). The association of KNOX and BEL is analogous to the functioning of MEIS proteins in humans which also interact specifically with another class of TALE homeodomain proteins, the PBC proteins (Joshi *et al.*, 1997; Knoepfler *et al.*, 1997; Ryoo *et al.*, 1999). KNOX proteins can also form transcriptionally active homodimers via the KNOX and homeodomain, though these are less effective at promoting transcription (Müller *et al.*, 2001; Nagasaki *et al.*, 2001a). In addition to promoting transcription, KNOX-BEL heterodimers may also be linked to nuclear localisation of KNOX (Cole *et al.*, 2006; Rutjens *et al.*, 2009).

Between the N-terminal MEINOX domain and C-terminal homeodomain of plant KNOX proteins is a central domain uncommon in other homeobox proteins (Figure 1.6). This region is enriched in Proline (P), Glutamic Acid (E), Serine (S) and, Threonine (T). Appropriately it is referred to as the PEST domain (Vollbrecht *et al.*, 1991). PEST sequences are purported to be involved in post-transcriptional control of KNOX proteins, targeting them for degradation (Rechsteiner & Rogers, 1996). Adjacent to the N-terminus

of the homeodomain is a second motif highly conserved in each member of the TALE superclass, the ELK domain, named after the first three amino acids; Glutamic Acid (E), Leucine (L) and Lysine (K). The ELK domain is thought to be involved in mediating protein-protein interactions and nuclear localisation of KNOX (Meisel & Lam, 1996; Mushegian & Koonin, 1996). For a detailed review of KNOX protein domain structure see Chan *et al.* (1998) and Hake *et al.* (2004).

1.4.3 *KNOX* genes are divided into two (or three) distinct classes

The *KNOX* gene family is divided into two phylogenetically distinct and well-supported classes based on subtle differences in sequence homology, expression patterns, the positions of conserved introns and the function of the encoded proteins (Kerstetter *et al.*, 1994). The distinction between class 1 and class 2 *KNOX* genes has been recovered in a number of subsequent phylogenetic analyses of *KNOX* genes (Figure 1.6); Hake *et al.* (2004); Jouannic *et al.* (2007); Magnani & Hake (2008); Mukherjee *et al.* (2009); Reiser *et al.* (2000); Sano *et al.* (2005). Recently the discovery of a functional *KNOX* gene lacking the homeodomain in *Arabidopsis* and poplar, *KNATM*, suggests that there may also be a third distinct class of *KNOX* genes (Magnani & Hake, 2008). The existence of this putative third class suggests that the MEINOX and KNOX domains may have evolved independently of one another (Magnani & Hake, 2008). Further characterisation and isolation of KNATM-like proteins from other seed plants is required to establish whether this is a truly novel class of *KNOX* genes or whether they are merely pseudogenes.

The homeodomain of class 1 and class 2 KNOX proteins show a high level of sequence identity, particularly in the recognition helix (helix 3), suggesting that they interact with similar DNA sequences. However, outside of the homeodomain, class 1 and class 2 *KNOX* genes vary significantly. This suggests that although the main contacts with the DNA backbone may be similar, there are likely to be significant differences in DNA binding and their protein partners (Chan *et al.*, 1998).

Class 1 and class 2 *KNOX* genes have now been isolated from all of the well-established genetic models including maize (Vollbrecht *et al.*, 1991), *Arabidopsis* (Chuck *et al.*, 1996; Lincoln *et al.*, 1994; Long *et al.*, 1996), barley (Müller *et al.*, 1995), rice (Matsuoka *et al.*, 1993), tomato (Hareven *et al.*, 1996; Janssen *et al.*, 1998b), snapdragon (Golz *et al.*, 2002) and petunia (Stuurman *et al.*, 2002). *KNOX* genes are also being increasingly identified in non-model plants including apples (Watillon *et al.*, 1997) and orchids (Yu *et al.*, 2000).

In addition to angiosperms, a number of *KNOX* genes have also been identified in more basal plant lineages such as the lycophytes (Harrison *et al.*, 2005b), gymnosperms (Sundås-Larsson *et al.*, 1998), ferns (Sano *et al.*, 2005), mosses (Champagne & Ashton,

2001), green (Serikawa & Mandoli, 1999) and red algae (Mukherjee *et al.*, 2009), suggesting that they are amongst the oldest plant homeobox gene families (Mukherjee *et al.*, 2009).

1.4.4 Class 1 and 2 *KNOX* genes arose by an ancient duplication event

Current phylogenetic analyses (Bharathan *et al.*, 1999; Champagne & Ashton, 2001; Jouannic *et al.*, 2007; Magnani & Hake, 2008; Mukherjee *et al.*, 2009; Reiser *et al.*, 2000; Sano *et al.*, 2005; Serikawa & Mandoli, 1999) firmly support the monophyly of class 1 and class 2 *KNOX* genes (Figure 1.6). Representatives of both classes can be found among angiosperms, gymnosperms, ferns and bryophytes. This strongly suggests that the last common ancestor of ferns and seed plants, the Trimerophytes (Stewart & Rothwell, 1993), had both class 1 and class 2 *KNOX* genes; indicating that these two classes diverged before the split between seed plants and ferns around 400 MYA.

KNOX genes isolated from green algae such as *Acetabularia acetabulum* (*AaKNOX1*), and a more recently identified *KNOX* gene isolated through genome sequencing of *Chlamydomonas*, are consistently recovered at the base of *KNOX* gene phylogenies (Champagne & Ashton, 2001; Jouannic *et al.*, 2007; Sano *et al.*, 2005; Serikawa & Mandoli, 1999). Individual algal *KNOX* genes possess features characteristic of both class 1 and class 2 *KNOX* genes, indicating that class 1 and class 2 *KNOX* genes may have diverged only shortly after the evolution of land plants from green algae 500 MYA. If this is the case the ancient duplication of class 1 and class 2 *KNOX* genes, and subsequent diversification in their function, may have occurred concomitantly with the evolution of complex plant bodies (Sano *et al.*, 2005).

1.4.5 Angiosperm class 1 *KNOX* genes appear to have duplicated three times

Angiosperm class 1 *KNOX* genes are divided into two distinct monophyletic clades, class 1a and 1b, both of which are further divided into two orthology groups each defined by a well-characterised *KNOX* gene member (Bharathan *et al.*, 1999; Reiser *et al.*, 2000). Class 1a *KNOX* genes consist of the *STM* and *KNAT1* groups, both defined by *Arabidopsis KNOX* genes. Class 1b *KNOX* genes consist of the *KNAT2* and *OSH6* groups, defined by *Arabidopsis* and rice *KNOX* genes respectively. Both monocot and eudicot sequences are represented in each subclass and subsequent orthology group. This strongly suggests that there have been three ancestral duplication events among angiosperm *KNOX* genes (Jouannic *et al.*, 2007). In the first duplication event an ancestral class 1 *KNOX* gene

was duplicated giving rise to class 1a and class 1b *KNOX* genes. The second and third duplication events subsequently generated the *STM, KNAT1* and *KNAT2, OSH6* groups from ancestral class 1a and class 1b *KNOX* genes respectively. According to the phylogenetic reconstruction of Jouannic *et al.* (2007), these duplication events occurred prior to the divergence of monocot and eudicot lineages (Figure 1.6).

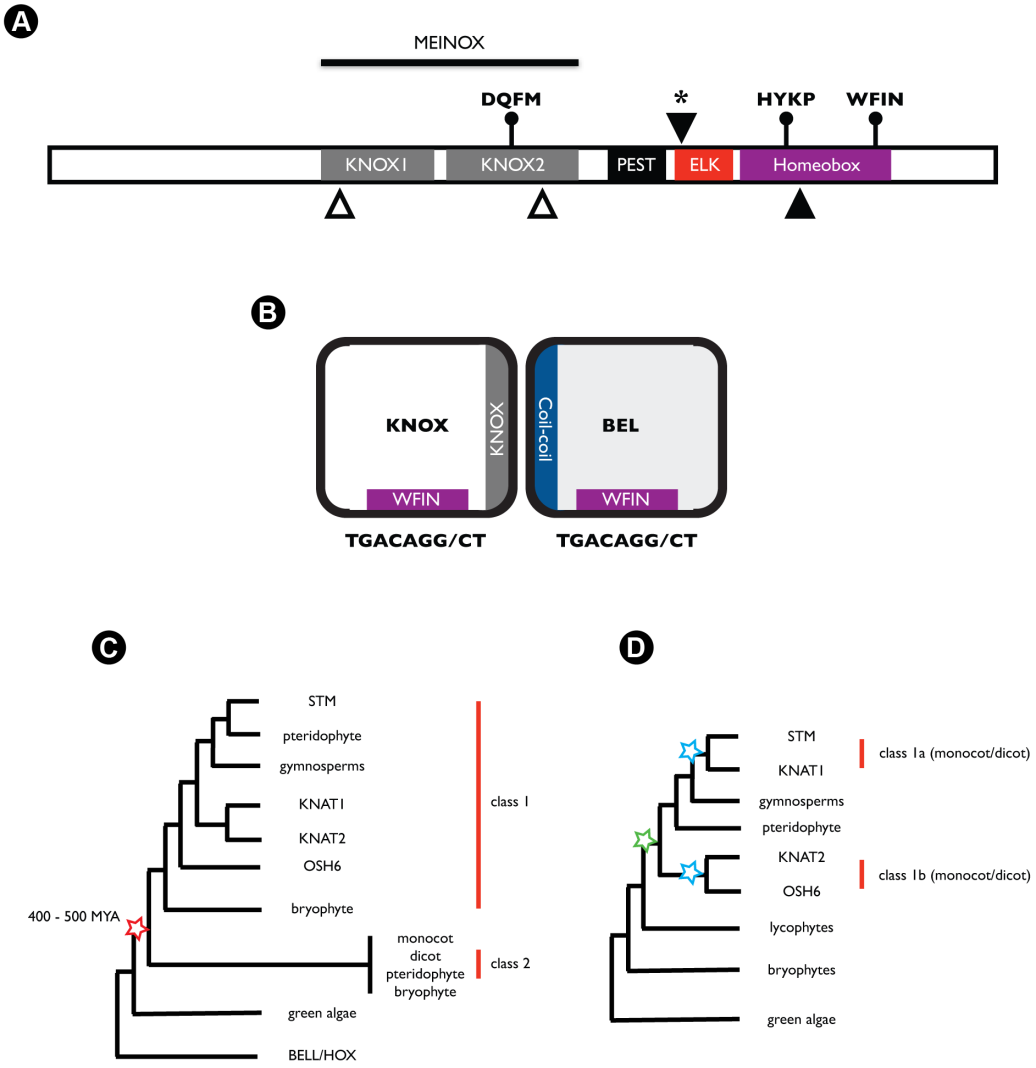


Figure 1.6: KNOX protein domain structure and evolution
A. Generalised *KNOX* gene schematic showing the conserved protein domains and amino acid motifs DQFM, HYKP and WFIN. Typical intron positions are indicated by triangles, open triangles indicate typical intron positions in class 1 *KNOX* genes. Conserved intron positions are indicated by closed triangles, the asterisk indicates the conserved intron position in class 2 *KNOX* genes. Conserved protein coding domains are labelled according to the scheme presented in Figure 1.5 with the addition of the PEST domain, black. B. BEL and KNOX proteins interact through their MEINOX domain forming heterodimer transcriptional units (modified from Hake *et al.* (2004). C, D. Cartoons representing the phylogenetic analyses of Sano *et al.* (2005) (C) and Jouannic *et al.* (2007) (D) showing the relationships between algal, bryophyte, pteridophyte, gymnosperm and angiosperm *KNOX* genes. The duplication of class 1 and class 2 *KNOX* genes occurred 400-500 MYA (indicated by a red star), subsequent duplication events common to monocot and eudicot class 1 *KNOX* genes are indicated by green (class 1a/1b) and blue stars.

1.4.6 Class 1 *KNOX* genes are present as multigene families in angiosperms

Class 1 *KNOX* genes are present as a multigene family in most plant species sampled. Five class 1 *KNOX* genes have been cloned from polyploid tobacco (Nishimura *et al.*, 1999), nine from the polyploid maize (Reiser *et al.*, 2000), five from rice (Sentoku *et al.*, 1999), four from *Arabidopsis* (Dean *et al.*, 2004), four from tomato (Reiser *et al.*, 2000) and five from snapdragon (Golz *et al.*, 2002). Further sampling in light of the recent availability of genome sequences for many of these models, has identified eight class 1 *KNOX* genes in *Arabidopsis*, 15 in poplar, 12 in rice (plus one pseudogene), 13 in maize, 5 (and 1 pseudogene) in *Selaginella*, 5 in moss, and 1 each in green and red algae (Mukherjee *et al.*, 2009).

Although sampling is incomplete for many of the species sampled to date, the relationships between these genes are well understood. Remarkably, each species tends to have at least one member of each of the four major orthology groups of class 1 *KNOX* genes (STM, KNAT1, KNAT2 and OSH6), in addition to one or more class 2 *KNOX* genes. Whilst multiple *KNOX* genes may provide functional redundancy in some cases (Byrne *et al.*, 2002), the phylogenetic distinctiveness between different groups of *KNOX* genes, their persistence over significant periods of evolutionary time and their differing expression patterns strongly suggests that *KNOX* genes have evolved by subfunctionilisation or through the acquisition of new roles in plant morphogenesis.

1.5 Putative roles for Class 1 *KNOX* genes in plant development

KNOX genes are fundamental to plant morphogenesis. Traditionally, class 1 *KNOX* genes have been viewed as being predominantly concerned with the maintenance of the shoot apical meristem (SAM), maintaining indeterminate cell fate and suppressing differentiation (Endrizzi *et al.*, 1996; Kerstetter *et al.*, 1997; Long *et al.*, 1996; Scofield & Murray, 2006; Vollbrecht *et al.*, 2000), through the action of cytokinins and gibberellins (Gordon *et al.*, 2009; Jasinski *et al.*, 2005; Sakamoto *et al.*, 2001). Maintenance of this group of self-renewing cells in the SAM is essential, as the derivatives of these cells give rise to all of the above ground organs of the plant (Veit, 2006). The fundamental role of *KNOX* genes in the SAM is demonstrated powerfully by the observation that down regulation of *KNOX* expression is required in the SAM prior to the differentiation of founder cells and

the subsequent development of leaf primordia (Jackson *et al.*, 1994; Smith *et al.*, 1992). For detailed reviews see Reiser *et al.* (2000), Hake *et al.* (2004) and Hay *et al.* (2009).

In direct contrast with the extensive characterisation of class 1 *KNOX* genes, the functions of class 2 *KNOX* genes are less well understood due to considerably lower experimental investment, largely as a result of a lack of available mutants and more ubiquitous expression patterns (Hake *et al.*, 2004; Reiser *et al.*, 2000; Scofield & Murray, 2006). Despite this, some authors have suggested that class 2 *KNOX* genes may function in late stages of plant organogenesis (Serikawa *et al.*, 1997) and root development (Truernit *et al.*, 2007).

1.5.1 Class 1 ***KNOX*** genes function in meristem maintenance

Evidence for this role is strong and predominantly comes from careful analysis of *KNOX* loss-of-function mutants such as *SHOOT MERISTEMLESS* (STM) of *Arabidopsis* (Barton & Poethig, 1993; Long *et al.*, 1996). Plants carrying the *stm-1* allele in a *Landsberg erecta* background produce cotyledons but no further components of the shoot system, strongly suggesting that *STM* is required to maintain and/or initiate the shoot apical meristem (SAM). Similar recessive mutants have since been discovered. In direct contrast to the original gain-of-function *Knotted1* mutant in maize, for example, recessive *kn1* mutants form only a limited shoot (Vollbrecht *et al.*, 2000) just like *stm-1 Arabidopsis* plants (Figure 1.7).

The precise roles of *KNOX* genes in the SAM are complex, and not only affect the maintenance of the apical meristem but a number of developmental processes related to the specific functions of the subdomains of the SAM. Often *KNOX* genes are expressed in specific subdomains of the SAM, *Arabidopsis KNOX* gene *STM*, maize *KN1*, rice *OSH1* and tobacco *NTH1* and *NTH15* are expressed in the central *corpus* zone of the meristem consistent with putative roles in meristem maintenance. By contrast, *Arabidopsis KNAT1*, tobacco *NTH20*, rice *OSH15*, *Arabidopsis BP* and maize *RS1* are predominantly expressed in the peripheral zone of the meristem, consistent with roles in internode elongation and lignin deposition (Nishimura *et al.*, 1999; Sentoku *et al.*, 1999), reviewed in Hake *et al.* (2004). More details of this are presented in Chapter 4.

1.5.2 Extra-meristematic functions of class 1 ***KNOX*** genes

In recent years a diverse array of additional roles for class 1 *KNOX* genes have been identified, some of which are related to the SAM, such as the development of vascula-

ture/internode growth (Smith & Hake, 2003) and lignin biosynthesis/deposition (Groover *et al.*, 2006; Hertzberg *et al.*, 2001; Mele, 2003).

However, a number of divergent roles have also been presented, most often in complex leaf morphogenesis, such as proximo-distal patterning and leaf serration, but specifically in the development of lobed and/or compound leaves. In plants with simple leaves such as *Arabidopsis* and snapdragon, *KNOX* gene expression is deactivated in the leaf primordium; however, in many plants with compound leaves such as tomato (Figure 1.7) and *Cardamine hirsuta*, *KNOX* expression is re-activated in leaves after leaf initiation where it facilitates leaflet formation by maintaining an indeterminate environment in late stages of leaf morphogenesis (Hareven *et al.*, 1996; Hay & Tsiantis, 2006; Shani *et al.*, 2009). Similar observations by a large number of authors in diverse taxa including the Brassicaceae, Solanaceae and the Legumes have confirmed these findings (Bharathan *et al.*, 2002; Champagne *et al.*, 2007; Chen *et al.*, 1997; Harrison *et al.*, 2005a; Hay & Tsiantis, 2006; Kimura *et al.*, 2008; Kumar *et al.*, 2007; Müller *et al.*, 2006; Parnis *et al.*, 1997; Shani *et al.*, 2009). It should be noted that *KNOX*-independent mechanisms of complex leaf development have also been proposed for other taxa such as pea, in which a pea ortholog of *Arabidopsis LEAFY* (*UNIFOLIATA*) converts leaves from simple to dissected (Gourlay *et al.*, 2000). Much of the evidence for such a broad phylogenetic role of *KNOX* genes in the development and evolution of complex leaf morphology comes from gain-of-function mutants and transgenic experiments, in which *KNOX* genes are constitutively expressed.

Despite well established roles for *KNOX* genes in apical meristem maintenance and complex leaf morphogenesis, altered expression of KNOX proteins can also have significant morphogenetic consequences in organs other than leaves. It has been known for some time that misexpression of the Barley ortholog of *KNOTTED* from maize results in a hooded floral phenotype, characterised by transformation of the awn of Barley flowers (a slender bristle-like structure found on the spikelets of many grasses) into a reiterative inflorescence structure (Müller *et al.*, 1995; Williams-Carrier *et al.*, 1997). Less well characterised roles have also been proposed for *KNOX* genes based largely on expression data. Additional roles for *KNOX* genes have been suggested in the development of the floral meristem (Long & Barton, 2000), carpels (Scofield *et al.*, 2007, 2008), lateral roots (Dean *et al.*, 2004), tubers (Chen *et al.*, 2003; Rosin *et al.*, 2003) and the development of novel axes of growth and subsequent floral organ parts such as nectar spurs (Golz *et al.*, 2002). More details of class 1 *KNOX* gene functions in plant morphogenesis are presented in Chapter 4.

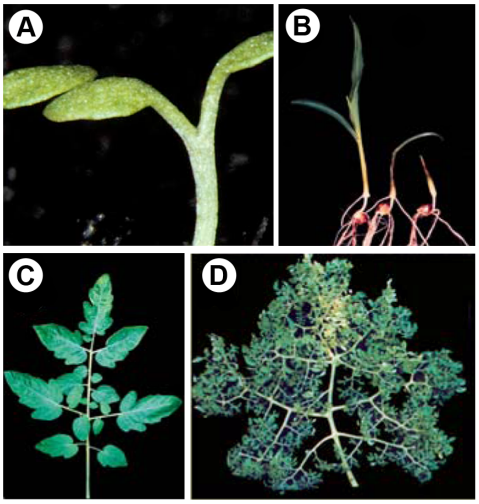


Figure 1.7: KNOX proteins function in meristem maintenance and compound leaf development

A. Extreme *stm-1* mutants of *Arabidopsis* lack a shoot apical meristem (SAM) and never develop non-seed leaves. B. Two-week-old maize seedlings. The wild-type (left) seedling has two visible leaves, *kn1-e1* mutants produce a normal root and a coleoptile but make no leaves (right) or only a single leaf (middle). C, D. Wild-type (C) versus super-compounded 35S::KN1 (D) tomato leaves. Images: B-D, from Hake *et al.* (2004).

1.5.3 Regulating *KNOX* activity

The range of phenotypes generated by aberrations in *KNOX* gene expression is strongly contextualised and can give rise to divergent morphologies depending upon the spatio-temporal pattern of expression and the developmental potential of the tissue they become expressed in (Tsiantis & Hay, 2003). In maize, the responsiveness of leaf tissue to ectopic *KNOX* expression changes with developmental stage (Muehlbauer *et al.*, 1999) and this is also found to be true for misexpression of *KNOX* genes from tomato in which the phenotype varies considerably with the location of ectopic expression (Chen *et al.*, 1997; Parnis *et al.*, 1997). The strongly contextualised nature of ectopic *KNOX* gene expression phenotypes makes us consider the factors that influence the activity of *KNOX* genes and suggests that the fundamental properties of KNOX proteins may be key to the evolution of a diverse range of complex vegetative and floral traits.

In order to carry out such complex roles in plant morphogenesis it is likely that class 1 *KNOX* genes are tightly integrated into complex gene regulatory networks including other transcription factors which may control, or themselves be controlled by, class 1 *KNOX* genes, such as the ABC or floral symmetry breaking genes. Often this is likely to occur with the help of interacting partners (e.g. BEL proteins), and in relation to plant growth substances. Much of what is known of the integration of class 1 *KNOX* genes into regulatory pathways is related to meristem maintenance and the development of complex leaf morphologies. The identity of precise targets of *KNOX* genes are only recently emerging and remain largely uncharacterised (Bolduc & Hake, 2009; Hake *et al.*, 2004; Hay *et al.*, 2009).

1.5.3.1 Intercellular KNOX protein trafficking

One possible control point of KNOX activity is shown by studies of KNOX protein transport. The maize KNOX protein KNOTTED is capable of moving intercellularly via plasmodesmata, a property that may be integral to its role in meristem maintenance (Jackson *et al.*, 1994; Kim *et al.*, 2003; Lucas *et al.*, 1995). Altered patterns of KNOX protein movement may be integral to the phenotype generated. There are, however, many other factors that influence the context of *KNOX* gene expression.

1.5.3.2 *KNOX* gene relationships with plant growth substances

In addition to KNOX protein movement, plant growth substances such as auxin, cytokinins and Gibberellic Acid (GA) have also been demonstrated to be involved in *KNOX*

gene dependent developmental processes. The roles of plant growth substances in morphogenesis are well-characterised. It is therefore not surprising that genes implicated in plant morphogenesis are likely to interact with growth substances that have been known to be of great significance for decades. Indeed, morphological phenotypes observed in both naturally occurring mutants and transgenic plants with altered *KNOX* gene expression are very similar to plants that exhibit imbalances in plant growth substances such as cytokinins, GA and auxin, reviewed by Hay *et al.* (2004); Shani *et al.* (2006).

Many authors have proposed a role for polar auxin transport and auxin gradients in leaf initiation in the SAM. Down regulation of class 1 *KNOX* genes in the founder cells of incipient leaf primordia is thought to require localised auxin maxima (Champagne & Sinha, 2004; Hay *et al.*, 2004; Piazza *et al.*, 2005). Furthermore, mutations that compromise auxin signalling, e.g. loss of function mutations of the auxin efflux transporter PINFORMED1 (PIN1), result in ectopic expression of the *KNOX* gene *BP* in leaves (Hay & Tsiantis, 2006).

Cytokinins are plant growth regulators critical for meristem maintenance, acting as positive regulators of cell division (Greenboim-Wainberg *et al.*, 2005; Shani *et al.*, 2006). As such cytokinins are also thought to be related to class 1 *KNOX* gene activity. Transgenic lines constitutively expressing *KNOX* genes and naturally occurring gain-of-function mutants often have delayed senescence, lack apical dominance and generate shoots on the surface of leaves (for more details see Chapter 7). Consistent with these phenotypes increased cytokinin levels are often observed in transgenic plants and mutants with higher and more widespread than normal expression of *KNOX* (Frugis *et al.*, 2001; Tamaoki *et al.*, 1997). *KNOX* genes have been shown to directly activate cytokinin biosynthesis by promoting transcription of the cytokinin biosynthetic gene isopentyl transferase-7 in transgenic *Arabidopsis* plants with constitutive *KNOX* expression (Yanai *et al.*, 2005).

GAs are a class of plant growth regulators that regulate and integrate a wide range of plant growth and developmental processes and are often associated with cell expansion and differentiation (Thomas *et al.*, 2005). Cytokinins normally act antagonistically to GA (Greenboim-Wainberg *et al.*, 2005; Shani *et al.*, 2006), external application of GA can reverse the unusual leaf phenotypes observed in plants constitutively expressing *KNOX* genes (Hay *et al.*, 2002). In the most current model of meristem maintenance, *KNOX* genes are thought to act to minimise GA levels in the meristem (Hay *et al.*, 2002; Jasinski *et al.*, 2005; Shani *et al.*, 2006). Recently, the interplay between *KNOX* and GA has been examined by extensive mutant analyses in *Arabidopsis* (Hay *et al.*, 2002), tobacco (Sakamoto *et al.*, 2001), potato (Chen *et al.*, 2003), maize (Bolduc & Hake, 2009) and pea (Singh *et al.*, 2010). *Arabidopsis* plants constitutively expressing the *KNOX* genes

STM and *BREVIPEDICELLUS* (*BP*), but not other *Arabidopsis KNOX* genes, showed significantly lower levels of *GA20ox1* (GA20-oxidase) mRNA (Hay *et al.*, 2002), a gene involved directly in physiologically active GA biosynthesis, and normally excluded from the SAM. *KN1* homologs of tobacco and potato have also been shown to down regulate expression of this key GA biosynthetic gene (Chen *et al.*, 2003; Sakamoto *et al.*, 2001).

Most recently the class 1 *KNOX* gene *KNOTTED* from maize has been shown to drastically increase the expression of the GA catabolism gene *Ga2ox1* which removes bioactive GA (Bolduc & Hake, 2009). The level of *Ga2ox1* expression is highest in the SAM from which GA is excluded. The way in which KNOTTED achieves bioactive GA catabolism is by up regulating *Ga2ox1* as a result of direct binding of the KNOTTED protein to a cis-regulatory region in the first intron of the *ga2ox1* gene. Chen *et al.* (2003) proposed a similar mechanism for the down regulation of *GA20ox1* by the *KNOX* gene *POTH1* in potato. This suggests that interactions of KNOX proteins with their downstream targets in this way may be common (Bolduc & Hake, 2009). By contrast, in pea, vegetative expression of the gene *Ga2-oxidase2* (*PsGA2ox2*), a GA inactivating enzyme, causes an increase in vegetative *KNOX* gene expression and induces typical *KNOX* over expression vegetative leaf phenotypes (Singh *et al.*, 2010), strongly indicating that GA/*KNOX* signaling may occur bi-directionally. To date no further KNOX protein targets have been demonstrated to this level of detail.

These roles are consistent with current models of meristem maintenance in which KNOX is thought to minimise GA levels in the meristem and increase levels of cytokinins, that work antagonistically, thereby delaying differentiation (Hay *et al.*, 2002; Jasinski *et al.*, 2005; Shani *et al.*, 2006).

1.5.3.3 Positive regulators of *KNOX* gene expression

A number of genetic factors have also been identified that positively or negatively regulate the expression of *KNOX* genes. These factors have been isolated largely as a result of mutant screens in genetic models such as *Arabidopsis*, based on their phenotypic similarities to *KNOX* loss or gain-of-function mutants respectively.

There are few examples of positive regulators of *KNOX* expression described in the literature, many of which have been identified as a direct result of the *KNOX*-like loss-of-function phenotypes that they generate. *CUP SHAPED COTYLEDON* (*CUC*), *CUPULIFORMIS* (*CUP*) and *NO APICAL MERISTEM* (*NAM*) are all members of the NAC gene family of transcription factors. Loss-of-function mutants such as *cuc*, from *Arabidopsis*, have a phenotype similar to *stm* mutants in that they fail to make embryonic meristems, but have more dramatic cotyledon fusion (Aida *et al.*, 1997). Furthermore,

cuc1, cuc2 double mutants lack any STM expression at all (Aida *et al.*, 1999). Conversely, transgenic *Arabidopsis* constitutively expressing *CUC* under the Cauliflower Mosaic Virus 35S constitutive promoter (CaMV 35S), have enhanced expression of *STM* and associated *KNOX* constitutive expression phenotypes (Lenhard *et al.*, 2002). Consistent with these data, the *Cuc1* allele has been implicated in the establishment of the SAM by activating *STM* (Hibara *et al.*, 2003) possibly under the regulatory control of the micro RNA miR164, essential for normal embryonic, vegetative and floral development, which directly regulates *CUC1* (Mallory *et al.*, 2004).

1.5.3.4 Negative regulators of *KNOX* gene expression

A number of negative regulators of class 1 *KNOX* genes have been identified from mutants with phenotypes similar to *KNOX* gain-of-function or constitutively expressing transgenic plants. The orthologous genes *PHANTASTICA* (snapdragon), *ROUGH SHEATH2* (maize) and *ASYMMETRIC LEAVES1* (*Arabidopsis*) all encode MYB domain transcription factors expressed in lateral organ primordia from which they restrict *KNOX* expression (Byrne *et al.*, 2000; Timmermans *et al.*, 1999; Tsiantis *et al.*, 1999b; Waites *et al.*, 1998). Loss-of-function mutants of these three orthologs generate subtly different phenotypes. Mutants of *phan* suffer a loss of adaxial-abaxial leaf polarity, generating abaxialized leaves with occasional radial symmetry (Waites *et al.*, 1998). The *rs2* and *as1* mutants, by contrast, have proximo-distal leaf polarity defects with leaf phenotypes reminiscent of *KNOX* gain-of-function mutants in maize (Timmermans *et al.*, 1999; Tsiantis *et al.*, 1999b) and *Arabidopsis* (Byrne *et al.*, 2000; Ori *et al.*, 2000) respectively.

There are a large number of additional mutants that generate *KNOX* misexpression phenotypes and reveal aspects of *KNOX* regulation. *ASYMMETRIC LEAVES2* (*AS2*) of *Arabidopsis* is a member of the lateral organ boundaries (LOB) family of transcription factors (Iwakawa *et al.*, 2002), which also leads to loss of dorsoventrality in loss-of-function mutants (Lin *et al.*, 2003). The genes *YABBY3* and *FILAMENTOUS FLOWER* (*FIL*), are both members of the *YABBY* gene family (Siegfried *et al.*, 1999). Whilst they do not generate a leaf phenotype of their own, *yabby3, fil* double mutants cause the formation of ectopic meristems and misexpression of *STM*, *BP* and *KNAT2* (Kumaran, 2002; Sawa *et al.*, 1999; Siegfried *et al.*, 1999). In tomato, the *clausa* locus causes misexpression of the class 1 *KNOX* gene *LeT6/TKN2* in mature leaves, resulting in excessive fusion of organs and ectopic meristems (Avivi *et al.*, 2000). In maize the *semiphore* mutants missexpress the class 1 *KNOX* genes *RS1* and *GNARLY1* in the endosperm and leaves (Scanlon *et al.*, 2002). *RS1* and *GNARLY1* can accumulate to extremely high levels and can be seedling lethal in some backgrounds.

Some more revealing insights into *KNOX* gene regulation have recently been identified from continuing work in *Arabidopsis*. Phylogenetic footprinting has identified a conserved *cis* regulatory element (the *K-box*) responsible for the persistent repression of *STM* transcription in leaves following initial downregulation in the leaf primordium (Uchida *et al.*, 2007).

Epigenetic regulation of *KNOX* genes has also been proposed recently. Two mutants, *serrate* and *pickle* (Eshed *et al.*, 1999; Ogas *et al.*, 1999) that encode a chromatin re-modelling factor and C(2)H(2) type zinc-finger protein, respectively, have been implicated in gene regulation through chromatin modification (Prigge & Wagner, 2001), suggesting that chromatin configuration is important in the regulation of *KNOX* genes. Further evidence for an epigenetic mode of *KNOX* gene repression comes from additional studies using *Arabidopsis* which implicate members of the polycomb group of proteins, e.g. *CURLY LEAF* (*CLF*), *SWINGER* (*SWN*) and *FERTILISATION INDEPENDENT ENDOSPERM*, all of which have methyltransferase activity that can repress *KNOX* gene transcription by trimethylation of histone H3 at lysine 27 (H3K27me3) and subsequent chromatin re-modelling (Katz *et al.*, 2004; Schubert *et al.*, 2006; Xu & Shen, 2008).

In addition to their accepted roles as transcriptional partners for KNOX, specific BEL proteins have also been demonstrated to negatively regulate KNOX activity, for example the BEL1-like homeodomain (BLH-HD) genes *SAWTOOTH1* (*SAW1*) and *SAWTOOTH2* (*SAW2*) repress KNOX activity in *Arabidopsis* leaves (Kumar *et al.*, 2007). The KNATM family of KNOX proteins, which lack a homeodomain, may also selectively interact with BEL proteins, competitively inhibiting other KNOX proteins via the formation of inactive KNATM-BEL heterodimers (Kimura *et al.*, 2008; Magnani & Hake, 2008).

Despite a plethora of mutants, it is still unknown precisely how *KNOX* expression is initially down regulated in the SAM prior to the initiation of incipient leaf primordia. However, these recent advances provide a basic framework for understanding how *KNOX* genes control an increasing number of plant developmental processes.

1.6 Defining the role of *KNOX* genes in nectar spur development

The implication of class 1 *KNOX* genes in the development of nectar spurs in close relatives of snapdragon (Golz *et al.*, 2002) suggests a novel role for *KNOX* genes well outside their accepted roles in apical meristem maintenance and complex leaf morphogenesis.

1.6.1 The evolution of spur development in Antirrhineae

In *A. majus HIRZ* and *INA* are considered duplicate genes that play a wild type role homologous to that of *STM* in Arabidopsis and *KN1* in maize, perhaps acting redundantly to maintain indeterminacy in the SAM (Golz *et al.*, 2002). The mutant alleles of *Hirzina* and *Invaginata* have been interpreted as neomorphic genes expressed ectopically in developing organs to 'organise' a novel proximo-distal axis of growth that duplicates the wild type petal tube of *A. majus*.

Past phylogenetic assessments of the tribe Antirrhineae have been unable to agree whether *A. majus* has arisen from a non-spurred ancestor or lost spurs during the course of evolution (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000; Vargas *et al.*, 2004). However, the presence of a swelling, the *gibba*, at the base of the petal tube of many spurless members of Antirrhineae, that might represent a vestigial spur, and the recent phylogenetic evidence that places *Antirrhinum* as more derived than *Linaria* (Oyama & Baum, 2004), suggests that nectar spurs have been lost in snapdragon (Golz *et al.*, 2002). This observation led Golz *et al.* (2002) to propose that in the *Hirz-d153* and *Ina-d1* mutants ectopically expressing *HIRZ* and *INA* as a result of transposon insertions, an organising process that was quiescent in spurless members of *Antirrhinum* had been reactivated (Golz *et al.*, 2002). In Antirrhineae, at least, it appears that nectar spurs have evolved as an additional petal tube formed by redeployment of an organiser that controls the elaboration of the ancestral petal tube, in combination with spatially restricting factors, to promote petal tube development in certain parts of the corolla. Over time, such a structure has evolved into the nectar spur that typifies many close relatives of snapdragon, such as *Linaria*.

Given that spurs have arisen many times among the angiosperms (Hodges, 1997; Hodges & Arnold, 1995), and the incredibly labile and rapid evolutionary appearance of the trait in *Aquilegia* (Hodges & Arnold, 1995) and orchids (Box *et al.*, 2008), it is possible that similar neomorphic mutations underlie spur development and evolution in a variety of angiosperms, reviving theories of a simple developmental genetic pathway.

1.6.2 Exploring the role of *KNOX* genes in nectar spur development: a candidate gene approach

A number of approaches can be taken to understand the development and evolution of biological structures. The most common methods employ a careful analysis of both naturally occurring and laboratory induced mutants using a suitable organism which is amenable to genetic analysis, i.e. one that has a rapid generation time, can be easily manipulated/mutagenised and, most importantly, is genetically tractable.

Many suitable plants have been identified with just such a suite of traits, most notably thale-cress (*Arabidopsis thaliana*), maize (*Zea mays*) and the garden snapdragon (*Antirrhinum majus*). Subsequently these plants have been developed into well-characterised model organisms. Traditional workhorses of molecular and developmental plant biology, such as thale-cress, have taught us a great deal about many common plant developmental processes and have provided some information pertaining to the evolution of morphological novelty. However, even when combined, these models cannot account for the sheer diversity of traits in the 250,000 plus species of angiosperm. As such, developing an understanding of many biologically significant traits necessitates the use of non-model organisms.

No current, well-characterised plant model exists with naturally occurring nectar spurs. While *Aquilegia* is currently being developed as new evolutionary developmental model (Kramer, 2009) that may provide significant insight into a broad range of additional floral traits, it is currently not sufficiently well developed to explore the molecular basis of spur development and evolution. Fortunately, the discovery of the two naturally occurring mutants of *A. majus* (Golz *et al.*, 2002) have indicated a potential role for class 1 *KNOX* genes in the development and evolution of floral nectar spurs in the tribe Antirrhineae. These mutants provide a foundation upon which detailed scrutiny of *KNOX* gene function may be conducted in closely related taxa that possess floral nectar spurs using a candidate gene approach.

However, while identifying causative alleles in mutants is a powerful demonstration of how mutants may be utilised to unravel molecular pathways that govern the development of complex floral traits, it does not, as such, constitute definitive evidence that alterations in expression of this gene are involved in generating new morphologies in nature, the so-called mutationalist fallacy (Coyne & Lande, 1985). It does, however, provide a good starting point for comparative studies.

1.6.3 Candidate systems for evolutionary-developmental studies of spurs

Choice of organism is key to the success of any evo-devo endeavour, in order to minimise experimental difficulties whilst maximising scientific interest. Many authors typically overcome this difficulty by investigating morphological features in close relatives of model species using a candidate gene approach, thereby improving the chances that currently well-established techniques may be applicable to the organism with morphological characters of interest.

1.6.3.1 *Linaria vulgaris* (Antirrhineae, Plantaginaceae, Lamiales) as a model system for nectar spur development

Much of what is currently known about floral nectar spur development has emerged from careful analysis of the *Hirz-d153* and *Ina-d1* mutants of the model plant *A. majus*. Despite the absence of a genome sequence and reliable genetic transformation, *Antirrhinum* is an excellent model for genetic studies with a high degree of phenotypic variation, hardiness, relatively short generation time of three months, ease of selfing and cross-pollination and the ability to carry out targeted mutagenesis via the use of transposon-tagging techniques (Schwarz-Sommer *et al.*, 2003).

To further our understanding of the role that *KNOX* genes may play in nectar spur development a number of authors have suggested that the orthologous genes of *HIRZINA* and *INVAGINATA* should be isolated, and their expression patterns characterised, in a close relative of *A. majus* which bears a true floral nectar spur. By far the best organism for this is common toadflax, *L. vulgaris* (Damerval & Nadot, 2007; Galego & Almeida, 2007; Whitney & Glover, 2007).

Linaria has been used previously in a number of developmental and evolutionary analyses of floral traits including flower colour variegation (Galego & Almeida, 2007) and floral zygomorphy (Almeida *et al.*, 1997; Cubas *et al.*, 1999; Hileman & Baum, 2003) as a natural progression from primary studies conducted using snapdragon. In many ways *Linaria* is a model system in its own right, possessing many of the properties that make snapdragon an excellent genetic model; although, like snapdragon, *Linaria* is not amenable to reliable genetic transformation and has no published genome sequence. Work by a number of authors has demonstrated that many of the molecular techniques established for snapdragon can be adapted effectively for use with *Linaria*. Even the transposon-tagging system, which makes snapdragon such a powerful genetic model, may be a useful tool that can be developed in the future (Galego & Almeida, 2007). Whilst *Linaria* is by no means a perfect model system, it represents a logical step to investigate the involvement of *KNOX* genes in nectar spur development.

1.6.3.2 Orchidoid orchids (Orchidoideae, Orchidaceae, Asparagales) as models to investigate the relationship between nectar spurs and speciation

Golz *et al.* (2002) proposed that *KNOX* genes may have been co-opted in the evolution of nectar spurs in the Antirrhineae. However, using close relatives of *Antirrhinum*, such as *L. vulgaris*, to infer the pattern of nectar spur evolution in this group is unreliable as current phylogenies do not reveal confidently whether *Antirrhinum* is more likely to have arisen from a non-spurred ancestor, or to have lost spurs (Albach *et al.*, 2005; Ghebrehiwet

et al., 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004). Therefore, although *Linaria* is an excellent model for understanding the role of *KNOX* genes in nectar spur development, it is a difficult system in which to address the role of nectar spur evolution in the rapid radiation of species in taxa that possess floral nectar spurs.

To better understand the role of nectar spurs in the evolution of new species a system with considerable phenotypic variation in nectar spur morphology, plus a robust and reliable phylogeny is required. Whilst orchids are by no means a good molecular model, lacking in many of the physiological, technological and life history traits present in genetic models, they are a group renowned for considerable phenotypic variation and high species diversity with more than 20,000 species divided into approximately 800 genera (Dressler, 1993). Furthermore, such diversity has attracted significant systematic interest as a result of which orchids are exceptionally well characterised phylogenetically (Bateman *et al.*, 2003).

In addition to parallel strands of research using *L. vulgaris*, this project focuses on a number of selected species within the subtribe Orchidinae (Orchidoideae). This group has a particularly well-resolved phylogeny and almost all of the species have been extensively sampled for morphological (Burns-Balogh & Funk, 1986; Dressler, 1993; Freudenstein & Rasmussen, 1999) and repeated nuclear *ITS* rDNA phylogeny to include representative DNA sequences from 190 species (Bateman, 2001; Bateman *et al.*, 1997, 2003). The subtribe Orchidinae is particularly interesting due to the considerable morphological diversity between the nectar spurs from a number of closely related species pairs. Despite considerable morphological diversity, there is very little *ITS* sequence variation between species (Bateman & DiMichele, 2002; Bateman *et al.*, 2003), strongly suggesting that Orchidinae is a group in the process of speciation, making it highly valuable to evolutionary-developmental studies.

Medium/long nectar spurs are considered ancestral in *Dactylorhiza* and more broadly among Orchidinae, as nectar spurs are a prominent feature among members of the genus *Gymnadenia*, sister genus to *Dactylorhiza*, *Platanthera*, sister to both, and *Pseudorchis*, sister to all three genera (Bateman *et al.*, 2003). Mapping spur morphology across the tree for the subtribe Orchidinae demonstrates a minimum of five losses of spurs during the evolution of this group plus a further seven cases in which the nectar spur has become considerably reduced in length (Bateman, 2005; Bateman & DiMichele, 2002). Additionally, in many other cases the spur is retained but nectar secretion is not. Whilst switches to deceptive pollination systems have been shown to radically alter pollinator behaviour and have been implicated in speciation within Orchidinae (Cozzolino & Widmer, 2005),

the potential influence of nectar spur reduction in Orchidinae has been less-well characterised. Bateman (2005); Bateman & DiMichele (2002) implicated changes in nectar spur morphology as a potential agent for recent speciation events in this group, perhaps by altering plant-pollinator interactions or in relation to a switch from allogamy (outbreeding) to autogamy (inbreeding) (Bateman & DiMichele, 2002). Furthermore, Bateman (2005); Bateman & DiMichele (2002) suggested that such length differences may have arisen by heterochronic shifts in the timing of expression of a neomorphic gene.

Understanding the genetic determinants of spur reduction in Orchidinae, and the subsequent effects on nectar spur morphology and speciation, is one of the main goals of this research. Pairs of orchidoid orchid species with contrasting spur morphologies from the genera *Gymnadenia*, *Orchis* and *Dactylorhiza* were used to investigate whether *KNOX* genes were involved in orchid spur development. Initially multiple species pairs were trialled and assessed morphologically as a continuation of previous work initiated at the Royal Botanic Gardens, Kew (see Box *et al.* (2008) cf. spur ontogeny in the genus *Gymnadenia*). However, due to practical difficulties obtaining plant material, only one species pair was examined in detail for this thesis. Nectar spur ontogeny and the patterns of *KNOX* expression were investigated in the ancestrally medium-spurred orchid *Dactylorhiza fuchsii* and compared to those of a short-spurred close relative, *D. viridis*, to investigate the potential role played by spatial-temporal shifts in *KNOX* gene expression in the phenomenon of spur reduction documented in Orchidinae and to examine the broader significance of *KNOX* genes in angiosperm nectar spur evolution and development.

1.6.4 Research objectives

The principal aim of this work is to investigate the role of *KNOX* genes in the development and evolution of floral nectar spurs. *Linaria vulgaris* has been chosen as the principal model in which to develop and further understand the genetic basis of floral nectar spur development. This will be achieved by building on the insights gained from the analysis of the *Hirz-d153* and *Ina-d1* snapdragon mutants.

It is possible that nectar spurs in different angiosperm taxa have evolved by different means. To investigate this possibility, the role of *KNOX* genes in nectar spur development has also been considered in an evolutionarily distant group of plants, the orchids in the subtribe Orchidinae (Orchidaceae). The aims of investigating the role of *KNOX* genes in orchid spur development are two-fold;

1. Examining spur development in such a divergent group of plants permits an assessment of how conserved the role of *KNOX* genes may be in angiosperm spur development.

2. The implication of changes in spur length in recent speciation events in Orchidinae provides an opportunity to investigate whether recent speciation events are related to changes in the spatio-temporal expression of *KNOX* genes.

By focusing on these two principal aims this research provides insight into larger processes that govern the evolution of biological novelty, specifically;

- Do similar molecular mechanisms underlie the repeated evolution of a single trait?
- Are existing developmental pathways redeployed in the evolution of novel biological structures?
- Can neomorphic mutations in key regulatory genes contribute to the origin of novel biological structures in plants, a phenomenon well characterised in animal models?

1.6.5 Experimental programme

Several areas of experimental focus have been identified in order to answer such ambitious evolutionary and developmental questions.

1. Nectar spur ontogeny will be examined using a range of imaging techniques such as light/electron microscopy and X-ray tomography to develop an understanding of the growth dynamics of floral nectar spurs in each of the study taxa and to provide targets for subsequent candidate gene isolation.

2. Candidate *KNOX* genes will be isolated from the study taxa using a combination of degenerate PCR and gene walking techniques such as 5′ and 3′ RACE-PCR.

3. Hypotheses pertaining to the putative functions of candidate *KNOX* genes will be generated using molecular phylogenetic approaches to place the newly identified genes in a phylogenetic context alongside previously characterised *KNOX* genes described in the literature.

4. Spatio-temporal patterns of candidate gene expression will be assessed using RT-PCR, quantitative RT-PCR and *in situ* hybridisation.

5. Transgenic experiments using a suitable heterologous host will be conducted to establish putative protein function of novel *KNOX* gene candidates.

Chapter 2

Materials and Methods

2.1 Materials

Details of specific solutions referred to in the text are summarised in Appendix A. All primers are detailed in Appendix B. The names and accession numbers of DNA sequences used in phylogenetic analyses are detailed in Appendix C.

2.1.1 Laboratory reagents and suppliers

Standard laboratory reagents and chemicals were obtained from Fisher Scientific (Loughborough, UK), VWR (France), BDH Laboratory Supplies (Poole, UK) and Sigma-Aldrich (St. Louis, MO, USA). Bacterial culture reagents and antibiotics were supplied by Becton, Dickinson and Co. (Sparks, MD, USA), Oxoid Ltd. (Basingstoke, UK) and Melford Laboratories (Ipswich, UK). The pGEM®-T Easy TA-Cloning kits used were obtained from Promega Corporation (Madison, WI, USA). Plasmid DNA purification and DNA Gel Extraction Kits were supplied by Qiagen (Crawley, UK). *Taq* polymerase, dNTPs and Bioscript® RNase H-Low reverse transcriptase were obtained from Bioline (London, England). Radioactively labelled probes were generated using the Stratagene Prime-It® II Random Primer Labelling kit (La Jolla, CA, USA). Radioactivity was supplied by MP Biomedicals (formerly ICN; Irvine, CA, USA). For extraction of ultra-pure genomic DNA (gDNA) the Puregene Genomic DNA Purification kit was used; supplied by Gentra Systems (Minneapolis, MN, USA). Endonuclease enzymes used were obtained from New England Biolabs (NEB; Hertfordshire, England) and Invitrogen Life Technologies (Paisley, UK). Oligonucleotide primers were obtained from Invitrogen Life Technologies and Integrated DNA Technologies (IDT, Coralville, IA, USA; formerly VH Bio Ltd. Gateshead, UK). 5′ RACE and SYBR® GreenER qRT-PCR kits were supplied by Invotrogen Life

Technologies. Microscopes and Histology equipment used were supplied by Leica (Milton Keynes, UK), Zeiss (Welwyn Garden City, UK) and Agar Scientific (Essex, UK). Reagents for *in situ* hybridisation were obtained from Roche (Welwyn Garden City, UK).

2.1.2 Sources of plant material and growth conditions

Plant material was obtained from a variety of sources and in a number of different forms depending upon the intended use.

2.1.3 Sources of preserved orchid material for morphological and ontogenetic analysis

Orchid material for morphological and ontogenetic analysis was obtained from the extensive preserved collections at RBG, Kew and collected from the wild by M.S. Box, Dr. P.J. Rudall and Prof. R.M. Bateman (RBG Kew, UK) in the Dolomites, northern Italy, and three sites in the Chiltern Hills (UK). Wild collected material included: *Gymnadenia conopsea* (L.) R.Br., *G. odoratissima* (L.) Rich., *G.* (formerly *Nigritella*) *austriaca* (Teppner Klein) Delforge, Dolomites, northern Italy, July 2005; *Dactylorhiza fuchsii* (Druce) Soó, Hertfordshire, UK and *G. conopsea*, Buckinghamshire, UK, May 2006; *D. viridis* Bateman, Pridgeon and Chase, formerly *Coeloglossum viride* Hartm, Hampshire, UK, June 2006. This material included whole inflorescences complete with flowers at various stages of development. All fresh material was fixed immediately after collection using Formalin-Acetic-Alcohol (section 2.2.1).

Additional orchid material was obtained from the RBG, Kew spirit collection. Material obtained from the spirit collection is listed below as SC, plus the collection number: *D. fuchsii* (SC - 29047.359; 36006; 70528), *D. viridis* (SC - 25974; 70512), *Gymnadenia conopsea* (SC - 7633), *G. densiflora* (Wahlenb.) Aver. (SC - 45915), *G. odoratissima* (SC - 5880; 31708; 40964) and *G.* (*Nigritella*) *austriaca* (SC - 31599; 40941; 70525).

2.1.4 Sources of preserved orchid material for molecular analysis

Silica-dried flowers of the orchids *G. rhellicani* (Teppner and Klein), *G. odoratissima*, *G. conopsea*, *D. fuchsii*, *D. incarnata* (L.) Soó, *D. viridis*, *Orchis italica* Poir., and *O. anthropophora*, were kindly supplied by Prof. R.M. Bateman for gDNA extraction. Tubers of *D. fuchsii*, *D. incarnata* and *G. conopsea* were donated by RBG, Kew for RNA extraction

and cDNA synthesis. Further tubers of *O. italica* and young *G. conopsea*, *G. odoratissima* and *O. anthropophora* plants were obtained from John Haggar, March 2007 (John's Orchids West Sussex, UK). Additional living plant material was collected from the wild (with permission of the landowner):*Orchis*(formerly Aceras) *anthropophora* R.Br., inflorescences and living plants of *D. fuchsii* and *G. conopsea* were obtained from two sites near Box Hill, Surrey, UK, and the Chiltern Hills, UK, May 2007. *Gymnadenia austriaca* and *D.virdis* material was collected under permit in the Nockberge National Park, Austria, June/July 2007-2008 (Permit Refs: SP3-NS-865/2007 (002/2007) SP3-NS-865/2007 (004/2008), Bezirkshauptmannschatt Spittal Naturschutzbehörde).

2.1.5 Growth conditions for living plant material

Living orchid plants obtained from growers and from the wild were placed in porous clay pots in a free-draining compost made according to a modified formula provided by John Haggar (John's Orchids, West Sussex, UK) designed to best simulate the natural habitat in which the orchids grow. Plants were kept in approximately natural conditions; outside in a shaded plunge bed at the Cambridge University Botanic Gardens.

Nicotiana tobaccum cv. Samsun, *Linaria vulgaris* and *Antirrhinum majus* plants were grown from seed in a controlled greenhouse environment at 26°C with a 16 hour light regime. Seed of each was surface sown in Intercept®-treated soil in 6-inch pots. Seedlings were thinned out after germination to prevent over crowding, these stock plants were grown in a controlled greenhouse environment at 26°C with a 16 hour light regime and watered daily. Genetically modified plants were grown in 6-inch pots containing a 2:1 (vermiculite:soil) mix containing 0.02g/L: Intercept®, watered initially with distilled water and grown in a controlled growth room at 20°C, 60% relative humidity, 200μM light with a 16 hour light regime. Plants were subsequently watered automatically four times daily.

2.2 Floral morphology and ontogeny

A variety of low and high-powered microscopy techniques were utilised to investigate the anatomy, morphology and ontogeny of the flowers of *L. vulgaris* and a variety of orchid species with particular emphasis on the morphology and ontogeny of the floral nectar spur.

2.2.1 Material preparation and preservation

Freshly collected flowers and inflorescences were fixed immediately using Formalin-Acetic-Alcohol for a minimum of 72 hours. Preserved orchid material obtained from the RBG, Kew Spirit Collection was stored in Kew Mix and transferred to Copenhagen Mixture prior to use.

2.2.2 Light microscopy (LM)

Flower buds were dissected in 70% ethanol and subject to a dehydration series and standard methods of Paraplast™ wax embedding using a Leica TP 1010 Tissue Processor. Serial sections were cut at a thickness of 14μm using a Leica RM 2155 rotary microtome and a disposable steel microtome knife. The sections were mounted onto Polysine™ glass slides and allowed to adhere in a slide oven at 25°C for 24 hours, stained with Safranin and Alcian blue using a Leica AutoStainer XL automated slide stainer, and mounted permanently using DPX mountant. Photomicrographs were taken using a Leica DMLB photomicroscope fitted with a Zeiss AxioCam™ digital camera.

2.2.3 Scanning electron microscopy (SEM)

Floral buds and inflorescences were dissected in 70% ethanol to expose floral organs and organ primordia, and dehydrated in an ethanol series to absolute ethanol. Dehydrated plant material was Critical Point Dried using a Tousimis Supercritical Autosamdri 815B critical point drier, mounted onto SEM stubs using double-sided adhesive pads and coated with platinum using an Emitech K550 sputter coater for two periods of four minutes. Subsequently material was visualised using a Hitachi S-4700-II cold field emission scanning electron microscope (FE-SEM) at 2.0KV. Photomicrographs and electron micrographs were processed and collated using Adobe Photoshop® CS2.

2.2.4 X-ray tomography (XRT)

Floral buds and inflorescences were prepared for X-ray tomography (XRT) in the same way as for SEM (section 2.2.3). Samples were kindly scanned by Denis Van Loo (Centre for X-ray Tomography, Ghent University (UGCT)). Samples were scanned using an in-house developed very high resolution XRT setup. The setup consists of a transmission type tube (Hamamatsu Photonics GmbH, Germany) that can achieve a resolution of 400nm and a 12 Mpix CCD camera with a Gadox scintillator with 7.7μmpixel size. During the scans, the detector was set to binning mode, resulting in approx. 1000x1000 pixel

images. The samples were rotated over 360° whilst taking projection images at 0.36° intervals. The average scan time was 1 hour. The projection images where reconstructed as virtual sections using an in-house developed software package (Octopus, University of Ghent). Virtual sectioning and 3D-reconstructions was performed with VGStudio software (Volume Graphics GmbH, Germany).

2.3 Cloning of candidate *KNOX* genes

The characterisation of candidate genes involved in spur development firstly involved the identification of conserved KNOX protein domains and the design of suitable primers for degenerate polymerase chain reaction (dPCR) based on the highly conserved C-terminal meinox and homeobox domains. The resultant dPCR-generated DNA fragments were cloned, sequenced and analysed. The sequences of candidate gene fragments viewed as likely candidates for spur development were then extended using a combination of 5′ and 3′ RACE-PCR (rapid amplification of cDNA ends; sections 2.3.12.2, 2.3.12.1).

2.3.1 Genomic DNA extraction

Genomic DNA (gDNA) was extracted by two different methods depending on the intended use of the gDNA extract. Confirmation of the presence of gDNA in the extract was made by agarose gel electrophoresis (section 2.3.6) to visualise the high molecular weight gDNA band.

2.3.1.1 gDNA extraction for PCR

For gDNA extractions intended for use as PCR template a modified DNA extraction protocol was utilised (Sambrook & Russel, 2001). Fresh, frozen or silica-dried plant material was ground to a fine powder in liquid Nitrogen using a sterile pestle and mortar or micro-pestle. The resulting ground plant material was transferred to a sterile 1.5ml eppendorf tube to which 500μl of autoclaved DNA Extraction buffer was added and then incubated at 65°C for 2 minutes, after which an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added.

The homogenate was vortexed and the gross cellular material, protein and nucleic acids separated by centrifugation at 21,460 x g. The upper aqueous layer, containing the nucleic acids, was removed to a fresh 1.5ml eppendorf and 500μl of chloroform added. The mixture was again vortexed and spun again at 21,460 x g, contaminant proteinaceous material dissolves in the lowermost phenol:chloroform phase whilst the purified nucleic

acids are removed in the upper aqueous phase to another fresh 1.5ml eppendorf. To this 350μl of isopropanol and 50μl of 3M NaAc was added, mixed thoroughly by pipetting to avoid shearing of the gDNA and the nucleic acids allowed to precipitate overnight at 4°C. The gDNA was then collected by centrifugation at 21,460 x g and washed with 500μl of 70% ethanol, allowed to air dry and resuspended in 20-30μl of autoclaved TE buffer or sterile dH_2O and stored at -20°C.

2.3.1.2 gDNA extraction using the Puregene® kit for Southern Hybridisation

High purity gDNA extracts (10-150μg DNA) intended for use in Southern Hybridisation (section 2.5) were made using the Puregene® Genomic DNA Extraction Kit (Gentra Systems). The extraction was made according to the manufacturer's instructions (pp. 47-48 Puregene® manual, 2002). The Puregene® gDNA extraction system extracts DNA from homogenised plant tissues previously ground under liquid Nitrogen. The supplied DNA extraction buffer contains an anionic detergent with a DNA stabiliser. Contaminant RNA is digested using RNase and residual protein contaminants removed by salt precipitation. gDNA is recovered by alcohol precipitation and dissolved in a buffered solution containing a DNA stabiliser. This protocol can be used to isolate high yields of ultra-pure, highly digestible gDNA, suitable for use in Southern blotting.

2.3.2 RNA extraction

Between 100μg and 2g of fresh plant material was harvested, flash frozen and ground to a fine powder under liquid Nitrogen using a sterile pestle and mortar. Throughout the procedure the RNA extraction was conducted on ice and using a refrigerated centrifuge at 4°C to minimise RNase activity and subsequent RNA degradation. In the cooled mortar 1mL of Tris-saturated phenol and 2mL of autoclaved RNA Extraction buffer were added and the mixture ground again until defrosted. 1mL of chloroform was then added to the homogenate and the mixture ground once more and transferred to a 15mL centrifuge tube. The homogenate was vortexed thoroughly and spun in a 4°C refrigerated centrifuge at 5,500 x g for 5 minutes. After spinning the upper phase, containing RNA, was transferred to a clean 15mL centrifuge tube containing an equal volume of chilled 4M LiCl and mixed by inversion and the RNA allowed to precipitate at 4°C overnight.

Precipitated RNA was pelleted in a 4°C refrigerated centrifuge and spun at 5,500 x g for 10 minutes; the pellet was resuspended in 500μl of autoclaved 1X DNase-I buffer to which 1μl (10 U/μl DNase-I (RNase-free) was added. The reaction was incubated at

37°C for 30 minutes and transferred to ice after which 500μl of phenol:chloroform:isoamyl alcohol (25:24:1) was added, mixed thoroughly by vortexing and the resultant mixture spun in a 4°C centrifuge at 5,500 x g for 5 minutes to remove contaminant proteins e.g. DNase-I in the phenol:chloroform phase. The upper aqueous phase was removed to a fresh 1.5mL eppendorf tube containing 2.5 volumes of chilled salty-ethanol and the RNA allowed to precipitate at -20°C for 60 minutes. RNA was pelleted for 5 minutes at 21,460 x g and 4°C . The resulting pellet was washed with 500μl of 70% ethanol, allowed to air dry and resuspended in 30μl of TE buffer or 0.1% DEPC treated dH_2O.

The presence of RNA in the extract was confirmed by agarose gel electrophoresis (section 2.3.6), RNA exhibits a characteristic banding profile due to the relatively high abundance of the 28S and 18S ribosomal RNAs.

2.3.3 First-strand cDNA synthesis

First-strand cDNA was synthesised using Bioscript® RNase H Low reverse transcriptase (Bioline) according to an optimised version of the manufacturers instructions. 2-8μl (0.5-5μg total RNA and either 0.5μg B26 Oilgo(dT) (carrying an adapter sequence for 3′ RACE-PCR, see Frohman *et al.* (1988)), 0.2μg random hexamer or 5-20pmole gene-specific primer (GSP) were added to a 0.5ml eppendorf tube, made up to a 12μl reaction volume with 0.1% DEPC treated dH_2O and incubated at 70°C for 5 minutes to denature the RNA. The reaction was rapidly cooled on ice and to this 0.5μl (10U) RNase inhibitor was added, 1μl 40mM dNTP mix (10mM each), 4μl of 5X Reaction buffer (provided with Bioscript® RT, Bioline), 0.25μl (50U) Bioscript® RT and 5.25μl RNase free dH_2O (20μl final reaction volume). The reaction mixture was mixed carefully using a pipette and incubated at 37-42°C for 60 minutes, Adding 20-100μl of dH_2O and freezing at -20°C or heating to 65°C for 15 minutes was used to quench the reaction.

2.3.4 Degenerate primer design

The protein sequences for a number of class 1 *KNOX* transcription factors from a variety of monocotyledonous and dicotyledonous plants including: *Oryza sativa* (OSH1; BAA03959.1), *Zea mays* (KN1; CAA43605.1), *Arabidopsis thaliana* (STM; NP 176426.1), *Antirrhinum majus* (HIRZINA; AAL67666.1, INVAGINATA; AAL67665.1) and the orchid *Dendrobium grex* Madame Thong-IN (DOH1; CAB88029.1) were obtained from the National Centre for Biotechnology and Information (NCBI) website. NCBI/GenBank accession numbers are indicated next to the abbreviated gene name.

Consensus-DEgenerate-Hybrid-Oligonucleotide-Primers, CODEHOP (Rose *et al.*, 1998), were designed based on three highly conserved protein motifs (DQFM, HYKP and WFIN; B.1) in and around the highly conserved KNOX2 and homeobox (HOX) domains (Figure 2.1). Primer design was based on reverse translation of multiply aligned sequences across the conserved regions of proteins (blocks) using Biology Workbench (San Diego Supercomputer Centre) and BlockMaker® (Weizmann Institute of Science).

The resulting block alignment was input to another web-based primer design program (CODEHOP; Weizmann Institute of Science) that selects primers with a non-degenerate 5′ clamp region of 18-25bp and degenerate 3′ core regions 11-12bp across four codons of highly conserved amino acids. Furthermore a codon bias (in this case for *Zea mays*) may be assigned in order to generate the most probable nucleotide primer sequence in the non-degenerate 5′ clamp region.

The combination of a long 5′ consensus clamp and short degenerate 3′ core provides several advantages over conventional degenerate primer design. The consensus clamp region significantly increases the annealing temperature at which degenerate PCR can be carried out and the short degenerate core region of the primer reduces overall degeneracy whilst focusing maximal annealing specificity on the 3′ end of the sequence to avoid mispriming and amplification of non-target gene fragments. Both modifications permit the use of higher annealing temperatures and lower concentrations of degenerate primer in the PCR reaction, which significantly reduces non-specific amplification.

PCR (section 2.3.5) was used to amplify 400-500bp candidate *KNOX* gene fragments containing the KNOX2/HOX protein coding domains using the CODEHOP degenerate primers described here from a variety of species specific cDNA templates consisting of a mixture of vegetative and floral tissues. Agarose gel electrophoresis (section 2.3.6) was used to confirm PCR amplification, in the event that multiple bands were present in a single PCR reaction nucleic acid fragments of interest were excised from the agarose gel using a QIAquick® gel Extraction Kit (Qiagen; section 2.3.7). Colony PCR and/or restriction digest was employed to check for the presence of nucleic acid fragments cloned into pGEM®-T Easy (section 2.3.11). Despite exhaustive efforts using gDNA template, dPCR was unsuccessful in the isolation of genomic copies of *KNOX* gene fragments.

Degenerate PCR was conducted on nine different species. *Linaria vulgaris* was targeted principally with the aim of identifying the orthologous genes of *A. majus Hirzina* and *Invaginata* in accordance with their potential roles in nectar spur development indicated by the work of Golz *et al.* (2002). Three groups of closely related species (*Dactylorhiza fuchsii*, *D. incarnata*, and *D. viridis*; *Gymnadenia conopsea*, *G. odoratissima*, and

G. rhellicanii; *Orchis anthropophora* and *O. italica*) were selected from the subtribe Orchidinae (Orchidaceae), each consisting of a long and short-spurred member, with the aim of establishing the evolutionary conservation of potential KNOX gene function in an unrelated angiosperm group and to increase the likelihood of identifying suitable KNOX gene candidates that my be involved in spur development. In addition to this, identification of KNOX gene candidates from multiple members of the subtribe Orchidinae was also conducted to facilitate the possibility of future work aimed at more fully investigating nectar spur evolution and associated speciation events in this group. This part of the research was started concurrently with morphological analyses (see section 3) and involved a broad range of species pairs, Unfortunately, the absence of time and sufficient plant material prevented this area of research from being completed. As such the depth of sampling is variable between species and only *KNOX* gene fragments isolated from the orchid species pair *D. fuchsii* (medium-length spur) and *D. viridis* (short spur) were targeted in continuing molecular analyses, owing largely to the relative ease with which plant material could be obtained for these species. A general program of dPCR was employed for the orchid taxa aimed at identifying as many *KNOX* gene candidates as possible.

2.3.5 Polymerase Chain Reaction (PCR)

PCR is a widely used molecular biology technique employed for a broad range of different experimental approaches used throughout this research project. The basic protocol is outlined here with minor alterations made according to the specific applications. Tables 2.1 and 2.2 outline the standard PCR reaction conditions and thermocycling profile. PCR was carried out using a Techgene TC-512 gradient PCR machine and Genius thermocyclers (Techne). Agarose gel electrophoresis (section 2.3.6) was used to confirm PCR amplification.

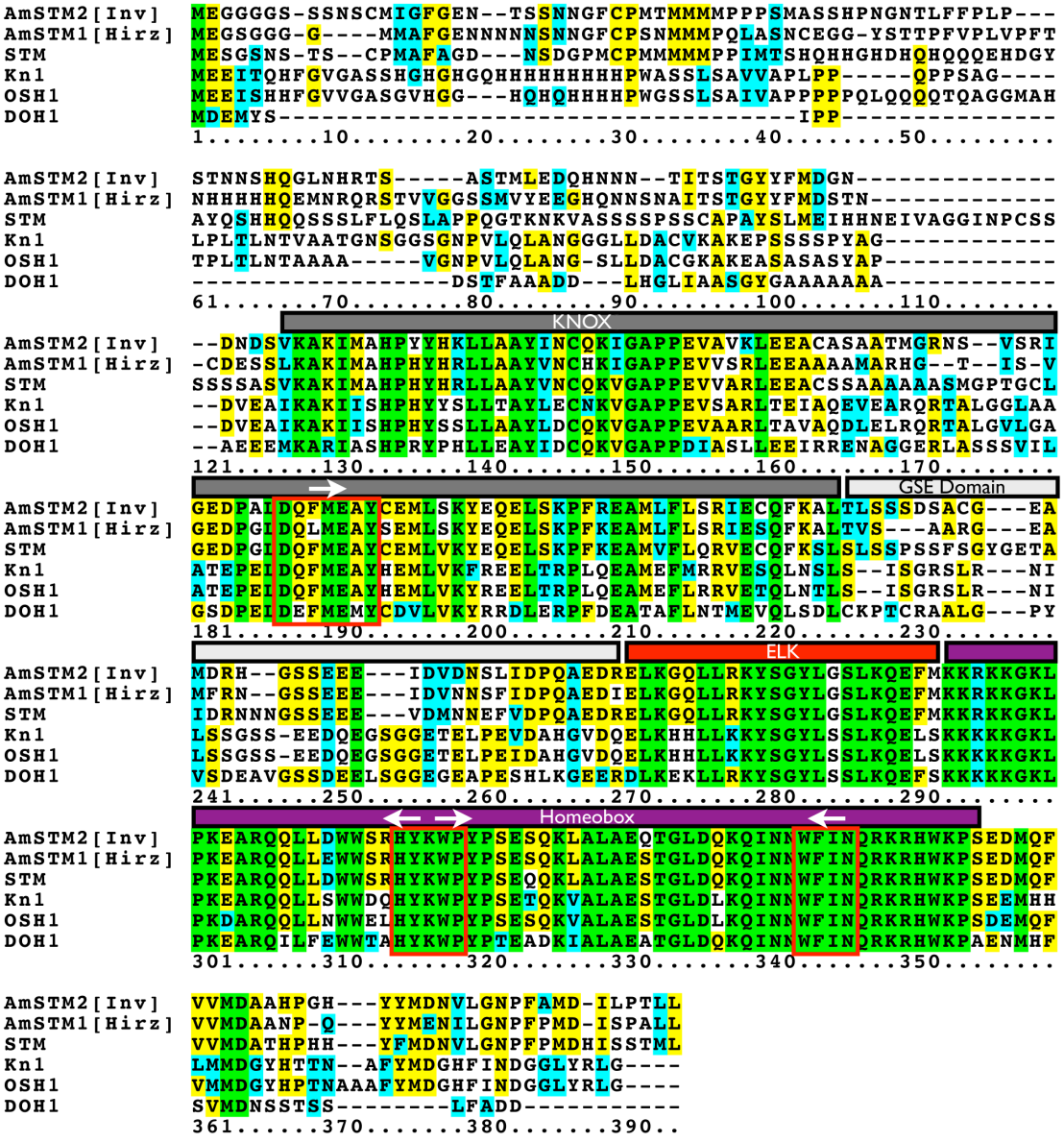


Figure 2.1: Protein alignment of landmark KNOX proteins used in dPCR primer design
Protein alignment of landmark class 1 KNOX transcription factors including *Oryza sativa* (OSH1; BAA03959.1), *Zea mays* (KN1; CAA43605.1), *Arabidopsis thaliana* (STM; NP 176426.1), *Antirrhinum majus* (HIRZINA; AAL67666.1, INVAGINATA; AAL67665.1) and the orchid *Dendrobium grex* Madame Thong-IN (DOH1; CAB88029.1) were used to identify highly conserved protein motifs for subsequent degenerate primer design. Green shading indicates full conservation of sequence, yellow indicates identical residues and blue indicates conservation of strong groups. Degenerate primers (white arrows) were designed around the highly conserved DQFM, HYKP and WFIN amino acid motifs (red boxes) in the KNOX2 and homeodomain regions of the proteins.

Table 2.1: Standard PCR reaction conditions

Reaction conditions	
Reaction component	*Final conc.*
10X Reaction buffer	1X
50mM $MgCl_2$	1.5mM
5mM dNTP	200M
100mM F/R Primer	200nM to 2M
Taq polymerase	1U
DNA template	1 to 5 ul

Table 2.2: Standard PCR thermocycle

Thermocycling profile			
Phase	*Temp (°C)*	*Time (mins)*	*Number cycles*
Initial denaturation	94	5	1
Denaturation	94	0.5	
Primer annealing	45 to 60	0.5	
Extension	72	0.5 to 3	35
Final extension	72	5 to 7	1
Hold	4	∞	

2.3.6 Agarose gel electrophoresis

Agarose gel electrophoresis is a key technique used to visualise nucleic acids and is employed as a confirmatory test of the success of several laboratory methods including gDNA/RNA extractions, PCR amplification and nucleic acid restriction digestion/Southern Hybridisation. Agarose gels were prepared by dissolving agarose in heated 0.5X TBE buffer to a final concentration of 0.8-1.5% (*w/v*), depending on the expected length of the nucleic acid fragments. The intercalating dye Ethidium Bromide (EtBr) was added to the cooled molten agarose solution to a final concentration of 0.1μg/ml before the gel was cast in an appropriately sized gel cassette. EtBr allows nucleic acid visualisation by fluorescing under UV light. Nucleic acid samples (1 to 10μl were mixed with 10X Orange-G gel loading buffer to generate sensible loading volumes. Gels were run between 60-100V as necessary in 0.5X TBE running buffer. The current running through the gel

draws the negatively charged nucleic acid fragments through the gel matrix towards the cathode, separating the fragments by size. The bands were then visualised under UV light and compared to between 3-5μg of an appropriate nucleic acid ladder of known size (e.g. 1-10Kb HyperLadder™ I DNA ladder (Bioline), 100bp DNA ladder (Invitrogen)).

2.3.7 Agarose gel extraction (QIAquick® Gel Extraction Kit)

On a number of occasions it was necessary to excise specific nucleic acid fragments from an agarose gel particularly after carrying out dPCR, which often results in the amplification of non-specific background products. The QIAquick® Gel Extraction Kit (Qiagen) was employed for this purpose. Agarose gel fragments of 100-400mg were excised with a scalpel under UV light and treated according to the manufacturer's instructions (pp. 25-26 QIAquick® manual, March 2006). The QIAquick® system uses a simple bind-wash-elute procedure based on ionic exchange. Gel slices are dissolved in a high-salt binding buffer to release the DNA from the agarose gel slice and promote efficient binding of the DNA to the silica membrane of the QIAquick spin column. Impurities are washed away using an ethanol based wash buffer and purified DNA eluted with a small volume of low-salt buffer or water. The resulting DNA is ready to use in all subsequent applications.

2.3.8 PCR purification (QIAquick® PCR Purification Kit)

In some instances PCR products or restriction digests were purified from excess salts, enzymes and/or primers using the QIAquick® PCR Purification Kit (Qiagen) according to the manufacturer's instructions. This kit operates as described in 2.3.7.

2.3.9 Ligation of PCR products into a plasmid vector

All DNA fragments generated by PCR were cloned using the pGEM®-T Easy (Promega) TA-cloning vector system. Cloning ensured that only a single fragment was sequenced and also generated cleaner sequencing traces than direct sequencing from PCR products. For PCR products generated by dPCR the absence of gene specific primers makes direct sequencing difficult. The linearised pGEM®-T Easy vector has single overhanging 3′ terminal thymidine (T) nucleotides at both ends. PCR amplified fragments generated using Taq polymerase (Bioline) have a single overhanging 3′ terminal adenine (A) at both ends, the resulting complementarity of these sticky ends greatly improves the efficiency of ligation with PCR products, generating a circular plasmid which may then be transferred into competent *E. coli* strain DH5α (section 2.3.10). The pGEM®-T Easy vector system contains a number of RNA polymerase promoters (e.g. T7/SP6/m13; Figure 2.2) flanking

a multiple cloning site within the α-coding region of the enzyme β -galactosidase. Successful ligations interrupt this coding sequence such that recombinant clones can be identified using blue-white screening on indicator plates containing the *lacZ* de-repressor isopropyl β-D-thiogalactoside (IPTG) and the colour changing substrate 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal). Successful recombinants with an interrupted *lacZ* operon fail to produce β-galactosidase and grow as white colonies. Unsuccessful recombinants generate β-galactosidase, which cleaves the colour changing substrate X-gal, generating blue colonies.

Each ligation reaction was set up according to the following protocol and manufacturer's instructions (p.8 pGEM®-T Easy technical manual, December 2005). The reaction was optimised to provide an approximately 3:1 molar ratio of PCR product to plasmid vector (Equation 2.1). A maximum of 3.5μl of PCR product (insert) was added to a 0.2ml PCR tube along with 5μl 2X pGEM®-T Easy Rapid Ligation Buffer, 0.5μl (25ng) of pGEM®-T Easy vector, and 1μl of T4 DNA Ligase (3 Weiss units/l). Xμl of dH_2O was added to produce a final reaction volume of 10μl. The reaction mixture was mixed gently by pipetting, incubated overnight at 4°C and stored at -20°C.

$$\text{ng insert} = \frac{(\text{ng vector} \times \text{Kb size of insert})}{\text{Kb size vector}} \tag{2.1}$$

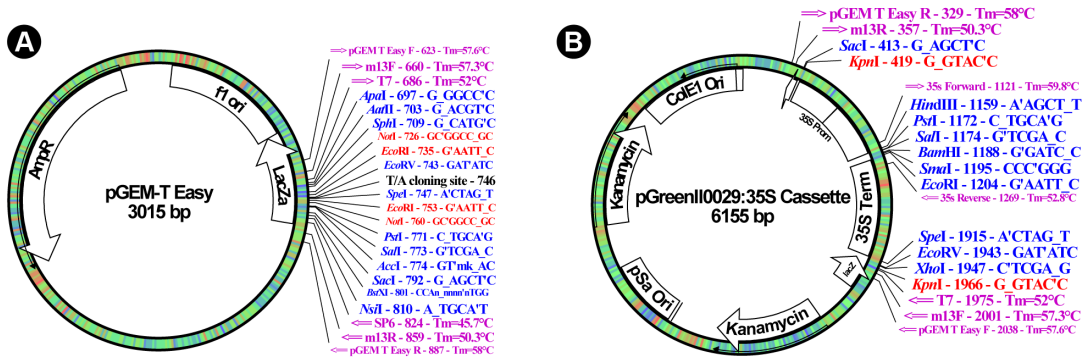


Figure 2.2: Vector maps of plasmids used in cloning experiments
Open reading frames are indicated by arrows on the plasmid backbone. Origins, promoters, terminators, operons and antibiotic resistance genes are indicated by block arrows. Primer annealing sites, restriction sites and multiple cloning sites are indicated by tags outside of the DNA backbone. A. pGEM-T Easy (Promega Corp.). B. pGreenII0029:35S constitutive expression cassette (Wilkins, 2004).

2.3.10 Transfer of ligated PCR products into competent *Escherichia coli* (DH5α)

Recombinant vectors such as that generated in section 2.3.9 were transferred into chemically competent *E. coli* DH5α (section 2.3.10.1). Successful recombinants were selected using a combination of antibiotic resistance and blue-white screening, where possible. For example, the pGEM®-T Easy plasmid vector contains a gene for ampicillin resistance, ensuring that only colonies that have been successfully transformed with the vector will be able to grow on antibiotic medium (Figure 2.2), plus blue-white screening to ensure that only colonies carrying the recombinant vector were selected.

Competent DH5α cells stored at -80°C were thawed on ice and 2μl of the ligation reaction was added to the solution and mixed gently. The mixture was incubated on ice for 20 minutes and then immersed in a 42°C water bath for 90 seconds and returned to ice for a further 5 minutes. 500μl of autoclaved SOC medium was added and the mixture was incubated at 37°C in a shaking incubator for 90 minutes at 180rpm. After incubation, 250μl of the *E. coli* culture was plated onto solid autoclaved LB agar plates containing IPTG (80μg /ml), X-gal (80μg /ml) and ampicillin (100μg /ml) for blue-white screening and antibiotic selection. Plates were incubated overnight at 37°C. A single white colony was picked and cultured in 5ml autoclaved liquid LB broth containing ampicillin (100μg /ml). The culture was incubated overnight at 37°Cand 180rpm in a shaking incubator.

2.3.10.1 Production of chemically competent *E. coli* (DH5α)

Chemically competent *E. coli* strain DH5α were produced according to the following protocol adapted from Sambrook & Russel (2001). DH5α is a strain of *E. coli* which has a low tendency to spontaneously recombine plasmid DNA; this makes it a useful tool for cloning of DNA fragments. A single colony of DH5α was picked from a plate of solid LB media and cultured overnight to an optical density (OD_{550}) of 0.3 in 5ml of autoclaved φ-broth. 1ml of this φ-broth culture was then sub-cultured into 100ml of LB broth and allowed to grow to an OD_{550} of 0.48 over approximately three hours in a shaking incubator at 37°C set to 180rpm. The 100ml culture was then divided between two 50ml sterile falcon tubes on ice and spun for 5 minutes at 5,500 x g in a refrigerated centrifuge at 4°C. The pelleted cells were resuspended in 40ml of chilled filter sterilised transformation buffer 1 (TFB I) gently by pipetting and incubated on ice for 5 minutes before being spun once more at 5,500 x g for 5 minutes at 4°C. Each resulting pellet was re-suspended in 2ml of ice cold filter sterilised TFB II and transferred to chilled, sterile 1.5ml eppendorf tubes in 100μl aliquots flash frozen using liquid Nitrogen and stored at

-80°C. The competency of these cells may be tested by transformation with a known mass of uncut plasmid carrying one or more antibiotic resistance genes and calculating colony-forming units (cfu) per μg DNA added. Only cells with 1 x 10^8 cfu/μg DNA or greater were used for cloning.

2.3.11 Screening for recombinant plasmids

In order to verify that potentially transformed bacterial colonies selected contained the recombinant plasmid carrying the desired PCR amplified DNA fragment two techniques were employed, restriction digests (section 2.3.11.1) and/or colony PCR (section 2.3.11.2). In each case once the presence of the desired insert had been verified, plasmid DNA was purified from a 3ml aliquot of the 5ml *E. coli* culture (generated in section 2.3.10) for use in DNA sequencing using the QIAprep® Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions which, though more efficient, is broadly similar to the method described in section 2.3.11.1. If larger quantities of plasmid were required the Qiagen® Plasmid Midiprep Kit was used according to the manufacturer's instructions.

2.3.11.1 Non-kit plasmid DNA miniprep and restriction digests

The cloned plasmids were purified from the bacteria and digested with an appropriate restriction endonuclease e.g. *EcoRI*. Of the 5ml *E. coli* culture generated from section 2.3.10, 1.5 to 3ml of culture was transferred to a sterile eppendorf tube and the cells pelleted by centrifugation using a microfuge at maximum speed. The cells were then resuspended in 150μlof sterile ribonuclease buffer (P1) by gently pipetting. An equal volume of sterile lysis buffer (P2) was added and the solution mixed by gentle inversion for no more than 5 minutes. The plasmid was then separated from the lysate by adding 200μl of precipitation buffer (P3), mixed thoroughly by inversion to generate a white, cloudy precipitate (containing proteins and gDNA) that was subsequently pelleted by centrifugation for 10 minutes in a microfuge at 21,460 x g. The supernatant was collected and transferred to a clean eppendorf tube and the purified plasmid DNA precipitated over 30 minutes at -20°C in 1mL of salty-ethanol, washed with 70% EtOH, allowed to air dry and subsequently re-suspended in 30μl of sterile dH_2O.

In the pGEM®-T Easy vector system the MCS is flanked either side by *EcoRI* restriction sites. The insertion of the desired PCR amplified fragment can therefore be assessed by purifying the plasmid DNA and digesting out the insert using *EcoRI*, or other appropriate restriction endonucleases (Figure 2.2). The presence of the insert may be confirmed by agarose gel electrophoresis (section 2.3.6). A 20μl restriction digest reaction

was assembled comprising 2μl of purified plasmid DNA, 2μl 10X *EcoRI* reaction buffer (NEB), 2μl 10X bovine serum albumen (BSA), 0.1μl (2U) *EcoRI* restriction endonuclease (NEB), and 13.9μl dH_2O. The reaction was incubated for 2 hours at 37°C and all 20μl loaded into a 1% agarose gel and visualised by gel electrophoresis with an appropriate DNA ladder. The size of the digested DNA fragment was compared to the original PCR product in each case.

2.3.11.2 Colony PCR

Colony PCR (cPCR) can be used as a more rapid alternative screening technique to restriction digest (section 2.3.11.1). A single bacterial colony was picked from a plate and dipped in 5μl dH_2O to provide the plasmid DNA template for the cPCR reaction. The same colony was pricked onto a duplicate plate from which a sequencing miniprep could be prepared using the QIAprep® Spin Miniprep Kit (Qiagen; section 2.3.11) if the colony was found to carry a recombinant plasmid with the desired insert. A pair of pGEM®-T Easy cPCR primers were designed in the regions flanking the MCS (Figure 2.2), these primers generate a 303bp product from **non**-recombinant circular pGEM®-T Easy in the absence of an insert. Selected colonies were screened for inserts by identifying cPCR-amplified fragments that were 303bp larger than the original PCR product visualised on a 1% agarose gel (section 2.3.6).

2.3.12 Extension of *KNOX* gene fragments

Subsequent to isolation of *KNOX* gene fragments by dPCR (section 2.3.4, Figure 2.1) the unknown sequence flanking either side of the fragment was determined by a combination of fragment extension methods. Unknown sequence upstream (i.e. 5′ of the sequence obtained from dPCR) was obtained by multiple rounds of 5′ rapid amplification of cDNA ends (RACE-PCR; section 2.3.12.2). Unknown downstream sequence (i.e. more 3′ sequence) may be obtained by 3′ RACE-PCR (section 2.3.12.1) (Frohman *et al.*, 1988). Subsequent to successful fragment extension by either method fragments were cloned and sequenced as outlined in sections 2.3.9 and 2.3.13. Full-length *KNOX* gene coding sequences were subsequently assembled using the manual alignment editor Se-Al v2.0a11 (Andrew Rambaut).

2.3.12.1 3′ RACE-PCR

Rapid amplification of cDNA ends (RACE) only requires a small amount of prior sequence knowledge to be effective. In 3′ RACE PCR, first-strand cDNA (section 2.3.3) is synthe-

sised from total RNA using reverse transcriptase and primed using the B26 Oligo(dT) primer containing a known adapter site for the primer B25 (Frohman *et al.*, 1988). Using standard thermocycling conditions (section 2.3.5) the RACE PCR reaction can then be carried out using the first-strand cDNA as template. A 10-fold higher specific primer concentration is used in conjunction with the relatively non-specific B25 adapter primer to favour amplification of specific products. In this case 3′ ends were obtained using a combination of B25 and a variety of gene specific sense primers designed for each *KNOX* gene identified, apart from the 3′ end of *LvKNOX2* (*LvINVAGINATA*) which was identified using a degenerate RACE-PCR approach utilising the degenerate HYKP F1 primer and B25. Additional nested gene specific primers were used in subsequent PCR using 1-5μl of a 10 to 100-fold dilution of the initial PCR product as template. The 3′ RACE PCR protocol is summarised in Figure 2.3.

2.3.12.2 5′ RACE-PCR

In order to obtain unknown upstream 5′ DNA sequence the GeneRacer® RACE PCR kit was used (Invitrogen). This kit employs an RNA ligase mediated Rapid Amplification of cDNA ends (RLM-RACE) procedure whereby an RNA adapter sequence is ligated to the 5′ end of de-capped, de-phosphorylated full-length mRNA transcripts prior to first-strand cDNA synthesis using Superscript®III supplied with the kit. A number of nested antisense specific primers were used in combination with supplied adapter primers to identify novel 5′ DNA sequence according to the manufacturer's instructions (Figure 2.4).

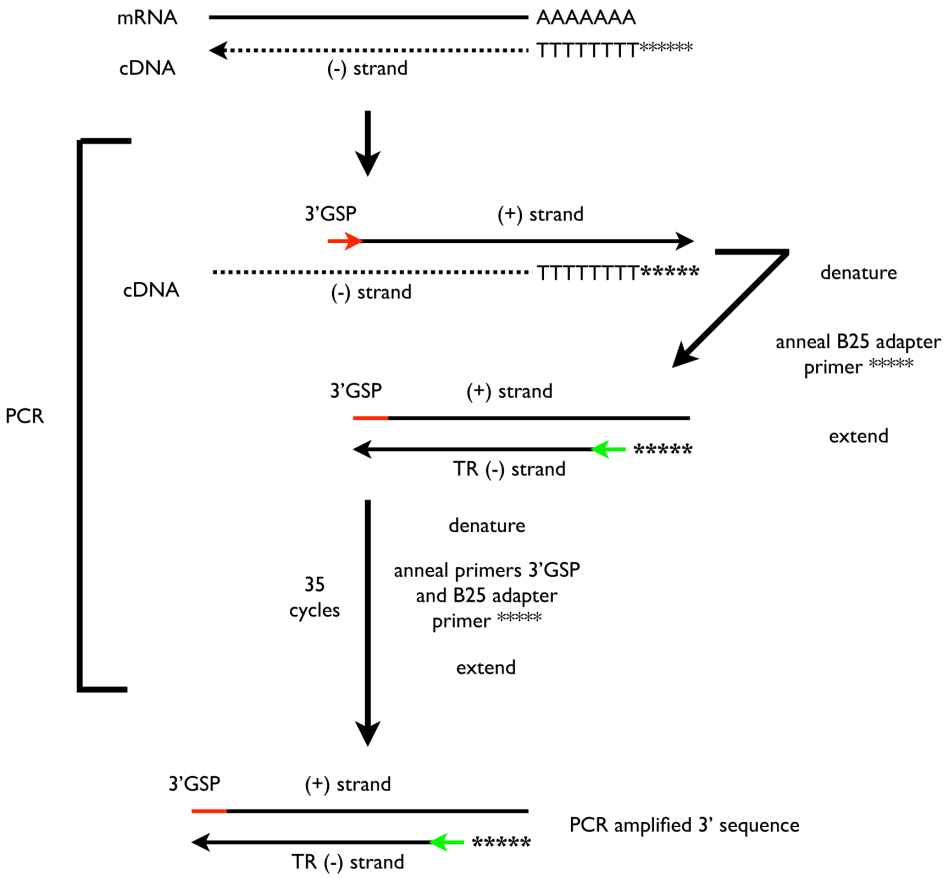


Figure 2.3: Schematic representation of the 3′ RACE protocol (Frohman *et al.*, 1988)
Schematic representation of the 3′ RACE protocol. First-strand cDNA is synthesised using the B26 oligo(dT) (TTTTTTTT*****) primer containing the B25 adapter primer site. The cDNA is used to amplify novel 3′ sequence by PCR using a 3′ gene specific primer (3′GSP; red) and B25 (*****; green). At each step the diagram is simplified to illustrate how only the new product formed during the previous step is utilized. TR indicates that the PCR product is truncated.

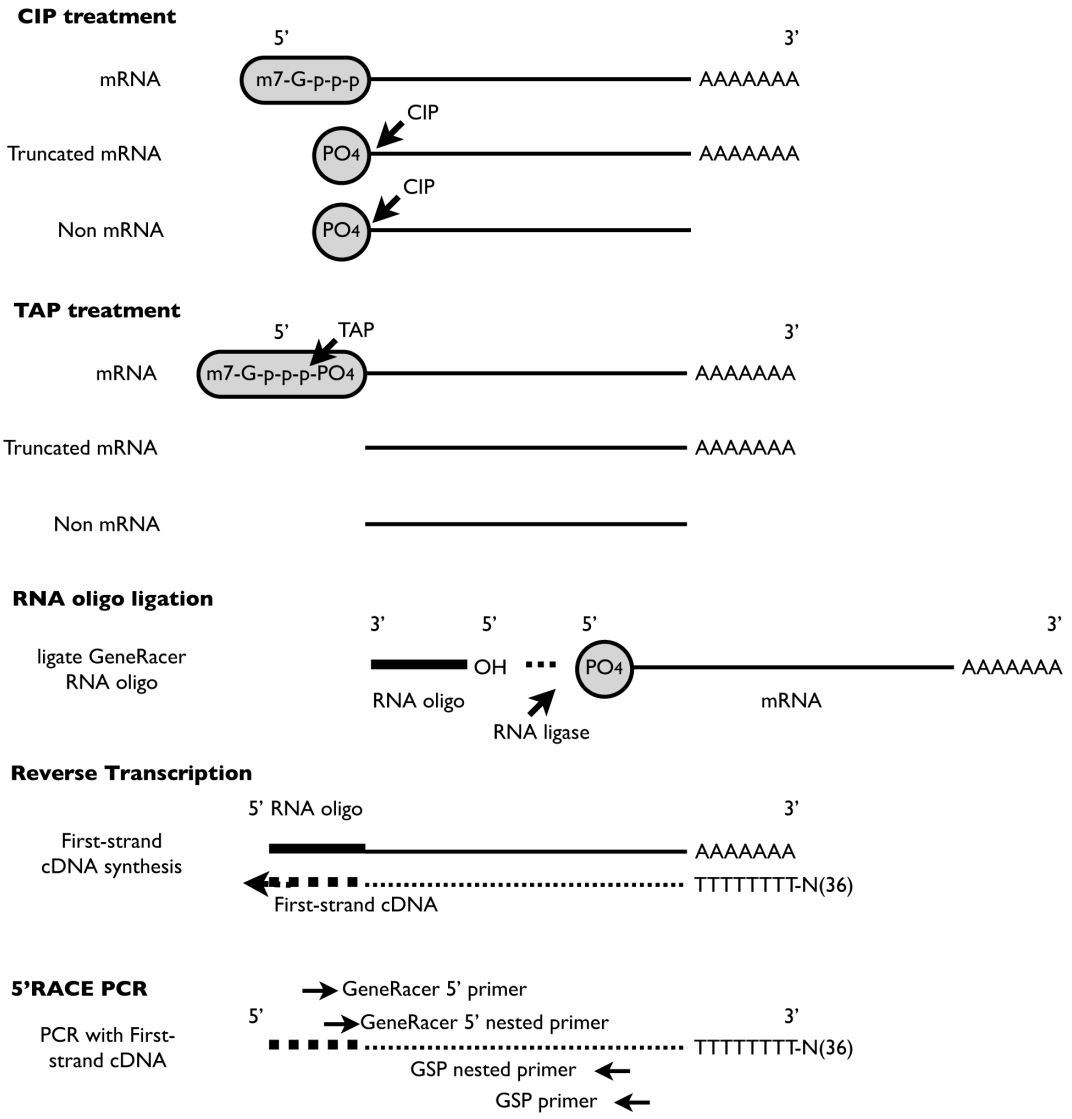


Figure 2.4: Schematic representation of the 5′ RACE protocol using the GeneRacer® kit
Schematic representation of the 5′ GeneRacer® RLM-RACE protocol. Treat total RNA with calf intestinal phosphatase (CIP) to remove the 5′ phosphates to eliminate truncated and non-mRNA from subsequent ligation. Treat dephosphorylated RNA with tobacco acid pyrophosphatase (TAP) to remove the 5′ cap structure from full-length mRNAs exposing the 5′ phosphate. Ligate the GeneRacer® RNA oligo adapter to the 5′ end of the de-capped mRNA to provide a known priming site for the GeneRacer® PCR primers. Reverse transcribe the ligated mRNA to generate cDNAs with known 5′ adapter sites that can be used subsequently in PCR consisting of supplied adapter primers and gene specific primers (GSP) to identify novel 5′ sequence.

2.3.13 DNA sequencing and analysis of purified DNA inserts

Purified recombinant plasmid DNA containing the target sequence was sent to the sequencing facility at the Department of Biochemistry, University of Cambridge. Sequencing of degenerate PCR products was done routinely using the m13 forward and/or reverse priming sites in the pGEM®-T Easy vector (Figure 2.2). The sequencing output files were analysed using the program 4Peaks v.1.7.2 (Griekspoor & Groothuis), which involved assessing the quality of the chromatogram and distinguishing the plasmid sequence from that of the insert. The edited DNA insert sequence was then identified using a translated nucleotide BLASTX query against the NCBI database (Altschul *et al.*, 1997), which identifies similar sequences by conserved domains or motifs.

2.4 Phylogenetic analysis of *KNOX* genes

Molecular phylogenetic methods were employed to assess the potential role of novel *KNOX* genes identified in this work. Novel genes were placed in a phylogenetic context with previously identified and more fully characterised *KNOX* genes available on the NCBI website. DNA and protein sequence information for a background of 88 class 1 and class 2 *KNOX* genes were sourced from previously published phylogenetic analyses of the *KNOX* gene family with representatives from all major land plant lineages (Appendix C).

2.4.1 DNA alignment

DNA sequences of all 31 newly isolated KNOX gene fragments and 86 published class 1 and class 2 *KNOX* gene sequences were aligned using ClustalW2 (Larkin *et al.*, 2007) and checked manually using the sequence alignment editor Se-Al v2.0a11 (Andrew Rambaut) and Bioedit v7.0.9 (Tom Hall, Ibis Biosciences, Carlsbad, CA, USA). A series of alignments were generated from nucleotide sequence data encoding both the meinox and homeodomain or the homeodomain only, including and excluding class 2 *KNOX* genes (Appendix C). Gaps were coded as '-' and missing data as '?' in the final alignment files.

Phylogenetic reconstruction was carried out using Bayesian Inference (BI; section 2.4.2) with the software PAUPv4.0, MrModeltest2.3 (Johan Nylander, Evolutionary Biology Centre, Uppsala University (Nylander *et al.*, 2004)) and MrBayesv3.1.2 (Huelsenbeck & Ronquist, 2001). All trees were rooted with the GenBank sequence for the green alga *Acetabularia acetabulum* (*AaKNOX1*: AF170172) and the in-group constrained to monophyly. Tree data files output from all analyses were viewed in the program Dendroscope (Huson *et al.*, 2007) and modified using the free vector graphics package Inkscape

(http://www.inkscape.org). The code used in these analyses can be found in the appendix (C).

2.4.2 Bayesian Inference analysis using PAUPv4.0b10, MrModeltest2.3 and MrBayes v3.1.2

BI analyses were conducted in accordance with the steps outlined in Huelsenbeck & Ronquist (2001) and the MrBayes v3.1.2 manual (published on-line; Ronquist *et al.* (2005)). The nucleotide substitution model was first determined using PAUPv4.0b10 and MrModeltest2.3, an optimised version of Modeltest 3.6 (David Posada) specifically developed for use with MrBayes v3.1.2. For the *KNOX* gene datasets Bayesian analyses employed a GTR+I+G model of sequence evolution selected by both the Akaike Information Criterion (AIC) and Hierarchical Likelihood Ratio Test (HLRT). Bayesian analyses were carried out using a Markov Chain Monte Carlo (MCMC) method as implemented in MrBayes v3.1.2. In each BI analysis four chains were run simultaneously for a sufficient number of generations until the value for the standard deviation of split frequencies fell below 0.01, indicating optimal convergence of the four independent chains. Once a split frequency below 0.01, was reached the analysis was terminated automatically according to the MrBayes code implemented (see appendix). Trees were sampled every 100 generations producing a total of Y trees (X generations/100). A burn-in fraction of 0.5 was used for each dataset, this represents the number of trees excluded from the analysis when computing the consensus tree and ensures that only convergent data sets are used to produce a majority rule consensus tree. MrBayes v3.1.2 assigns posterior probability values to Bayesian trees that indicate the reliability of the relationships between taxa/genes (Harrison & Langdale, 2006; Holder & Lewis, 2003; Nylander *et al.*, 2004). Analyses were repeated to ensure the test trees were recovered.

2.5 Analysis of *KNOX* gene copy number by Southern blotting

Localization and copy number of particular sequences within gDNA is usually accomplished by the transfer techniques first described by (Southern, 1975) in which gDNA is first digested with one or more restriction endonuclease enzymes (section 2.3.11.1) and the resulting fragments separated by agarose gel electrophoresis (section 2.3.6). The DNA is then denatured *in situ* and transferred from the gel to a solid support (section 2.5.1), usually a nitrocellulose membrane or nylon filter, this preserves the relative positions of

the nucleic acid fragments within the gel matrix. The gDNA fragments may then be hybridised to radio-labelled DNA or RNA (sections 2.5.2; 2.5.3), and a phosphoimager is then used to locate the position of the bands complementary to the probe.

2.5.1 Transferring nucleic acid fragments to a solid support: Southern blotting

Digested gDNA fragments (c. 10μg total DNA) were run slowly (40V) overnight on a 1% agarose gel as outlined and the DNA was transferred to a nitrocellulose (Schleicher and Schuell) or nylon membrane using a common Southern blot procedure (Southern, 1975). Prior to probing, the gel was visualised under UV after re-staining the gel with EtBr in 0.5X TBE containing 0.5μg/ml EtBr over 30-45 minutes on an agitator plate, excess EtBr was then removed by washing briefly with 1M $MgSO_4$ (Southern, 1975).

The DNA was denatured by submerging the gel twice in 0.25M HCl for 15 minutes, twice in Denaturation solution and then twice in Neutralisation solution. Each step was undertaken for 15 minutes on a shaker. The gel was then blotted onto a nitrocellulose/nylon membrane using the apparatus described in (Southern, 1975). A double layer of 3MM™ Whatman paper was placed beneath the agarose gel, which acts as a wick over a tank containing 10X SSC. The gel was surrounded by cling film to prevent short-circuiting and a piece of wetted nitrocellulose/nylon membrane was placed on the exposed gel. Two pieces of 3MM™ Whatman paper were wet in 2X SSC and placed on top of the membrane, along with two more pieces of dry 3MM™ Whatman paper and a stack of absorbent paper towels, approximately 7-10cm thickness. A glass plate and weight of c.1Kg were placed on top of the paper towels and left to blot overnight at room temperature in a fume cupboard. Once the DNA had transferred to the support the membrane was washed for 30 seconds in 2X SSC and left to air-dry before being baked for two hours at 80°C in a 3MM™ Whatman paper envelope to cross-link the DNA to the membrane.

2.5.2 Synthesis of ^{32}P α-dCTP labelled probe

Radio-labelled probes were prepared from PCR products (section 2.3.5) using the Stratagene Prime-It® II Random Primer labelling kit. 5μl (c.100ng DNA) of PCR product and 10μl random 9-mer primers, made up to 24μl with sufficient dH_2O, were added to a 1.5ml eppendorf and boiled for 5 minutes to denature the probe template DNA and allow the primers to anneal. To this 10μl of 5X dCTP buffer (Stratagene), containing unlabelled dATP, dTTP and dGTP, was added and 1μl of Klenow polymerase prior to the addition

of 5μl of ^{32}P α-dCTP (MP Biomedicals). The reaction mixture was then incubated for 30 to 90 minutes at 37°C depending upon the size of the probe used (c.300bp to 1Kb).

The labelled probe DNA was separated from the unincorporated ^{32}P α-dCTP nucleotides using a Sephadex (G50) column prepared in a 1mL syringe fitted with 3MM™ Whatman paper disc inside the syringe at its base. The reaction mixture was added to the column with 30μl of dye (1.5% dextran blue and 0.5% Orange-G) and the DNA was eluted with a gravity flow of TE buffer. The blue dye elutes first along with the radio-labelled probe DNA, whilst the unincorporated ^{32}P α-dCTP nucleotides are eluted second of all with the Orange G dye. Each elution layer was collected in a 1.5mL screw-cap eppendorf. The blue elution layer containing the labelled DNA probe was boiled with 500μl of herring sperm DNA (10mg/ml; previously sheared by sonication) for 5 minutes prior to addition of the probe to the pre-hybridised nitrocellulose/nylon membrane.

2.5.3 Probing nitrocellulose/nylon membranes

Nitrocellulose or Nylon membranes obtained from Southern blotting (section 2.5.1) were placed in an appropriately sized plastic box with a tight fitting lid or rotisserie hybridisation tube and incubated with 20mμl pre-hybridisation buffer containing 500μl of herring sperm DNA (10mg/ml; previously sheared by sonication) at 65°C , or other suitably stringent temperature, for 2 hours in a rotisserie oven (Techne). The pre-hybridisation buffer was then replaced with 10-20mL of pre-heated hybridisation buffer containing the probe and the membrane was incubated at 42-65°C in a low/high stringency buffer overnight. After probing the membrane was washed 2-4 times with preheated wash solution (high/low stringency) for 5 to 20 minutes, until a suitable level of radioactivity was detectable.

2.5.4 Detection of radioactive probe

Following probing and washing (section 2.5.3), the membrane was allowed to air-dry and subsequently attached to a 3MM™ Whatman paper backing, covered with saran wrap and placed securely in a phosphor imager cassette at room temperature for 24-36 hours or over a period of several days at 4°C, depending on the level of radioactivity. Phosphor imager plates were read using a Typhoon scanner (Molecular Dynamics, Amersham Pharmacia Biotech Ltd).

2.6 Gene Expression Analysis

A variety of techniques including Reverse Transcription PCR (RT-PCR; section 2.6.2), Quantitative Real Time PCR (QRT-PCR; section 2.6.3) and *in situ* hybridisation (section 2.6.4) were utilised to identify candidate spur-related genes in each of the taxa studied, based on determination of their expression patterns.

2.6.1 Obtaining positive control genes for RT and QRT-PCR

A number of novel positive control gene sequences had to be obtained for use in RT and QRT-PCR analyses of *KNOX* gene expression in orchids and *L. vulgaris*.

2.6.1.1 Obtaining orchid actin

Actin sequences of approximately 823bp were successfully isolated from five orchid species; *D. viridis, D. fuchsii, G. conopsea, G. odoratissima* and *G. rhellicani*. Sequences spanned an intron and were obtained using primers designed from an alignment of *A. thaliana* (*ACT11*; U27981) gDNA sequence and *Phaelonopsis* hybrid cultivar actin (*PACT4*; AY134752) complete cDNA sequence. Specific actin sequences for each of the five orchid species were isolated from 1 to 5μl of cDNA template with the reaction conditions and thermocycling profile outlined in tables 2.1 and 2.2. Species specific PCR products were cloned, sequenced and identified using a translated nucleotide BLAST-X query against the NCBI database to confirm specific amplification of the target genes. Novel orchid actin sequences were aligned using ClustalW2 and universal orchid actin primers designed to amplify a 388bp fragment of actin common to all five Orchidinae orchid species. The primers were designed to amplify across a predicted intron region (Appendix B.1).

2.6.1.2 Obtaining *L. vulgaris* positive control genes

The nucleotide sequence of *L. vulgaris* 18S Ribosomal RNA (U38315) was obtained from the NCBI website. Primers were designed to amplify a 265bp fragment of this gene. PCR amplification was conducted with 1 to 5μl of cDNA template with the reaction conditions and thermocycling profile outlined in tables 2.1; 2.2. The PCR products were cloned, sequenced and identified using a translated nucleotide BLAST-X query against the NCBI database to confirm specific amplification of the target gene.

For QRT-PCR analysis of *L. vulgaris KNOX* gene expression an additional low abundance positive control gene was identified. Primers were designed to PCR amplify a 908bp fragment of *L. vulgaris* Tubulin alpha5-chain (*LvTUA5*). The primers were designed

across introns based on a ClustalW2 nucleotide alignment of several TUA5 genes previously identified from several mono/dicot species; *A. thaliana* (NM 121983), *Oryza sativa* (DQ683571), *Solanum tuberosum* (DQ294259) and *Zea mays* (NM 001111854). To establish intron/exon boundaries the gDNA sequence of *A. thaliana AtTUA5* (*AT5G19780.1*) was also included in the alignment. PCR amplification was conducted with 1 to 5μl of cDNA template with the reaction conditions and thermocycling profile outlined in tables 2.1; 2.2. The PCR products were cloned, sequenced and identified using a translated nucleotide BLAST-X query against the NCBI database to confirm specific amplification of the target gene. Specific primers were then designed for *LvTUA5* for use in QRT-PCR to generate a 358bp product.

2.6.2 Reverse Transcription PCR (RT-PCR)

Non-quantitative RT-PCR was used to assess the presence or absence of target gene expression in a number of different tissues at appropriate stages of floral development. A bank of specific RT-PCR primers (Appendix B.1) were designed to amplify c. 300-500bp of sequence for each of the *KNOX* genes obtained from dPCR (section 2.3.4) and as a result of fragment extension by RACE-PCR (section 2.3.12). Primers were designed to amplify across predicted introns to control against potential gDNA contamination. PCR was performed using 1 to 5μl of cDNA template with the reaction conditions and thermocycling profile outlined in tables 2.1; 2.2.

2.6.3 Quantitative Real Time PCR (QRT-PCR)

Quantitative Real Time PCR (QRT-PCR) was conducted on a variety of *L. vulgaris* floral/non-floral tissues to explore the precise expression patterns and levels for two *L. vulgaris KNOX* genes. QRT-PCR was carried out with the SYBR® GreenER qRT-PCR kit (Invitrogen Life Technologies, UK) according to the manufacturer's instructions, using 1μg of purified total RNA (section 2.3.2) and between 0.25-0.625μl of the subsequently derived cDNA in each 12.5μl QRT-PCR reaction, using the thermocycling profile in table 2.3. QRT-PCR was carried out using a Chromo4 Real Time Detector and DNA Engine Peltier Thermocycler and the reaction monitored using Opticon Monitor™ v3.1 software (BioRad).

To accurately quantify fold changes in expression between *LvHirz* and *LvIna* in different tissues standard curves were prepared using a 2-fold serial dilution series of cDNA from developing floral buds in order to determine the reaction efficiencies for each of the primer sets LvHirz, LvIna and LvTUA5 (Appendix B.1) according to the guidelines outlined by

Pfaffl (2001, 2004). For each dilution series three technical replicates were prepared and the average C(t) value (cycle number at which the fluorescence threshold was reached) plotted on a log-linear scale against the starting concentration of cDNA template added in each reaction. Linear regression was performed on the results for each primer set and the slope of the regression line calculated for the LvHirz (Equations 2.2), LvIna (2.3) and LvTUA5 (2.4), see Figure 2.5).

Using a simple calculation outlined in Pfaffl (2004), the LvHirz primer set had an estimated reaction efficiency of 89.4%, LvIna 63.5% and LvTUA5 78.6%. While these QRT-PCR reaction efficiencies are highly consistent, to compare two or more different genes with differing reaction efficiencies it was necessary to perform an efficiency correction calculation. There are a plethora of methodologies for applying efficiency correction (EC) to QRT-PCR data (Pfaffl, 2004) but one of the most straightforward is to make a simple modification to the $2^{-\text{delta, delta Ct}}$ method (Equations 2.5, 2.6).

In order to apply the actual reaction efficiencies derived from a serial dilution series a quantity (X) for each gene can be calculated using equation 2.7 from an average Ct value generated from three technical replicates and the gradient of the standard curve (N). This calculation must be made for both the target and control genes and the difference in quantity calculated generating an expression value of the target gene as a proportion of the control (Y; Equation 2.8). Quantifying expression relative to an internal control is necessary to make comparisons of transcript abundance between different tissues, treatments or to compare the transcript abundance of different genes. Using this method the expression of *LvHirz* and *LvIna* was measured relative to *LvTUA5* from three independent biological replicates for a range of different tissues.

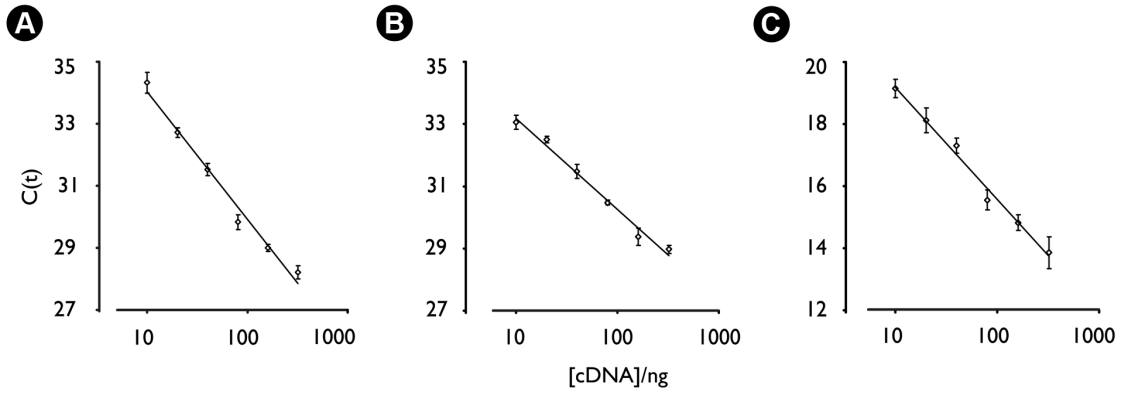


Figure 2.5: Standard curves for *LvHirz*, *LvIna* and *LvTUA5* QRT-PCR primer sets
Standard curves for each of the *LvHirz*, *LvIna* and *LvTUA5* QRT-PCR primer sets were prepared using a 2-fold dilution series of cDNA starting template. Each dilution is represented as the average Ct value of three technical replicates plotted against the starting cDNA concentration on a log-linear scale. A regression line was fitted to each data series to calculate the gradient of the slope (see equations 2.2 to 2.4). A. *LvHirz*, $R^2 = 0.98$. B. *LvIna*, $R^2 = 0.99$. C. *LvTUA5*, $R^2 = 0.99$.

$$\text{y} = \text{-}1.787\text{Ln(x)} + 38.146 \tag{2.2}$$

$$\text{y} = \text{-}1.2701\text{Ln(x)} + 36.103 \tag{2.3}$$

$$\text{y} = \text{-}1.5719\text{Ln(x)} + 22.809 \tag{2.4}$$

Table 2.3: QRT-PCR thermocycle

QRT-PCR Thermocycling profile			
Phase	*Temp (°C)*	*Time (mins)*	*Number cycles*
UDG incubation	50	2	1
Initial denaturation	95	10	1
Denaturation	95	0.15	
Primer annealing/extension	60	1	40

$$\Delta \text{ Ct} = \text{Ct[control]} - \text{Ct[target]} \tag{2.5}$$

$$\text{relative expression} = 2^{-\Delta \text{ Ct}} \tag{2.6}$$

$$EC \text{ quantity (X)} = \text{N}^{-\text{averageCt}} \tag{2.7}$$

$$EC \text{ relative expression (Y)} = \frac{\text{X[target]}}{\text{X[control]}} \tag{2.8}$$

2.6.4 *In situ* hybridisation (ISH)

In situ hybridisation is a powerful technique for determining spatial patterns of mess anger RNA localization. It is the only high resolution method of localising the site of endogenous gene expression at the tissue/cellular level. Although ISH was attempted no signal was detectable, as a result no ISH results are presented in this thesis and therefore the method used will not be elaborated on in this chapter. However, the ISH protocol employed was adapted from that used in the Langdale (University of Oxford) and Byrne (John Innes Centre) labs and subsequently modified by myself.

2.7 Constitutive expression analysis in heterologous hosts

To assess potential protein function full-length protein coding sequences of candidate *KNOX* genes were cloned into the pGreenII0029::35S constitutive expression vector (generated by Wilkins (2004)) and transferred to *Agrobacterium tumefaciens* strain GV3101 prior to transformation of *Nicotiana tabacum* cv. Samsun with the desired construct.

2.7.1 Generating constructs for Tobacco transformation

Full-length protein coding sequences were cloned into the pGreenII0029::35S constitutive expression vector which contains a 35S cassette consisting of a multiple cloning site flanked by a strongly active double CaMV 35S promoter and a single CaMV 35S terminator (Figure 2.2). The vector also contains genes for plant and bacterial-expressed Kanamycin resistance.

The complete full-length protein coding sequences of several *KNOX* gene candidates were cloned directly into pGreenII0029::35S via the addition of suitable restriction sites to the ends of specific full-length gene specific primers. Individual cloning strategies were designed for each of the candidate *KNOX* genes but all involved PCR amplification with full-length primers modified with appropriate restriction sites. The full-length coding sequences of *AmHIRZINA*, *AmINVAGINATA* and *LvHIRZINA* were cloned directly using primers modified with the restriction sites HindIII and BamHI. The full-length coding sequence of *LvINVAGINATA* was cloned directly using PstI and BamHI modified primers. Primers carrying the EcoRI restriction site were used to clone the orchid gene *DfKNOX2* in a non-directional manner. In each case the pGreenII0029::35S vector was digested with the same enzymes as the intended insert that allowed directional ligation into pGreenII0029::35S.

Digested insert and vector were run on an agarose gel (section 2.3.6) and the desired DNA fragments excised from the gel and/or column purified (sections 2.3.7, 2.3.8). Purified insert and vector were ligated using T4 DNA ligase. Each 10μl reaction contained: 1μl 10X T4 DNA Ligase buffer, 1μl T4 DNA ligase, 1μl 50ng/μl vector, Xμl insert such that there was a 5:1 ratio of insert:vector (made up to 10μl with sterile dH_20). The ligation was incubated overnight at room temperature and then transferred to competent *E. coli* DH5α (section 2.3.10). The presence of the desired insert was confirmed by CPCR/restriction digests and DNA sequencing (section 2.3.11).

2.7.2 Production of electrocompetent *Agrobacterium tumefaciens* strain GV3101

To prepare *Agrobacterium tumefaciens* strain GV3101 cells for electroporation a single colony was grown up overnight in 100mL of LB with 50μg/mL rifampicin (in 50% methanol), 25μg/mL gentamycin, and 10μg/mL tetracycline (in 95% ethanol) in a 30°C shaking incubator. The culture was transferred to two sterile falcon tubes and the cells spun down for 20 minutes at 5,500 x g at 4°C. The pelleted cells were then resuspended in 50mL sterile 10% glycerol or 1mM sterile Hepes (pH 7) and the cells spun down as before. This process was repeated three times, after the fourth spin the cells were resuspended in 1mL of 10% glycerol and divided into aliquots of 80μl, then snap frozen in liquid Nitrogen and stored at -80°C.

2.7.3 Transformation of *Agrobacterium tumefaciens* strain GV3101 by electroporation

Transformation of constructs into *Agrobacterium tumefaciens* strain GV3101 was conducted by electroporation (Mattanovich *et al.*, 1989) at 1.8Kvolts. Strain GV3101 is appropriate for transformation of tobacco and carries a chromosomal rifampicin resistance gene, but is sensitive to kanamycin selection, making it ideal for use with binary vectors that confer kanamycin resistance such as pGreen (Hellens *et al.*, 2000). It contains a disarmed Ti plasmid (pMP90) that possesses the virulence gene needed for T-DNA (transferred gene) insertion and gentamycin resistance, but has no functional T-DNA region of its own. The GV3101 strain used in the lab also contains the helper plasmid pSOUP that conveys tetracycline resistance and provides the machinery for the pGreen plasmid to replicate in *Agrobacterium*.

Transformation of constructs (section 2.7.1) into electrocompetent *Agrobacterium* was achieved using a chilled electroporation cuvette (Bio-Rad, Chicago, IL, USA) kept on ice,

to which 1μg of plasmid DNA was added to an 80μl aliquot of electrocompetent cells (section 2.7.2), mixed gently and transferred to the cuvette. After electroporation 500μl of SOC medium was added to the cuvette and transferred to a microcentrifuge. The transformed cells were allowed to recover by incubation at 30°C in a shaking incubator for 3 to 4 hours after which the culture was plated out onto two LB 1% agarose plates with 25μg/mL gentamycin and 50μg/mL kanamycin in a 30°C oven for 48 hours. 10μg/mL Tetracycline may also be included to retain the pSOUP helper plasmid but as pGreen will not replicate in the absence of pSOUP Tetracycline can be omitted.

Single colonies were screened for the construct by colony PCR (section 2.3.11.2) and/or restriction digest (section 2.3.11.1). Those colonies shown to have been successfully transformed were subsequently grown in 5mL of LB broth plus 25μg/mL gentamycin and 50μg/mL kanamycin in a 30°C shaking incubator overnight in preparation for tobacco leaf disc transformation (section 2.7.4).

2.7.4 Transformation of tobacco leaf discs

Leaf segments of tobacco var. Samsun were transformed using a modified protocol from Horsch *et al.* (1985). Several young tobacco leaves were harvested and immersed in 5% commercial bleach for 15 minutes, then rinsed several times in sterile distilled water. The leaves were cut into squares approximately 1cm^2 in a petri dish containing sterile distilled water. Leaf segments were then blotted dry and placed adaxially onto Murashige and Skoog regeneration media (MS9). This step ensured maximum wound response from the leaves which produce the phenolic acetosyringone after wounding. This compound is recognised by *A. tumefaciens* and induces transcription of virulence genes facilitating efficient transformation. After 24 hours pre-incubation the leaf segments were then transferred to a petri dish containing 25mL of a 100mL overnight culture of transformed *Agrobacterium* (section 2.7.3), each 25ml of culture is sufficient to transform leaf discs from two small leaves, and transferred to fresh MS9 plates for a co-cultivation period in the dark at 25°C for 48 hours. Leaf segments from approximately eight small tobacco leaves could be transformed in this way.

Agrobacterium was then removed by transferring the tobacco leaf segments to fresh MS9 media containing 50μg/mL kanamycin to select for transformants and a mixture of 100μg/mL ampicillin, to retard general microbial growth, and 250μg/mL cefotaxime to inhibit growth of *Agrobacterium* itself. Leaf segments were incubated at 25°C in 16 hours light, 8 hours darkness and transferred to fresh selective MS9 plates every 10 to 12 days and allowed to callus. Once the callus regenerated shoots these were transferred to selective root initiation media (MS0). The transformation status of rooted plants was

subsequently checked by PCR using genomic DNA template (section 2.3.1.1) before the plants were transferred to a 3:1 mixture of potting compost and medium grade vermiculite and allowed to grow to maturity in a phytotron (section 2.1.5). Expression of the transgene in target tissues was further characterised by RT-PCR (section 2.6.2). Gross morphological phenotypes were recorded photographically.

2.8 Viral Induced Gene Silencing of *Linaria vulgaris* (VIGS assay)

Virus vectors carrying host-derived sequence inserts induce silencing of the corresponding genes in infected plants. Virus Induced Gene Silencing (VIGS) is a manifestation of an RNA-mediated defense mechanism that is related to post-transcriptional gene silencing (PTGS) in transgenic plants. The RNA-mediated defence is triggered by modified virus vectors carrying specific gene sequences and targets both the viral genome, as with wild type viruses, and the host genome corresponding to the target gene insert. As a result the symptoms phenocopy loss of function or reduced expression mutants in the host gene.

Silencing of candidate *KNOX* genes in *L. vulgaris* was attempted using the tobacco rattle virus (TRV) VIGS vector system (Ratcliff *et al.*, 2001). However, although constructs were generated for TRV-VIGS in *L. vulgaris* it was not possible to optimise the TRV-VIGS protocol and achieve silencing. As a result no TRV-VIGS results are presented in this thesis and only elements of the protocol relevant to generating the silencing constructs are described.

2.8.1 The TRV-VIGS vector system

TRV is a positive-strand RNA virus with a bipartite genome consisting of RNA1 and RNA2 (Figure 2.6). The TRV-VIGS vector system therefore consists of two plant Ti binary vectors, pBinTRA6 and pTV00. pBinTRA6 contains a full-length infectious clone of TRV RNA1 (strain PPK20) cloned into the vector pBin61 (Bendahmane *et al.*, 2000) which is a derivative of pBin19, between a single CaMV 35S promoter and terminator (Ratcliff *et al.*, 2001). Proteins encoded by RNA1 are sufficient for replication and movement in the host plant. In the pBinTRA6 vector the RNA polymerase ORF had been interrupted by intron 3 of the nitrate reductase *NIA1* gene to improve stability of RNA1 in *E. coli* (Wilkinson & Crawford, 1993).

pTV00 was derived from the previously described TRV RNA2 vector pCaK20-2T7 (Hernandez *et al.*, 1995) and contains a modified clone of TRV RNA2 which encodes pro-

teins for virion formation and nematode-mediated transmission between plants (Ratcliff *et al.*, 2001). The non essential 29.4K and 32.8K proteins have been removed and replaced with a MCS leaving only the 5′ and 3′ untranslated regions of the viral coat protein flanked by a single CaMV 35S promoter and nopaline synthase NoS terminator. This modified TRV RNA2 has been subsequently cloned into the plant binary transformation vector pGreen0000 (see Ratcliff *et al.* (2001) for further details). The TRV VIGS vectors used were kindly provided by Prof. David Baulcombe (University of Cambridge).

2.8.2 Propagation of TRV-VIGS constructs in *E. coli* and *A. tumefaciens*

For propagation of pBinTRA6 and pTV00 constructs in *E. coli* strain DH5α was used, as pBinTRA6 is unstable in most strains of *E. coli* (Lu *et al.*, 2003). The constructs were selected with 50μg/mL Kanamycin in LB media. For propagation of pBinTra6 the *A. tumefaciens* strain C58C1 carrying the pCH32 helper (Tetracycline resistance) was used, for pTV00 the *A. tumefaciens* strain GV3101 carrying the pSoup helper was used. Selection in *A. tumefaciens* was conferred with a combination of 50μg/mL Kanamycin, 5μg/mL Tetracycline and 50μg/mL Rifampicin (may be excluded) in LB media (Figure 2.6).

2.8.3 Generating constructs for TRV-VIGS

Several *L. vulgaris KNOX* gene constructs were generated for VIGS including a positive control construct carrying a fragment of the endogenous phytoene desaturase gene (PDS).

A fragment of the endogenous *PDS* gene was PCR amplified from *L. vulgaris* cDNA using primers originally developed to amplify *PDS* from *N. benthamiana* (Ratcliff *et al.*, 2001). The primer sequences were modified by the addition of BamHI and KpnI primer sites (Appendix B.1) to facilitate direct cloning into the pTV00 vector. Three additional *L. vulgaris KNOX* gene constructs were also generated for VIGS. Full-length *LvHIRZINA* was PCR amplified from *L. vulgaris* cDNA using modified primers carrying the restriction sites BamHI and HindIII, this construct carries the highly conserved homeodomain and was developed to silence all *KNOX* gene activity. To generate specific silencing smaller fragments of *LvHIRZINA* and *LvINVAGINATA* were PCR amplified from *L. vulgaris* cDNA each consisting of the less well conserved meinox domain. Primers were modified with the restriction sites BamHI and HindIII for cloning of the meinox domain of

LvHIRZINA and BamHI/KpnI for cloning the meinox domain of *LvINVAGINATA* (Appendix B.1). In each case the pTV00 plasmid was digested with the appropriate restriction enzymes (Figure 2.6). Digested insert and vector were run on an agarose gel (section 2.3.6) and the desired DNA fragments excised from the gel and column purified (section 2.3.7). Purified insert and vector were ligated using T4 DNA ligase (section 2.3.9). The presence of the desired insert was confirmed by restriction digest using the same enzymes implemented in the cloning steps and DNA sequencing (section 2.3.11). The presence of pBinTRA6 was confirmed by digest with the restriction enzyme BamHI only.

2.8.4 VIGS assay

The TRV-VIGS assay was conducted using a protocol derived from the lab of Elena Kramer (Harvard University, USA) but could not be optimised. The protocol is available on-line: http://www.oeb.harvard.edu/faculty/kramer/Site/Protocols.html

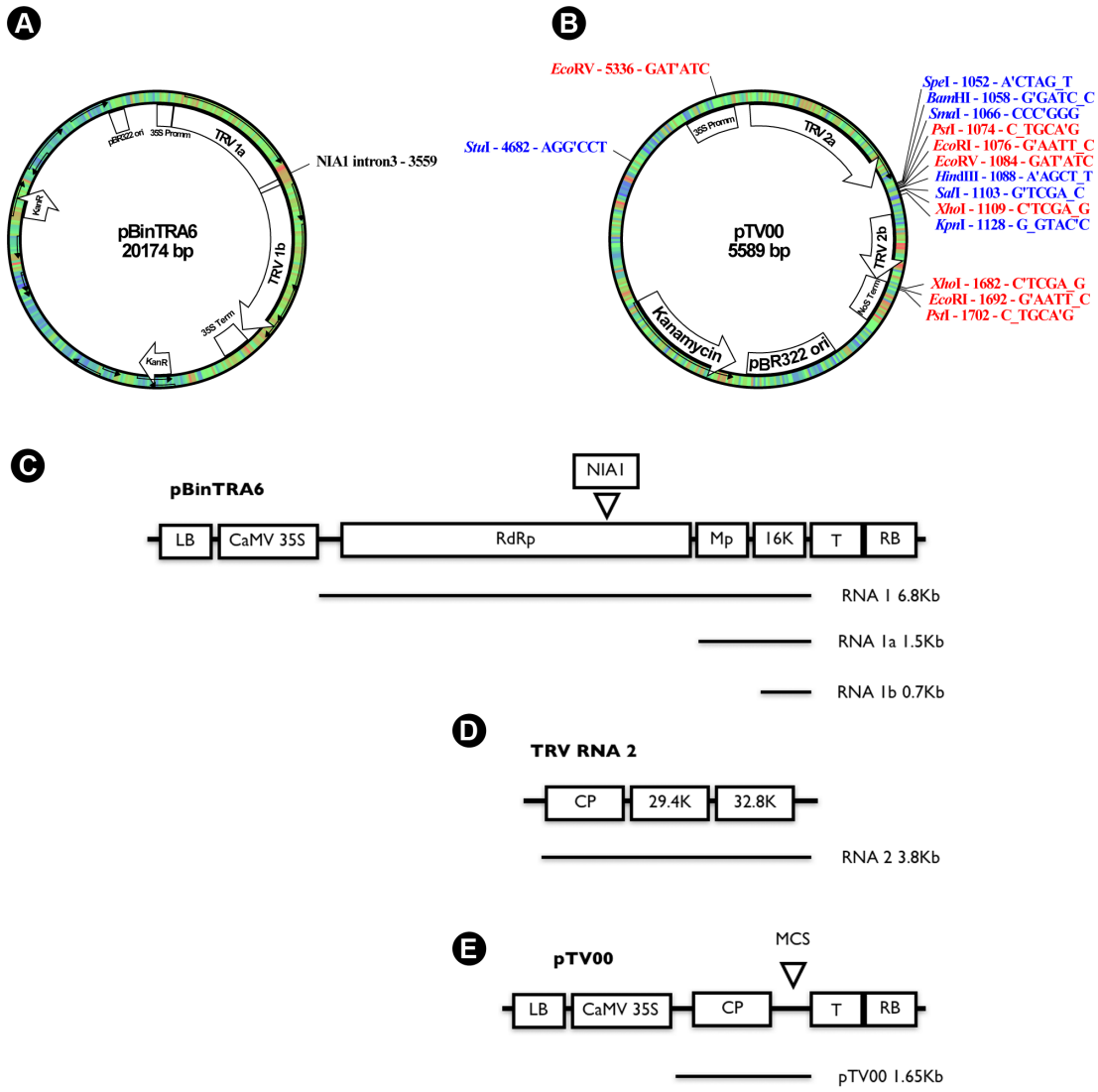


Figure 2.6: The TRV-VIGS vector system

A, B. TRV-VIGS vector maps. Open reading frames are indicated by arrows on the plasmid backbone. Origins, promoters, terminators, operons and antibiotic resistance genes are indicated by block arrows. Primer annealing sites, restriction sites and multiple cloning sites are indicated by tags outside of the DNA backbone. A. pBinTra6. B. pTV00. C-E. Schematic representation of the genomic organization of wild type TRV RNAs 1 and 2 and the modifications made during the development of the TRV-VIGS vector system (Re-drawn from Ratcliff *et al.* (2001)). C. The T-DNA organization of pBinTRA6, a plant binary transformation vector containing a full-length clone of TRV RNA 1. D. Genomic organization of wild type TRV RNA2. E. T-DNA organization of pTV00, a plant binary transformation vector containing a full-length clone derived from TRV RNA 2. The clones are positioned between the left and right borders (LB and RB) of the T-DNA and the CaMV35S promoters and transcriptional terminators (T). Open reading frames correspond to the RNA-dependent RNA polymerase (RdRp), movement protein (MP), 16K protein (16K), coat protein (CP) and the 29.4 and 32.8K proteins. The NIA1 intron, multiple cloning site (MCS) and sizes of viral genomic and sub-genomic RNA species are also shown.

Chapter 3

Morphological and ontogenetic analysis of floral nectar spur development in *Linaria vulgaris*, *Dactylorhiza fuchsii* and *D. viridis*

3.1 Introduction

The nectar spur is a structure of acknowledged function(s) that is known to influence speciation in some groups, is unusually evolutionarily labile (in terms of both acquisition and loss), commonly subjected to strong directional and/or disruptive selection, and is frequently discussed in the literature as a key innovation and an example of parallel and/or convergent evolution that has played a critical role in determining close co-evolutionary relationships with specific plant pollinators. Investigating the genetic determinants of such an important floral innovation in an evolutionary-developmental context requires an intimate knowledge of the morphology of that structure and the sequence of events that occur during its morphogenesis (ontogeny). Such an understanding provides essential clues about the types of genes that may be involved and the probable times at which they are expressed. In terms of phylogeny and evolution, considerable morphological diversity can be generated as a result of changes in the timing (heterochrony) and/or location (heterotopy) of organ development. Understanding the morphology and ontogeny of a structure is particularly important in this regard, as often ontogeny is the most reliable indicator of homology. In the absence of an ontogenetic framework it is difficult to understand the broader significance of focused insights gleaned from developmental genetics.

The eudicot *Linaria vulgaris*, a close relative of the developmental model *A. majus*, and the orchids *Dactylorhiza fuchsii* and *D. viridis* have been identified as candidates to develop an understanding of the genetic determinants of nectar spur ontogeny and morphological diversification. Despite long-term use as a developmental-genetic model, reports of floral ontogeny in the *Antirrhinum* and *Linaria* literature are sparse. A detailed account of floral ontogeny in snapdragon has only recently been published (Vincent & Coen, 2004). Likewise, reports of floral ontogeny in *L. vulgaris* are restricted only to early events, associated with work aimed at understanding floral zygomorphy (Cubas *et al.*, 1999). However, many key floral features arise at later stages of development, such as organ shape, size and complex floral elaborations including nectar spurs. Furthermore, it is variation in these features that provides much of the diversity in floral form. Although floral ontogeny in *L. purpurea* has been characterised in some detail (P.J. Rudall, unpublished), despite the emphasis placed upon five-spurred actinomorphic *L. vulgaris* mutants in understanding genetic determinants of floral zygomorphy, there are currently no reports of nectar spur ontogeny for *L. vulgaris* in the literature.

Developmental studies of nectar spur formation are few and far between, particularly among eudicots. Tepfer (1953) studied patterns of spur development in *Aquilegia formosa* and noted that the spur was among the last of the floral organs to develop in the ontogeny of the flower, almost as if it had been added as an innovation to the end of floral ontogeny. Gottlieb (1984) and other authors have noted the same late development of the spur in *Aquilegia*. More recently, nectar spur development has been more fully characterised using scanning electron microscopy in a comparative developmental context, using spurred *Aquilegia olympica* and non-spurred *A. ecalcarata* and *Semiaquilegia adoxoides*, all Ranunculaceae (Tucker & Hodges, 2005). In *Aquilegia* the nectar spur starts development as an abaxial bulge late in petal ontogeny (Tucker & Hodges, 2005).

For orchids, detailed ontogenetic and developmental-genetic investigations have only recently been initiated on a large scale (Kurzweil, 1998; Mondragón-Palomino & Theißen, 2008). However, well documented though it is for a number of orchid species, only early stages of floral development have been considered and vital data pertaining to nectar spur development are lacking in the published literature. In related work, initiated at the Royal Botanic Gardens, Kew (published alongside results reported in this chapter by Box *et al.* (2008)), detailed comparative developmental studies have been conducted between long and short-spurred species in the *Gymnadenia* clade. *Gymnadenia conopsea* and *G. odoratissima* have long/intermediate length nectar spurs (the ancestral condition) whilst *G. austriaca* has a short nectar spur (the derived condition). Nectar spur development in *Gymnadenia* has been shown to occur in a similar manner to that described by Tucker

& Hodges (2005). Furthermore, the short spurs of *G. austriaca* are the result of early cessation of nectar spur development (Box *et al.*, 2008), producing a shortened nectar spur representative of earlier stages of nectar spur ontogeny in ancestral species, heterochrony (see section 3.4.2), a phenomenon that is also apparent in other parts of the *G. austriaca* flower.

3.1.1 Research aims and objectives

Floral morphology and ontogeny will be examined in the eudicot *L. vulgaris* and the closely related orchids *D. fuchsii* and *D. viridis*, with particular focus on the nectar spur. A range of imaging techniques will be employed including light/electron microscopy and X-ray computed tomography, in order to develop an understanding of floral morphology, nectar spur ontogeny, and to identify key stages of morphogenesis that can be targeted for subsequent developmental-genetic analyses, e.g. candidate gene isolation and gene expression analysis. Prior to discussing the results presented in this chapter a brief introduction to the study taxa is presented below.

3.1.2 *Linaria vulgaris* (Antirrhineae, Plantaginaceae, Lamiales)

Common toadflax, *Linaria vulgaris* (L.) Mill. (Figure 3.1), is a hardy, perennial, clonal herb commonly found along roadside verges, railway tracks and uncultivated sites throughout most of Europe, northern Asia and, after introduction from Europe in the 1700s, much of North America (Arnold, 1982; Blamey & Grey-Wilson, 1989; Saner *et al.*, 1995; Takhtajan, 2009). It has short, spreading roots, erect to decumbent stems 15-90cm high, with fine, threadlike, glaucous blue-green leaves 2-6cm long and 1-2 mm broad. The flowers of *L. vulgaris* are similar to those of the snapdragon, 25-33mm long, pale yellow except for the lower lip which is bright yellow/orange. Ramets bear several racemose inflorescences with numerous yellow flowers (15-20), each with a nectar spur 15-20mm long. *Linaria vulgaris* reproduces clonally by aggressive rhizomatous growth and sexually by prolific seed production (Newman & Thomson, 2005), as such many regard toadflax as a weed, albeit an attractive one.

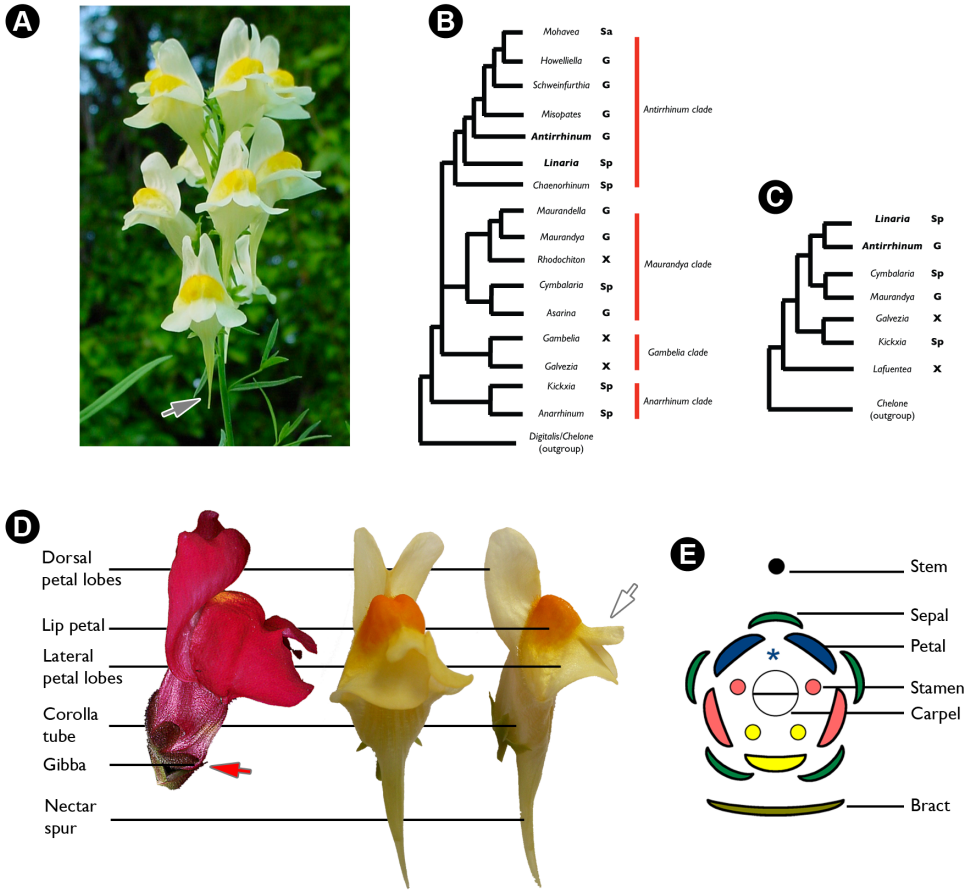


Figure 3.1: ***Linaria vulgaris***
A. Wild-type *L. vulgaris* inflorescence. B, C. Phylogenetic position of *L. vulgaris* and *Antirrhinum majus* in the Antirrhineae, corolla elaborations are marked next to genera; (B) combined morphological and molecular *ndhF* strict consensus (Ghebrehiwet *et al.*, 2000), (C) *matK-trnK* intron strict consensus tree (Albach *et al.*, 2005). D. Gross floral morphology of *L. vulgaris* compared to *A. majus* (left), red arrow indicates the snapdragon gibba located in an identical position to the nectar spur of *L. vulgaris*, white arrow indicates the ventral petal lobe. E. *L. vulgaris* floral schematic, blue indicates dorsal organs, pink lateral organs and yellow, ventral organs. The asterisk indicates the dorsal staminode. *Figure abbreviations: G - gibbous, Sa - saccate, Sp - spurred and, X - simple corolla tubes.*

Traditional circumscriptions placed the genus *Linaria* in the family Scrophulariaceae along with *Antirrhinum*, which is morphologically similar. More recently however, combined molecular and morphological phylogenetic analyses have shown that Scrophulariacae *sensu lato* is not a monophyletic family (Olmstead & Reeves, 1995; Olmstead *et al.*, 2001; Oxelman *et al.*, 1999, 2005). Consequent re-circumscription of the Scrophulariaceae has led to the re-distribution of some of the genera to other previously recognised families, and the adoption of novel nomenclature to describe newly identified assemblages. For example, the names Antirrhinaceae (Reeves & Olmstead, 1998; Reveal *et al.*, 1999) and Veronicaceae (Olmstead *et al.*, 2001) have both been proposed to identify the monophyletic clade containing *Linaria* and *Antirrhinum*. However, as a number of authors still uphold previous circumscriptions and since this taxonomic debate remains unresolved, to avoid confusion, the oldest conserved name, Plantaginaceae, is recognised by APGIII (Bremer *et al.*, 2009) and will be used in this thesis.

Within Plantaginaceae, *Linaria* and *Antirrhinum* are closely related genera in the tribe Antirrhineae (Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004). Antirrhineae has remained largely unaffected by taxonomic re-circumscription of Scrophulariaceae and the revision of the tribe by Sutton (1988). However, the supra-generic relationships within Antirrhineae (Figure 3.1B, C) have been the subject of numerous publications and remain uncertain (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004). Members of the tribe are characterised, most notably, by the presence of poricidal capsule dehiscence and apparently unique iridoid glycosides. Otherwise, members of the tribe are highly variable in terms of floral morphology (Sutton, 1988).

Floral structure in Antirrhineae is very diverse but is typical of other eudicots in terms of gross composition, with four clearly defined whorls of floral organs (sepals, petals, stamens and carpels) each consisting of up to five parts. Often the corolla is tubular, with a basal appendix (spurred, gibbous or saccate) (Sutton, 1988). The corolla is highly variable, including both zygomorphic and actinomorphic forms that can be broad, narrow, flaring or personate as is the case for *Linaria* and *Antirrhinum* (Figure 3.1).

Most Antirrhineae, including *Antirrhinum* and *Linaria*, are self-incompatible and dependant on insect pollinators in the wild, although laboratory strains are mostly self-compatible. The flowers produce a nectar reward, from a nectary encircling the base of the carpel, that accumulates in the gibbous basal appendix of *Antirrhinum* and the long, narrow, floral nectar spur of *L. vulgaris* (Elisens & Freeman, 1988; Sutton, 1988; Vogel, 1998). Pollination is effected by long-tongued bumblebees (melittophily) that enter the tubular corolla by manipulation of the palate. As the nectar reward is located at the

very base of the gibbous/spurred corolla tube, pollen is brushed onto the dorsal side of the bee as it enters the flower and probes for nectar (nototribic pollination) (Elisens & Freeman, 1988; Macior, 1967; Nepi *et al.*, 2003; Newman & Thomson, 2005; Stout *et al.*, 2000; Sutton, 1988).

3.1.3 *Dactylorhiza fuchsii* and *D. viridis* (Orchidinae, Orchidaceae, Asparagales)

Dactylorhiza is a Eurasian genus consisting of approximately 60 species of densely flowered, small to large, tuber forming, winter dormant perennial herbs distributed principally in the boreal and temperate zones (Delforge, 2006). Most grow in full sun or shade on dry to wet, alkaline substrates in short grassland, marshes, woodland edges, secondary woodland and forests. Surprisingly, many *Dactylorhiza spp.*, including *D. fuchsii*, are common in disturbed sites such as roadside verges.

The common spotted orchid, *D. fuchsii* (Figure 3.2), is a slender plant, 15-70cm in height characterised by densely spotted cauline leaves, common throughout much of south-east England such as the Chiltern Hills, Hertfordshire, Cambridgeshire and much of Northern Europe. The inflorescence is compact to lax, conical to sub-cylindrical, 4-10cm in height bearing 15-20 medium sized flowers, often purple in colouration with a broad, near cylindrical nectar spur 6-10mm in length (Delforge, 2006). The frog orchid, *D. viridis* (Figure 3.2), is somewhat smaller in stature, ranging in height from 5-40cm, with unspotted leaves. It has become relatively uncommon in much of the UK but can still be found in small populations in the Chiltern Hills. However, in Austria, *D. viridis* remains relatively common at altitudes between 1,800 and 2,000m, whereas *D. fuchsii* appears to be more common at lower altitudes (Box, personal observation). The inflorescence is sub-lax and cylindrical, bearing approximately 25-30 medium sized, inconspicuous green flowers with a short, bulbous nectar spur, 2-3mm in length. Each flower is subtended by a bract equalling or exceeding the flower with a red-brown coloured margin (Delforge, 2006).

Both *D. fuchsii* and *D. viridis* are members of the Orchid subtribe Orchidinae (Orchidoideae, Orchidaceae) *sensu* Bateman *et al.* (2003). Although the circumscription of Orchidinae is contentious and not universally recognised, it will be used throughout this thesis in accordance with the classification of Bateman *et al.* (2003). The subtribe is exceptionally well characterised from a phylogenetic perspective (Figure 3.2). Almost all of the species have been extensively sampled for morphological (Burns-Balogh & Funk, 1986; Dressler, 1993; Freudenstein & Rasmussen, 1999) and repeated nuclear *ITS* rDNA

phylogeny to include representative DNA sequences from 190 species (Bateman, 2001; Bateman *et al.*, 1997, 2003). Although intra-generic relationships within Orchidinae are exceptionally well resolved, uncertainty of species relationships in *Dactylorhiza* make it difficult to determine the precise phylogenetic relationships between *D. fuchsii* and *D. viridis*. What is clear from the phylogenetic analysis of Bateman *et al.* (2003), is that the medium-length nectar spur of *D. fuchsii* is representative of the ancestral condition in this group and that the short spur of *D. viridis* is apparently a derived feature (Figure 3.2). Although not sister species, the closely related nature of *D. fuchsii* and *D. viridis* and the confident ancestral status of longer nectar spurs, allow valid evolutionary-developmental comparisons to be made between these species.

Detailed reports of pollination biology in *Dactylorhiza* are few and far between and provide little conclusive proof for any suggested pollinator. However, some authors e.g. Scopece *et al.* (2007); van der Cingel (1995), have suggested that *Dactylorhiza spp.* are pollinated predominantly by long-tongued bees as they probe the longer nectar spur that characterises most *Dactylorhiza spp.*, for a reward. Whilst this is consistent with personal observations of flower visitation during field collection, other authors such as Gutowski (1990), have observed pollination of *D. fuchsii* by male Longhorn beetles (*Alostera tabacicolor*). Scopece *et al.* (2007) also recognised diverse pollinator visitation among *Dactylorhiza spp.* including representatives from the insect orders Coleoptera, Diptera and Lepidoptera, however by far the most frequent visitors were long-tongued bees. Morphological assessments of *Dactylorhiza spp.* (Bell *et al.*, 2009), suggest that whilst the pollinators probe deeply for a nectar reward, the flowers of most *Dactylorhiza spp.* do not secrete nectar and appear to attract pollinators by food deception, in which the orchid promises (by possession of a floral nectar spur), but fails to deliver, a nectar reward (Schlüter & Schiestl, 2008; Scopece *et al.*, 2007). Such a pollination mechanism is common among members of the subtribe Orchidinae (Cozzolino & Widmer, 2005; Schiestl, 2005; Scopece *et al.*, 2007).

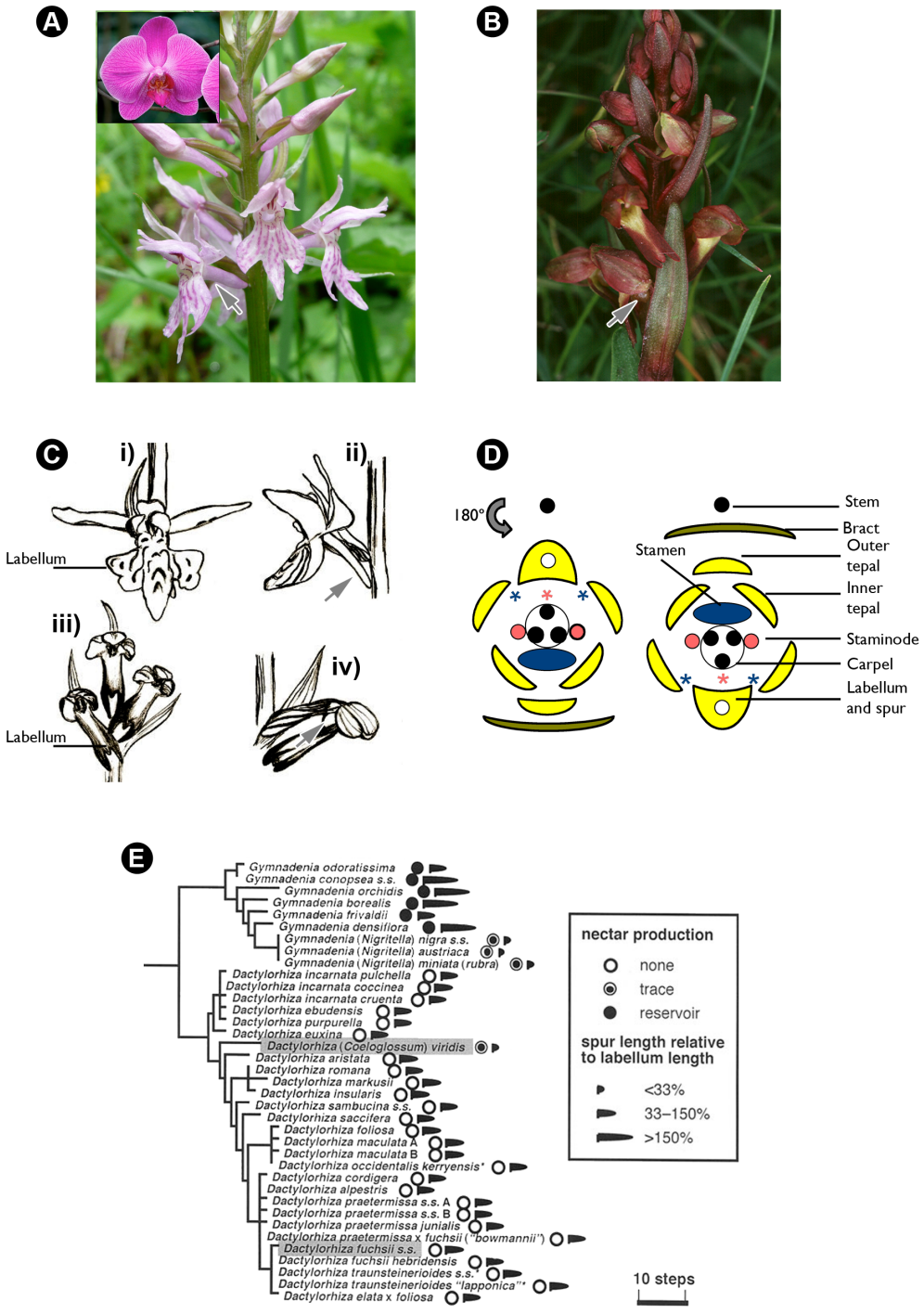


Figure 3.2: *Dactylorhiza fuchsii* and *D. viridis*
A. *Dactylorhiza fuchsii* inflorescence, with the ornamental *Phalaenopsis* inset. B. *D. viridis* inflorescence. C. Line drawings of *D. fuchsii* (i, ii) and *D. viridis* (iii, iv) flowers, kindly provided by S. Dodsworth. D. Schematic floral diagram of a generalised orchid flower indicating the enlarged labellum petal, gynoecium, resupination of the ovary and the floral nectar spur. E. Spur and nectar characteristics in the *Dactylorhiza-Gymnadenia* clade, *ITS* phylogeny modified from Bateman *et al.* (2003) (cf. Bell *et al.* (2009); Box *et al.* (2008)). A-C, arrows indicate the nectar spur.

3.1.3.1 Orchid floral biology

Orchid floral structure is highly variable, often making it difficult to interpret (Kurzweil, 1998), compare the flowers of the ornamental orchid *Phalaenopsis* with those of *Dactylorhiza spp.* (Figure 3.2). Despite their apparent complexity, orchid flowers have a relatively simple trimerous organisation that consists of five successive floral whorls each with three segments. There are two perianth whorls (outer and inner), two staminal whorls, and one central carpel whorl. However, orchid flowers have become highly modified compared to the general trimerous condition (Dressler, 1993; Rudall & Bateman, 2002). In this sense orchid flowers are considered to be highly elaborate, specialised structures that strongly demonstrate the importance of co-evolutionary relationships between plants and pollinators. As a result, a significant amount of specialist terminology has emerged to describe the peculiarities of orchid flowers. Whilst it is not appropriate to discuss the validity of one term over another, it is necessary to present a brief overview of orchid floral structure in order to introduce terminology that will be used subsequently in this work.

The orchid flower is zygomorphic and epigynous, that is the floral parts are situated on top of the ovary. There are two whorls of three perianth parts; the outer most whorl comprises two lateral sepals and a dorsal or median sepal, usually all of similar size. The inner perianth whorl consists of two often relatively large and showy lateral petals, variously coloured and bearing scent glands (osmophores) and/or hairs, plus the distinctive labellum. The labellum is highly modified in many species, elaborately adorned with scent glands and a highly sculptured adaxial epidermis (Stpiczynska, 2001), that functions in pollinator attraction by forming a landing platform. In most orchid species the labellum is located lowermost, however the labellum is initiated uppermost and adaxial in bud, but comes to occupy its lowermost abaxial position at anthesis owed to resupination of the ovary and/or pedicel (i.e. the ovary twists through 180°; Kurzweil (1998)). In approximately half of the 20,000 species of orchid, including most Orchidinae, the labellum is associated with a saccate to filiform nectar spur that exhibits remarkable diversity in form (Figure 3.2).

Androecial and gynoecial parts of the orchid flower show the greatest deviation from the basic monocotyledonous condition (Dressler, 1993; Rudall & Bateman, 2002). Reduction in the number of androecial parts (Figure 3.2), commonly one or two anthers are fertile whilst the remaining stamens are non-fertile (i.e. staminodial), and fusion of the male and female organs into a single reproductive structure, the gynostemium (Szlachetko & Rutkowski, 2000), have been the predominant trends in orchid floral evolution (Kurzweil, 1998; Rudall & Bateman, 2002). The gynostemium is an immensely variable

structure. Homologies of the gynostemium parts are notoriously difficult to interpret in the absence of careful ontogenetic observations. With some debate as to its relevance this entire structure is commonly referred to as the *column*.

3.2 Results

Floral morphology and ontogeny were investigated in the eudicot *L. vulgaris* and the orchids *D. fuchsii* and *D. viridis* using a combination of microscopical techniques. Early floral ontogeny for *L. vulgaris* has been reported previously in the literature by Cubas *et al.* (1999) but has been repeated here for completeness. Floral morphology and ontogeny for the orchids *D. fuchsii* and *D. viridis*, presented in this chapter, have been published in the *Botanical Journal of the Linnean Society*, alongside similar results obtained from the closely related orchid genus *Gymnadenia*, which was the subject of a related project at RBG, Kew, see Box *et al.* (2008) and Bell *et al.* (2009).

3.2.1 Floral morphology and ontogeny in *Linaria vulgaris*

3.2.1.1 Morphology

Floral morphology in *Linaria* is highly similar to that of snapdragon. The flowers are bracteate, hermaphrodite and hypogynous. The corolla tube is formed from five proximally fused petals which are distally separate to form the characteristic two-lobed upper lip and a three-lobed lower lip (Figure 3.1). The lip complex is bi-functional, the lower lip *masks* the interior of the corolla tube (hence the term *personate* flower) limiting access to the tube and the reproductive structures within, whilst the outer surface and petal lobes of the enlarged lower lip presents an attractive outer face that acts as a landing platform for pollinators.

Within the corolla tube four stamens are inserted at its base and extend along the dorsal side of the corolla tube (Figure 3.3A). The posterior, dorsal-most, stamen is reduced to a staminode, not readily visible at maturity. The remaining four fertile stamens are arranged in two pairs; the lateral stamens are shorter and less hairy than the two ventral stamens, which have a considerably longer filament. Anthers are dithecous, introrse (turned towards the axis) and dehisce lengthwise (Figures 3.3A, B). The gynoecium is formed from two fused carpels that generate a superior ovary that is bilocular, enclosing numerous ovules on axile placentae where ovules form on carpel septa (Figure 3.3C). The style is simple and intermediate in length between the lateral and ventral stamen filaments

such that the capitate stigma is located at the centre of a reproductive complex bordered, top and bottom, by fertile anthers (Figures 3.3A, B).

Whilst the five petals are united proximally to form the corolla tube, the highly zygomorphic nature of the corolla is apparent in the petal lobes and lip complex. The two dorsal-most petals divide at the neck of the tube to form the upper lip that consists of two relatively long, strap-shaped, erect dorsal petal lobes that form a hood over the entrance to the corolla tube (Figures 3.1D, 3.3B, D, E). The lower lip of the corolla tube is coloured bright yellow/orange and consists of two proximally fused lateral petals and the ventral petal (Figure 3.3B). The ventral and lateral petals of the corolla tube divide distally to form three petal lobes. The ventral petal lobe is small and rounded with a bright yellow/orange lip, whilst the lateral petal lobes are considerably larger with a partially bright yellow/orange lip (Figure 3.1D).

The lower lip is formed as a result of curvature in the region of the common basal parts of the corolla lobe so that the lower lip is touching the upper lip, closing the entrance to the corolla tube (section 3.2.1.2). The curvature of the lower lip corresponds with the formation of a moveable extension of the ventral corolla tube, the *palate*, which becomes exposed when the flower is opened by a suitable pollinator. Opening personate flowers such as this is made possible by a springy articulated zone at the juncture of the lower lip and corolla tube, the *hinge*. Trichomes are prominent on the dorsal side of the corolla tube interior and the palate. Two highly conspicuous lines of bright yellow/orange trichomes act as nectar guides and draw the pollinator along the palate and toward the centre of the corolla tube opening (Figures 3.3B, D, E).

In *Linaria*, the base of the corolla tube is extended to form a long narrow nectar spur derived from the ventral petal (Figure 3.1D). The dorsal side of the spur interior is dappled with numerous short hairs continuous with those that surround the base of the gynoecium and becoming sparse towards the lower third of the nectar spur as it narrows towards its tip from which they are absent (Figure 3.3A). A nectary can be found at the ventral base of the gynoecium (Figure 3.3C), the secretions from which accumulate in the tip of the floral nectar spur in *Linaria*. A similar structure is present in *Antirrhinum* flowers, the secretions from which accumulate in the small gibbous swelling located in the same position as the spur of *Linaria* (Figure 3.1D).

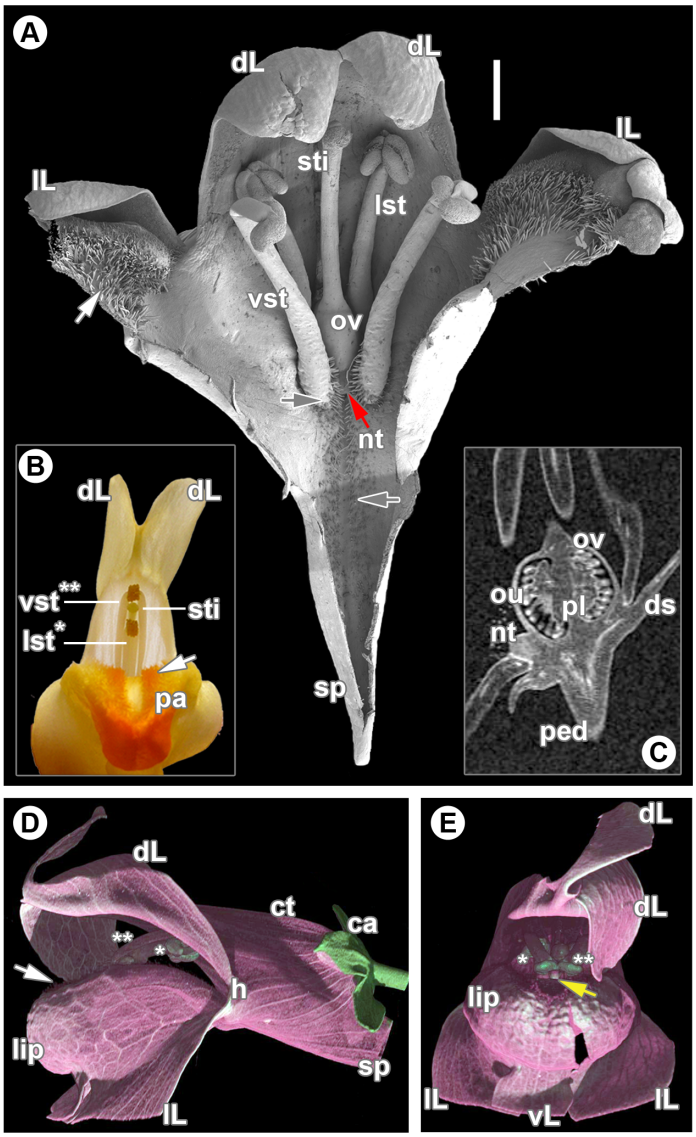


Figure 3.3: *L. vulgaris*, floral morphology (SEM, XRT)
A. Pre-anthesis *L. vulgaris* flower bud split longitudinally to reveal the internal morphology of the corolla tube and nectar spur, scale bar = 1mm. B. Opened personate flower of *L. vulgaris* demonstrating the relationship between the anthers and style at anthesis. Numerous trichomes cover the palate (white arrow), two lines of bright orange trichomes act as nectar guides. C. LS X-ray computed tomography (XRT) section of the mature ovary, clearly demonstrating the two-locular ovary with numerous ovules attached by axile placentae. D, E. Side (D) and top (E) view of a 3D-representation of the *L. vulgaris* flower reconstructed using XRT, false coloured to clarify the calyx (green) and corolla (pink) demonstrating the complex arrangement of corolla parts to form the personate flower. Note the nectar spur has been removed. *Figure abbreviations: ca - calyx, ct - corolla tube, dL, lL, vL dorsal, lateral and ventral petal lobes, ds - dorsal sepal, h - hinge, nt - nectary (red arrow), pl - placenta, sti - stigma, vst, lst - ventral and lateral stamens, ov - ovary, ou - ovule, pa - palate (trichomes, white arrow), ped - pedicel, sp - spur (papillae, grey arrow). Yellow arrow indicates the corolla tube entrance.* The asterisks ** and * in D and E replace the labels vst and lst from B.

3.2.1.2 Ontogeny

Early floral ontogeny is similar to that described for snapdragon, overall similarity in floral ontogeny between *Linaria* and *Antirrhinum* is reflected in the adoption of the ontogenetic phases described by Vincent & Coen (2004). Note, phase A pertains to initiation of the inflorescence apex rather than floral ontogeny *per se* and is not described below.

Phase B Floral zygomorphy is evident in the very earliest stages of floral ontogeny as the dome-shaped floral apex, formed in the axil of a bract, gives rise to five sepal primordia in quick succession (Figures 3.4A, B).

Phase C Following inception of the sepal primordia the floral apex becomes flattened, assuming a pentagonal shape as a result of inception of five petal primordia inner to and alternate with the sepal primordia. The growth of the petal primordia is slower than that of the sepals which will soon enclose the developing petals in bud (Figure 3.4C).

The petal primordia are shortly followed by the inception of the congenitally fused stamen primordia, inner to and alternate with the petal primordia (Figure 3.4C). The lateral pair of stamens appear to be initiated shortly before the ventral pair. The dorsal-most, fifth stamen, is formed shortly after the other four stamens and develops more slowly than the fertile stamens. At this stage a gynoecial cup primordium formed from two carpels is clearly visible in the centre of the floral apex.

Phase D Shortly after initiation of the stamen primordia, there is an extension of the bases of the fused petal primordia generating the rudiments of a corolla tube with distally free petal lobes that fold over one another, enclosing the interior reproductive parts of the flower (Figures 3.4D, E). At this stage petal growth is accelerated such that the extending corolla tube and more slowly enlarging petal lobes, catch up with the development of the sepal primordia. The dorsal petal lobes appear to develop at a greater rate than the ventral and lateral petal lobes, which remain smaller.

Phase E The beginning of the next phase of development is marked by the formation of a ventral furrow at the junction between the corolla tube and the lateral-ventral corolla lobes (Figure 3.4F). Trichome cells are clearly visible on the surface of the fused ventral and lateral petals around the ventral furrow. At the base of the corolla tube on the ventral petal, a small bulge starts to become clear, this is the nectar spur primordium. Estimates of cell sizes from SEM images indicate that cells of the nectar spur primordium

(and surrounding corolla tissues), are uniformly small and isodiametric, approximately 10 x 10μm in size (Figure 3.5A).

Phase F Enlargement of the nectar spur primordium occurs alongside continued growth of the corolla tube (Figure 3.4G). Yellow pigmentation starts to colour the developing corolla tube, distinguishing it from the green sepals of the calyx. Dividing the bud longitudinally at this stage clearly demonstrates that the initial outgrowth of the nectar spur primordium is not driven by the enlargement of internal structures such as the ventral stamens. The sacs of the dithecous anther are now clearly visible, although the filament of the anthers are yet to develop and elongate. At this stage the dorsal stamen aborts development to form the small staminode barely visible at floral maturity. Surrounded by the enlarged developing anther thecae, the component parts of the ovary are clearly discernable. A short style supports a capitate stigma on the apex. Longitudinal sections of the ovary at this stage clearly show the presence of two locules containing numerous distinct ovules attached via axile placentae (Figure 3.4G, I).

Phase G The floral nectar spur primordium continues to develop from the base of the ventral petal of the corolla tube. However, at this stage the bud is still generally spherical in shape despite the initiation of the nectar spur and some growth of the corolla tube (Figure 3.4H). Continued bud development is characterised by proximo-distal elongation of the floral bud.

A rapid increase in the size of the corolla tube, petal lobes and nectar spur dominates later stages of floral ontogeny, generating a bud that more closely resembles the proximo-distally elongate flower at maturity (Figure 3.4H, J). The corolla tube now starts to become the predominant floral structure as the petals exceed the growth of the surrounding sepals. Longitudinal extension of the corolla tube is apparent as cells increase in size predominantly along their longitudinal axis (20 x 15μm) whilst cells in the region of the elongating nectar spur remain small and isodiametric, approximately 10 x 10μm in size (Figure 3.5B), indicating growth by cell division.

As the corolla tube elongates there is a deepening of the ventral furrow as it grows inwards to form an indentation or cleft, such that the ventral and lateral corolla lobes start to become folded back on themselves and distinct parts of the corolla tube, such as the palate, become apparent (Figure 3.4H, J). The upper and lower lips of the corolla tube are now clear, the dorsal part of the corolla tube and dorsal petal lobes form the upper lip of the flower, which extends over the lower lip formed from the ventral and lateral petals/petal lobes (Figure 3.4J). Cells in the central region of the corolla tube,

and at its base, continue to increase in size longitudinally, attaining a size of 35 x 25μm and 30 x 15μm respectively (Figure 3.5C). At the tip of the developing nectar spur there is also an apparent longitudinal increase in cell size to approximately 20 x 15μm. As the corolla tube continues to elongate the bud starts to resemble the mature flower (Figure 3.4J). Upon the upperside of the lower lip, which is exposed to the pollinator at anthesis, a dense growth of trichomes can now be seen as a deep orange pigmentation starts to develop. Cell sizes now increase dramatically, particularly along the longitudinal axis, reaching sizes of up to 90 x 30μm in the centre of the corolla tube, 75 x 20μm at the base of the corolla tube/top of the spur, and 55 x 20μm at the tip of the developing spur (Figure 3.5D).

Coincident with the elongation of the corolla tube the anther filaments and style also increase in length so that the stigma and anther thecae are positioned at the top of the opening to the corolla tube at maturity. As the stamen filaments begin to elongate, trichomes start to develop at their base. During elongation of the anther filaments they twist so that the anther thecae are presented toward the ventral part of the flower. The lateral and ventral stamens develop differentially, such that the ventral stamens surpass the height of the lateral stamens and style to generate the reproductive complex that is masked by the enlarged lower lip of the mature personate flower of *L. vulgaris* (see Figure 3.3).

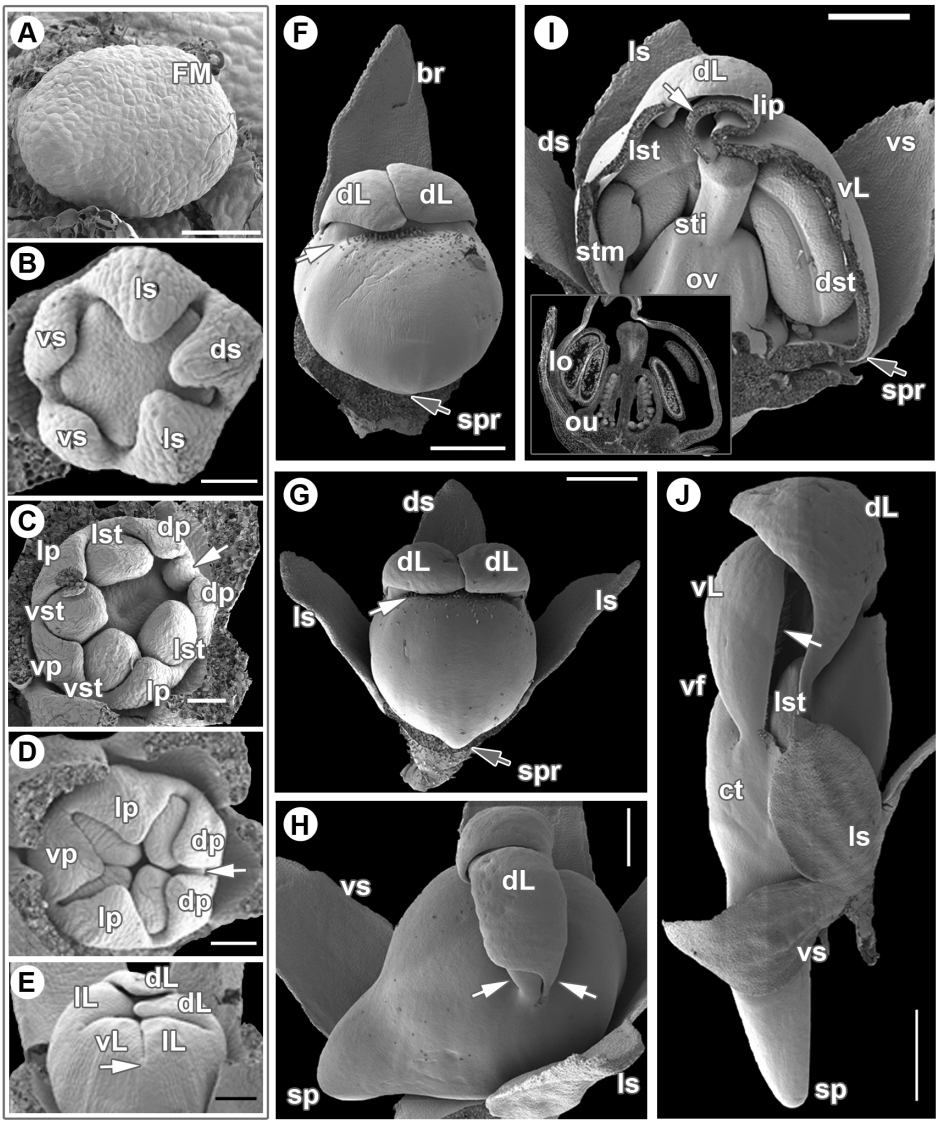


Figure 3.4: *L. vulgaris*, floral ontogeny (SEM, XRT)
Floral ontogeny can be divided into six phases, phase A pertains to initiation of the inflorescence apex (cf. *A. majus*, Vincent & Coen (2004)) A, B. Phase B - initiation of the floral meristem, sepal and petal whorls. C-E. Phase C- initiation of the stamen (C, D, arrow indicates staminode) and the gynoecium. Phase D - accelerated petal growth encloses the reproductive organs, the corolla tube and petal lobes are clearly marked by a tube-lobe boundary (arrow in E). F, G, I. Phase E - formation of a ventral furrow (white arrows) and nectar spur primordium (grey arrows). Longitudinal division of the bud (I, inset LS XRT image) shows outgrowth of the spur is not driven by enlargement of internal bud structures. H. Phase F - sepals stop growing, the ventral furrow deepens (white arrows, H), corolla tube and spur length increase longitudinally. J. Phase G - maturation and enlargement of floral organs. Scale bars = 50μm in A-E; 500μm in F-H ; 1mm in H. *Figure abbreviations: ct - corolla tube, dL, lL, vL dorsal, lateral and ventral petal lobes, ds, ls, vs - sepals, dp, lp, vp - petals, lo - anther locule, sti - stigma, vst, lst - ventral and lateral stamens, ov - ovary, ou - ovule, sp - spur, spr - spur primordium, stm - staminode, vf - ventral furrow.*

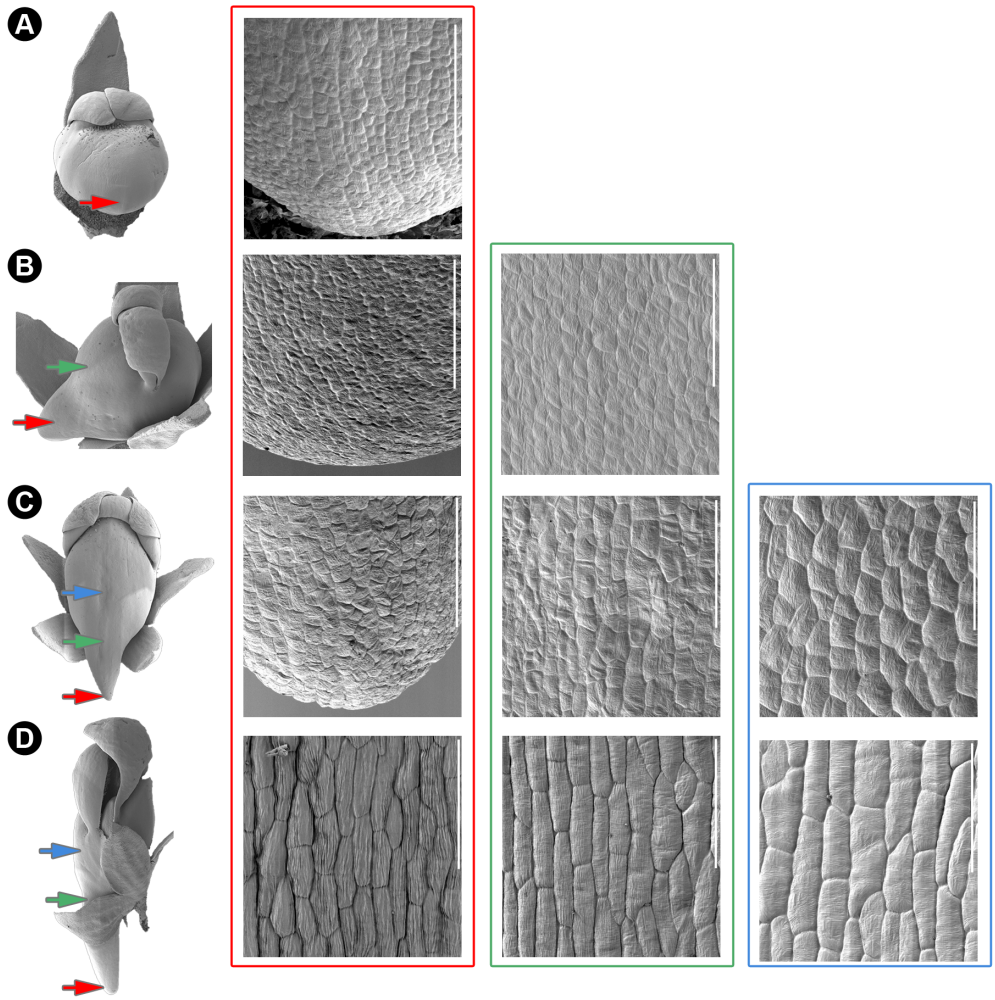


Figure 3.5: Cellular dimensions during *L. vulgaris* ontogeny
A. During spur initiation (phase E) cells are small and isodiametric, ca. 10 x 10μm. B. Cells during phase F remain relatively small at the tip of the floral nectar spur (red, ca. 10 x 10μm) whilst those in the corolla tube start to lengthen significantly (green, ca. 20 x 15μm). C, D. Subsequent floral development (phases F/G) is characterised by significant longitudinal increases in cell size. (C.) spur tip ca. 20 x 15μm, spur base ca. 30 x 15μm, corolla tube ca. 35 x 25μm, (D.) spur tip ca. 55 x 20μm, spur base ca. 75 x 20μm, corolla tube ca. 90 x 30μm. Coloured arrows and borders indicate the position of the corolla tube or nectar spur investigated at high SEM magnification, red indicates spur tip, green indicates spur base and blue indicates the corolla tube. Scale bars = 100μm.

3.2.2 Comparative floral morphology and ontogeny in *D. fuchsii* and *D. viridis*

3.2.2.1 Morphology

In *Dactylorhiza* (Figures 3.2A, B) the inflorescence is usually densely flowered, as in *D. fuchsii*, with leafy floral bracts often exceeding the flowers in length (Figures 3.2A, 3.6A-C). The inflorescence is compact to lax, conical to sub-cylindrical, varies considerably in height and bears numerous medium-sized white to pink flowers (Figure 3.2A). Lateral sepals are ovate to lanceolate, asymmetrical, often spotted on the inner surface, spreading to suberect. The dorsal sepal forms a hood with the lateral petals that encapsulates the gynostemium (Figure 3.6A).

In *D. viridis* (Figure 3.2B), the inflorescence is cylindrical and relatively lax, typically bearing fewer flowers than in *D. fuchsii*. The flowers are yellowish-green or greenish-brown, each subtended by a bract equalling or exceeding the flower in length (Figures 3.2B, 3.7A). Sepals are oval to triangular in shape (Figures 3.2B, 3.7D, E); the lateral sepals are asymmetrical, green, often with a purple-brown margin, and converge to form an obtuse hood that encapsulates the gynostemium. Lateral petals are linear and enclosed within the hood (Figures 3.2B, 3.7A, 3.8A, B).

The labellum of *D. fuchsii* is large (Figures 3.2A, 3.6A-D), 6-10 mm long and 8-16 mm wide; it is deeply trilobed, variously patterned with purple loops, spots and dashes (Figure 3.2A). Lateral lobes of the labellum are wrapped around the sides of the gynostemium in bud but are broadly spread at anthesis (Figures 3.6A-D); they are rounded, occasionally with a moderately crenulate margin. The triangular median lobe is almost as wide as, and typically longer than, the lateral lobes. Detailed examination of the interior surface of the *D. fuchsii* labellum revealed that the cells are distinctly conical, apart from the very base of the labellum surrounding the entrance to the spur (Figure 3.8H). The spur of *D. fuchsii* is a stout, narrowly conical structure approximately as long as the ovary, 6-10 mm in length (Figures 3.6A-D). Electron and light micrographs of the interior epidermis reveal numerous well-developed papillae (Figures 3.6F, 3.8G); these are distributed along the entire length of the spur, though they are concentrated toward the apex. In *D. fuchsii* the papillae have a bulbous tip, suggesting that they are secretory (Figures 3.6D, F).

The elongate labellum of *D. viridis* (typically 10 mm long and 3 mm wide) is flat, pendent or sometimes bent backwards, with a median longitudinal groove (Figures 3.8A, C, D, E). Like other perianth parts, the labellum is mostly green often with a brown/purple margin (Figure 3.2B). The elongate, parallel-sided labellum terminates in three lobes, the lateral lobes exceeding the tooth-like median lobe (Figure 3.7C). Prior to anthesis the

lateral lobes form a tight-fitting cap over the short, erect gynostemium (Figures 3.8A, B, F, 3.12G, H). Numerous domed cells, visible even by eye, are evenly distributed over the labellum surface. These cells are less elaborate in structure than those that characterise the labellum of *D. fuchsii* (Figures 3.6E, 3.7C, 3.8A, B, 3.12H) and are unlikely to be osmophores. The labellum possesses a globose (almost spherical), basally bi-lobed spur (Figures 3.7A, B, E, 3.8C, D). Cells of the interior epidermis of the spur form a smooth surface with no evident papillae, even in mature spurs dissected from flowers at anthesis (Figures 3.7F, C, 3.12I).

Dactylorhiza fuchsii is characterised by a short, erect gynostemium (Figures 3.6G, 3.8E). The anther is erect, tapering towards the base, consisting of a pair of large anther locules united by a broad and pointed anther connective; towards its base the viscidia are formed in the pouched bursicula (Figure 3.6A). The gynostemium has two small, but rather prominent, lateral staminodes (auricles), located on either side of the fertile stamen above two relatively undifferentiated basal bodies, to which they have become fused (Figures 3.8F). The median carpel forms the three-lobed rostellum, the median lobe of which is pleated and usually situated between the large parallel anther loculi (Figures 3.8E, F). The receptive lateral carpels that form the broad, bi-lobed stigmatic surface are clearly visible within the neck of the broad spur entrance (Figure 3.6G).

The gynostemium of *D. viridis* is similar to that of *D. fuchsii* in terms of gross structural organisation, though superficially it looks substantially different (Figures 3.7A, B, F). The anther is broader and square in shape; the anther loculi are relatively small and shallow (Figures 3.8A, B). Below the anther the median carpel lobe consists of a trilobed rostellum, the median part of which is located between the anther loculi in a fairly open pleat. Clearly visible within the entrance of the spur, and beneath the median part of the rostellum, are the lateral carpel lobes that form the receptive stigmatic surface (Figures 3.7A, F). The stigmatic surface is a broad bi-lobed structure, similar in shape to that of *D. fuchsii* but smaller.

In both *D. fuchsii* and *D. viridis* the ovary is cylindrical-fusiform, sessile, resupinate and glabrous, the single locule containing numerous small ovules with parietal placentation (Figures 3.6A, 3.7B, E, 3.8D, I). The three parietal placentas twist around the interior surface of the ovary, providing a further demonstration of the extent of 180° torsion in the ovary and resupination of the flower (Figures 3.7B, D, E, 3.8C, D, I).

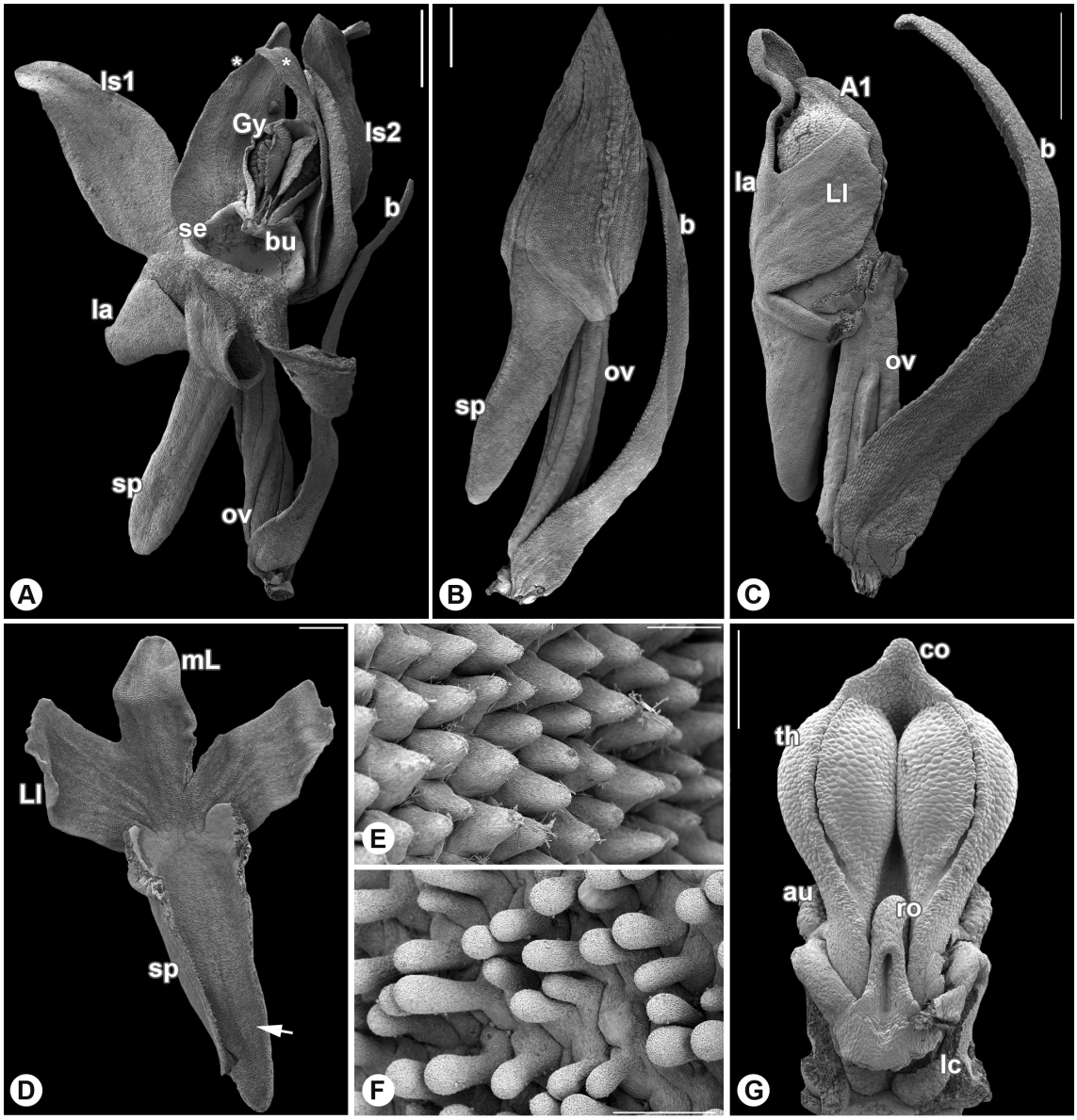


Figure 3.6: *D. fuchsii*, floral morphology (SEM)
A. Mature flower at anthesis, showing deeply trilobed labellum and stout, conical spur. B. Mature bud prior to anthesis and resupination showing stout, conical spur. C. Mature bud prior to anthesis, dissected to show link between labellum and gynostemium; the lateral lobes of the labellum are wrapped around the thecae of the gynostemium. D. Dissected spur/labellum shortly prior to anthesis, showing conical cells (arrowed) on labellum surface. E. Conical cells on labellum surface. F. Papillae on the internal surface of the spur. G. Mature gynostemium, showing elongate, tapering anther with lanceolate connective, sculptured auricles, rostellum and the receptive lateral carpel lobes in the neck of the spur. Scale bars = 1 mm in A-D; 50 μm in E, F; 500 μm in G. *Figure abbreviations: * - lateral petals, A1 - median adaxial fertile stamen, au - auricle, b - bract, bu - bursicula, co - anther connective, Gy - gynostemium, la - labellum (lip), lc - lateral carpel lobe, Ll - lateral labellum lobe, ls1 - lateral sepal 1, ls2 - lateral sepal 2, mL - median labellum lobe, ov - ovary, ro - rostellum, se - spur entrance, sp - spur.*

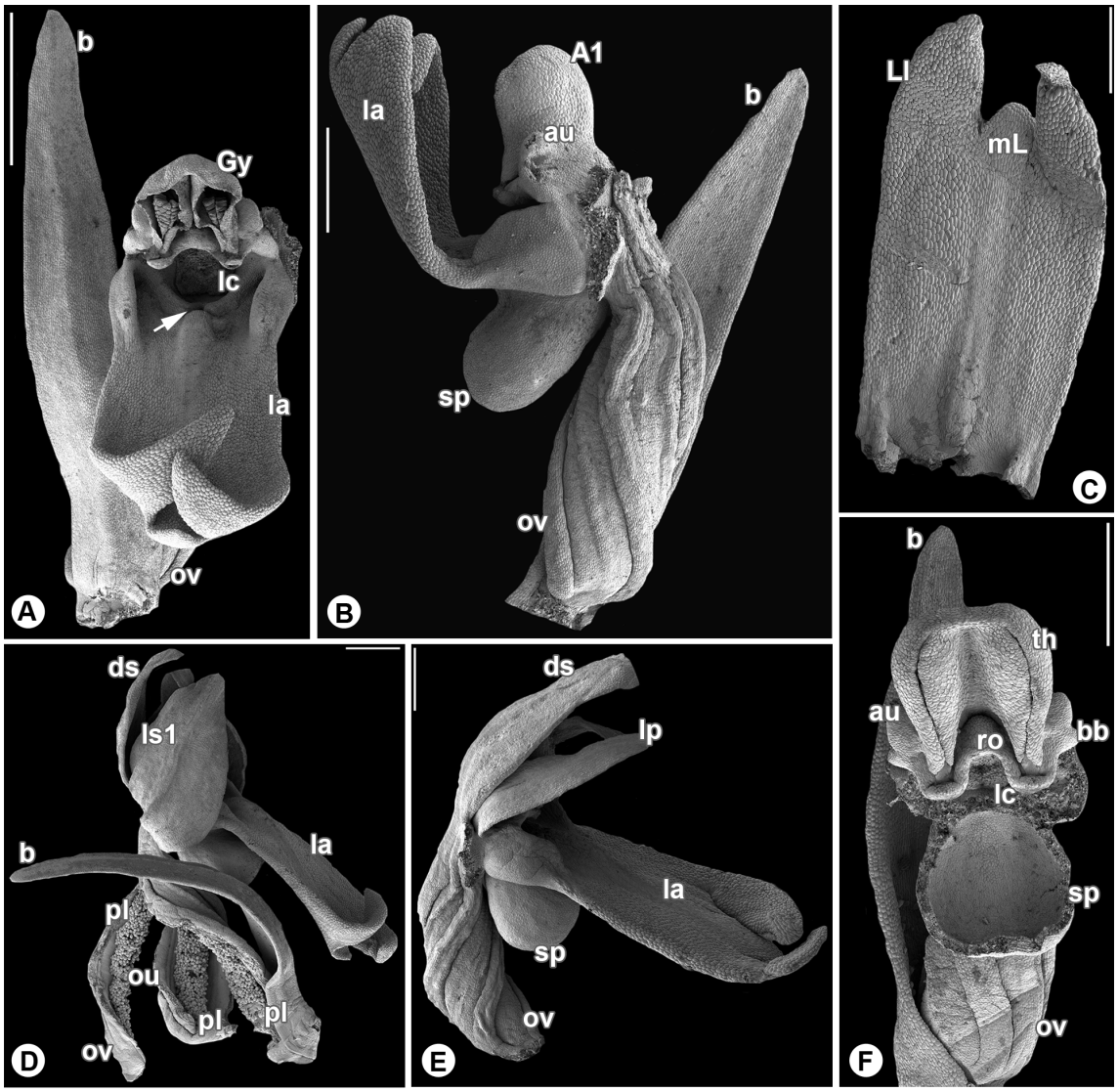


Figure 3.7: *D. viridis*, floral morphology (SEM)
A. Mature flower, perianth parts dissected, showing details of elongate labellum, gynostemium, receptive stigmatic surface and spur entrance (arrow); the ovary is resupinate. B. Lateral view of mature column just prior to anthesis; the spur is reflexed and globose. C. Dissected mature labellum; note lateral labellum lobes are much longer than median one, which is reduced and rounded to triangular in shape. D. Details of 180° torsion of the ovary, which had been split longitudinally to reveal three parietal placentas bearing numerous small ovules; the dorsal and lateral sepals are erect, forming a hood that encloses the gynostemium. E. Lateral view of mature flower, lateral sepals removed, showing linear lateral sepals enclosed in hood formed by erect petals. F. Details of mature gynostemium showing prominent auricles and basal bulges; the anther is short and square, as is the pleat of the rostellum; dissection of the spur in this mid-mature bud has revealed the absence of papillae. Scale bars = 1 mm. *Figure abbreviations: * - lateral petals, A1 - median adaxial fertile stamen, au - auricle, b - bract, bb - basal bulges, ds - dorsal sepal, Gy - gynostemium, la - labellum (lip), lc - lateral carpel lobe, lc1 - lateral carpel lobe 1, Ll - lateral labellum lobe, ls1 - lateral sepal 1, mL - median labellum lobe, ou - ovule, ov - ovary, pl - placenta, ro - rostellum, sp - spur, th - anther thecae.*

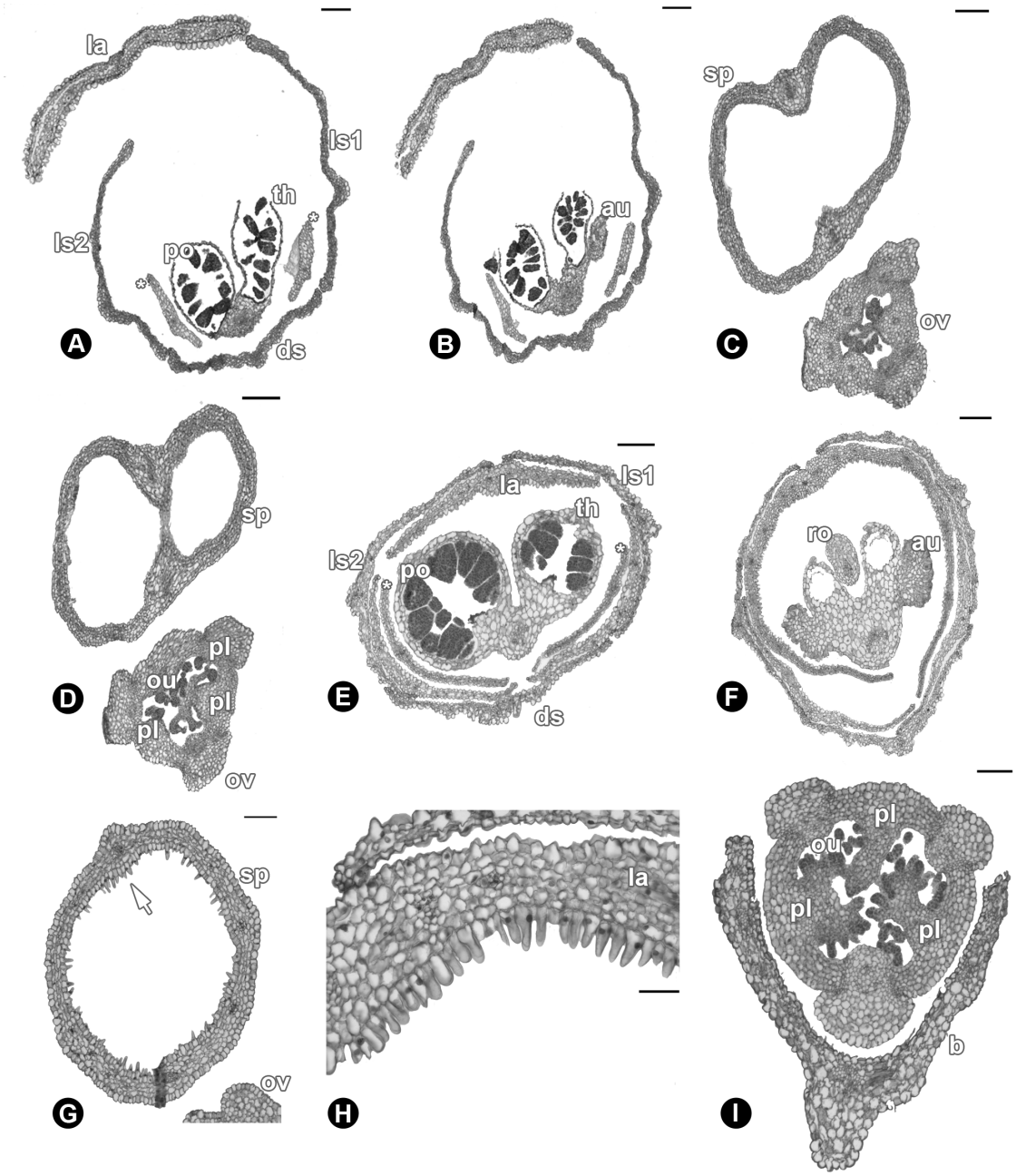


Figure 3.8: *D. viridis* (A-D), *D. fuchsii* (E-I), transverse sections of flowers (LM)
A, B. TS mature gynostemium. C. TS spur and ovary; papillae absent from interior spur surface. D. TS ovary showing parietal placentation and numerous small ovules. E, F. TS mature gynostemium. G. TS spur, showing numerous bulbous papillae lining interior of spur (arrow). H. TS mature labellum showing conical cells. I. TS ovary showing parietal placentation and numerous small ovules. Scale bars = 10 μm. *Figure abbreviations: * - lateral petals, au - auricle, ds - dorsal sepal, la - labellum (lip), ls1 - lateral sepal 1, ls2 - lateral sepal 2, ou - ovule, ov - ovary, pl - placenta, po - pollinium, ro - rostellum, sp - spur, th - anther thecae.*

3.2.2.2 Ontogeny

In both *D. fuchsii* and *D. viridis*, the first perianth organs to be initiated are the outer perianth parts, which are notably larger than the structures observable in the inner perianth whorl (Figures 3.9A, 3.12A). The primordium of the labellum is clearly visible between the lateral sepals and the slightly larger lateral petals. Based on the relative size of the organs it is likely that the lateral petals are initiated at the same time as the lateral sepals and labellum but that development proceeds at different rates. The gynostemium develops subsequent to the perianth segments (Figures 3.9B, C, 3.12B, C) and appears to follow a developmental program typical of other orchidoid orchids (cf. Box *et al.* (2008); Kurzweil (1998)).

In flowers of *D. fuchsii* the spur is initiated very late in floral ontogeny from the base of the labellum at approximately the same time that auricles are initiated on either side of the developing anther. At this point the lateral lobes of the labellum have yet to form (Figures 3.9D-F) but much of the remainder of the flower is in an advanced stage of development. Interestingly, in *D. fuchsii*, auricles are initiated while the gynostemium is still relatively immature, as revealed by dissection of young buds with already initiated spurs (Figures 3.9B, C, E, F)(Figs 12B, C). Although the auricles are relatively well differentiated in these buds, the median carpel apex has yet to extend into the deeply pleated rostellum; nonetheless, it has already extended between the anther thecae to form a deeply trilobed structure typical of orchidoid orchids. Lateral carpel apices are also relatively unelaborated at this point (Figures 3.9B, C), but resembling the broadly spread, bi-lobed receptive structure that sits within the spur entrance at maturity (Figures 3.6A, 3.9C,

In *D. fuchsii*, later development of the labellum proceeds by intercalary growth of the lateral lobes, coincident with extension of the spur and ovary (Figures 3.9D-F, 3.10A-C). When the spur is initiated the cells are isodiametric and approximately 12 x 10μm in size (Figure 3.11A). The conical cells that characterise the mature labellum surface are initially absent, forming later by elaboration of the domed cells, which extend apically throughout the duration of organ maturation (Figures 3.10H, I). The ovary initially elongates at a greater rate than the spur, but later the spur catches up, prior to resupination of the ovary (Figure 3.10E).

Spur elongation is associated with a significant increase in the size of the cells along the longitudinal axis throughout much of the length of the spur, increasing from approximately 12 x 10μm at spur initiation to cells in excess of 170μm long and 35μm wide at anthesis. Depite considerable increases in cell length, cells at the tip of the spur remain relatively small by comparison, attaining a length of only approximately 60μm but with a similar

width (Figure 3.11). As the spur elongates details of the accessory organs characteristic of the mature gynostemium become elaborated, particularly the developing rostellum, auricles and the anther, which itself increases considerably in girth, generating the steeply tapering anther toward its base (Figure 3.10D). Dissection of the developing spur (Figures 3.10D-G) reveals that papillae are absent from *D. fuchsii* prior to spur elongation but are amongst the last structures to develop prior to resupination of the ovary (Figures 3.10E, F).

In *D. viridis*, early floral development is similar to that of *D. fuchsii* up until the very latest stages of floral ontogeny predominantly concerned with later development of the nectar spur and the labellum. Spur development ceases relatively early in *D. viridis* in relation to that of *D. fuchsii*, shortly after intitiation of the spur, rapidly attaining a size representative of that at maturity whilst other organs, such as the gynostemium and semi-mature nature of the lateral stigma lobes, which typically characterise the latest stages of floral development in orchidoid orchids (Figures 3.12B-G), are yet to complete their development. This contrasts markedly with spur ontogeny in longer-spurred *D. fuchsii*. Comparison of different stages of spur development also suggests that the interior of the mature spur retains immature features in that no papillae develop (Figure 3.12I). Floral development in *D. viridis* includes resupination of the ovary, despite the apparent lack of maturity in the spur (Figures 3.7B, D, E, 3.12G, I). The lateral labellum lobes and spur are initiated at approximately the same time, at a point when the gynostemium is approaching maturity (Figure 3.12B-F). The different size of the lateral and median labellum lobes, which are initiated coincident with the development of the spur, is caused by extended growth of the lateral labellum lobes (Figures 3.12E-G). The domed cells that characterise the labellum of *D. viridis* (Figure 3.12H) resemble those of the early developing *D. fuchsii* labellum, suggesting that the labellum of *D. viridis* may also have experienced premature arrest of cellular differentiation.

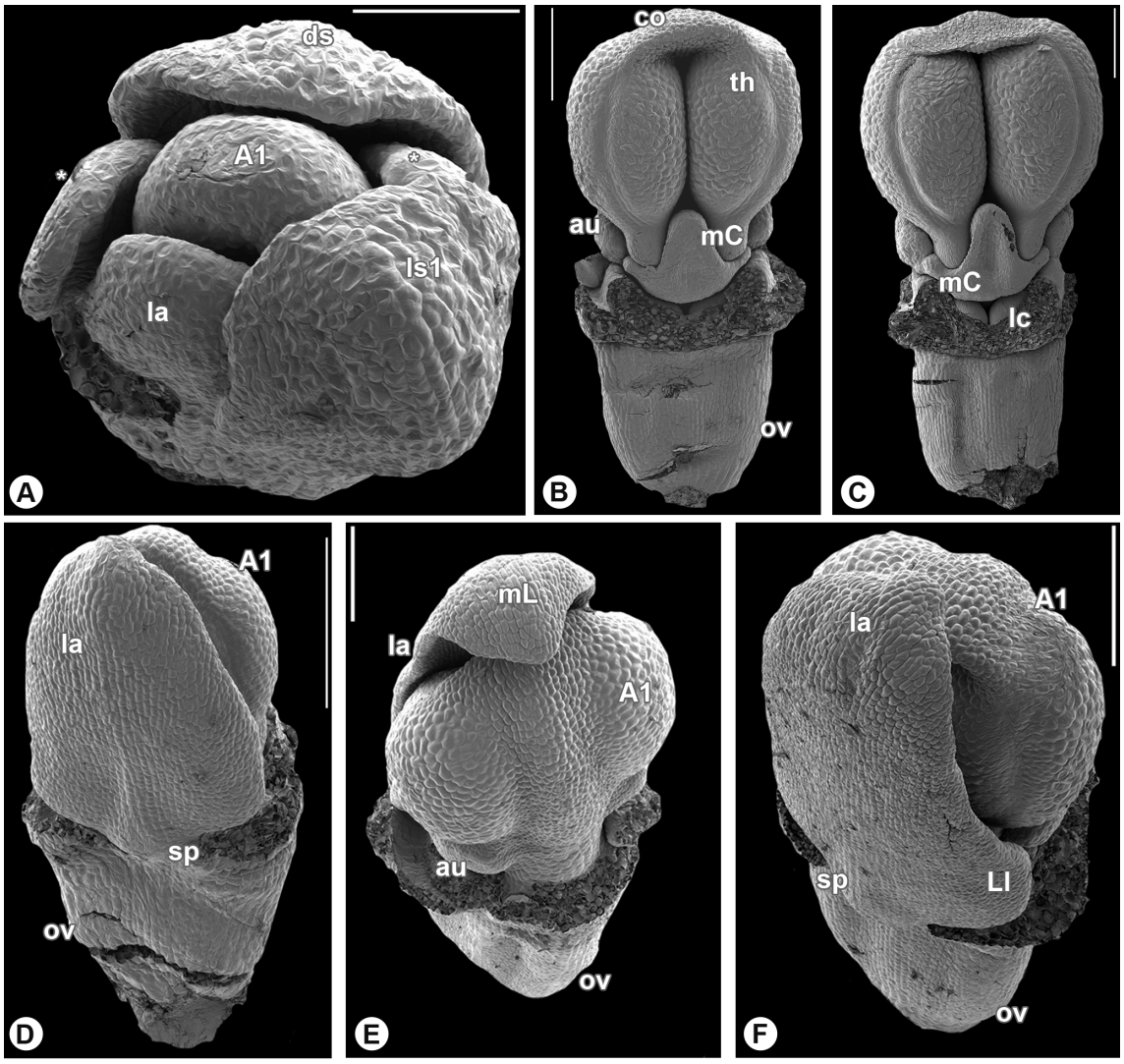


Figure 3.9: *D.fuchsii*, early floral ontogeny (SEM)
A. Young bud. B, C. Details of late gynostemium development, the anther, auricles and lateral carpel lobes are already formed; meanwhile, the median carpel lobe grows between the anther thecae and differentiates into the rostellum. D. Spur initiation at base of labellum; lateral labellum lobes are yet to be initiated. E. Auricles are initiated at approximately the same time as the spur. F. Lateral lobes of the labellum are initiated and grow prior to further extension of the initiated spur and ovary. Scale bars = 100 μmin A; 200 μmin BF. *Figure abbreviations: * - lateral petals, A1 - median adaxial fertile stamen, au - auricle, co - anther connective, ds - dorsal sepal, la - labellum (lip), lc - lateral carpel lobe, Ll - lateral labellum lobe, ls1 - lateral sepal 1, mC - median carpel lobe, mL - median labellum lobe, ov - ovary, sp - spur, th - anther thecae.*

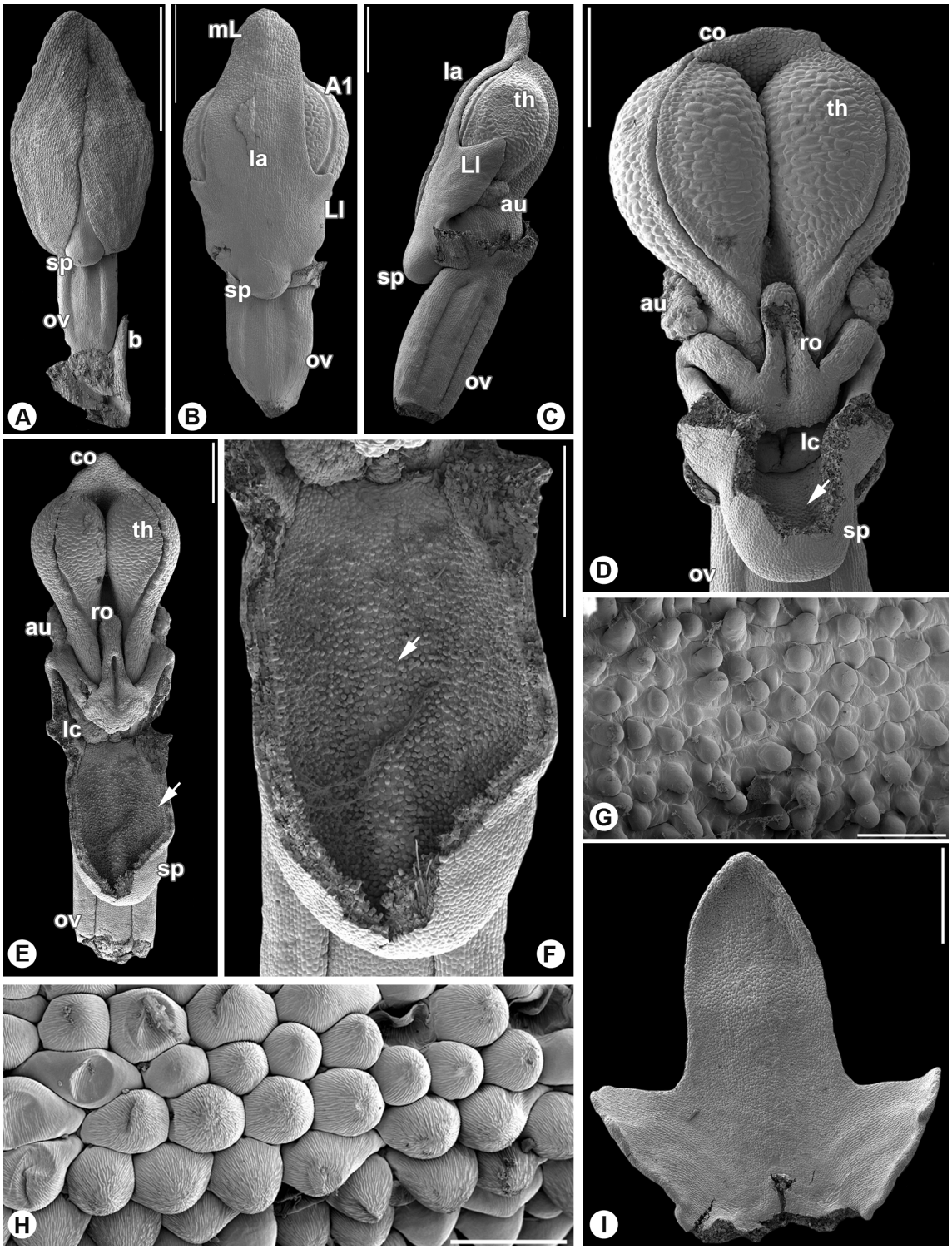


Figure 3.10: ***D.fuchsii*****, late floral ontogeny (SEM) continued over the page...**

Figure 3.10: *D.fuchsii*, late floral ontogeny (SEM) cont.
A. Details of spur initiation in a non-dissected bud, the spur is initiated between the lateral sepals at the base of the labellum. B. The labellum continues to develop prior to elongation of the spur, lateral lobes of the labellum are visible and the median lobe has extended. C. Auricles are visible at the side of the fertile anther. D-G. Spur extension occurs relatively late after development of the gynostemium (D); numerous developing papillae visible on the inner epidermis (G), as the spur elongates, papillae become more pronounced over the entire inner epidermis of the spur (E, F), arrows indicate spur papillae. H-I. The conical cells of the labellum surface develop late by apical extension of the dome-shaped cells that cover the labellum surface at an early stage of its development. Scale bars = 1 mm in A; 50 μmin G, H, 500 μmin B-F, I. *Figure abbreviations: A1 - median adaxial fertile stamen, au - auricle, co - anther connective, la - labellum (lip), lc - lateral carpel lobe, Ll - lateral labellum lobe, mL - median labellum lobe, ov - ovary, ro - rostellum, sp - spur, th - anther thecae.*

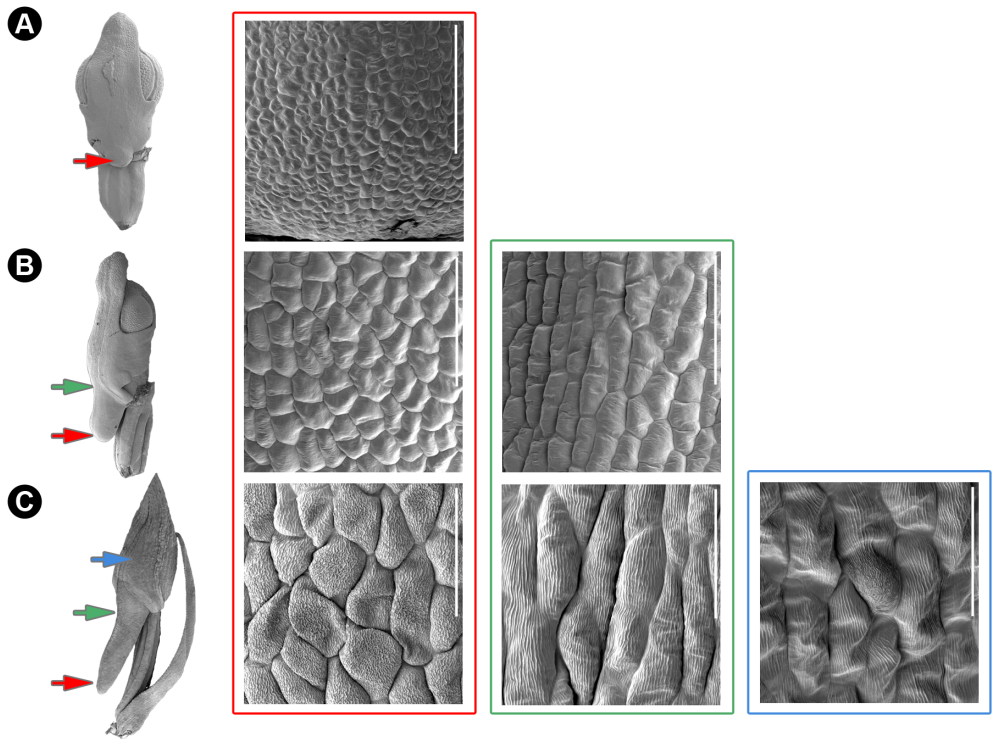


Figure 3.11: Cellular dimensions during *D. fuchsii* ontogeny
A. During spur initiation cells are small and isodiametric, ca. 12 x 10μm. B. Cells during nectar spur ontogeny remain relatively small at the tip of the floral nectar spur (red, ca. 30 x 20μm) whilst those in the corolla tube start to lengthen significantly (green, ca. 35 x 25μm). C. Subsequent floral development is characterised by significant longitudinal increases in cell size, cells at the spur tip remain smallest ca. 55 x 35μm, whilst those of the spur base (ca. 170 x 35μm) and corolla tube (ca. 100 x 35μm) increase in size dramatically. Coloured arrows and borders indicate the position of the corolla tube or nectar spur investigated at high SEM magnification, red indicates spur tip, green indicates spur base and blue indicates the corolla tube. Scale bars = 100μm.

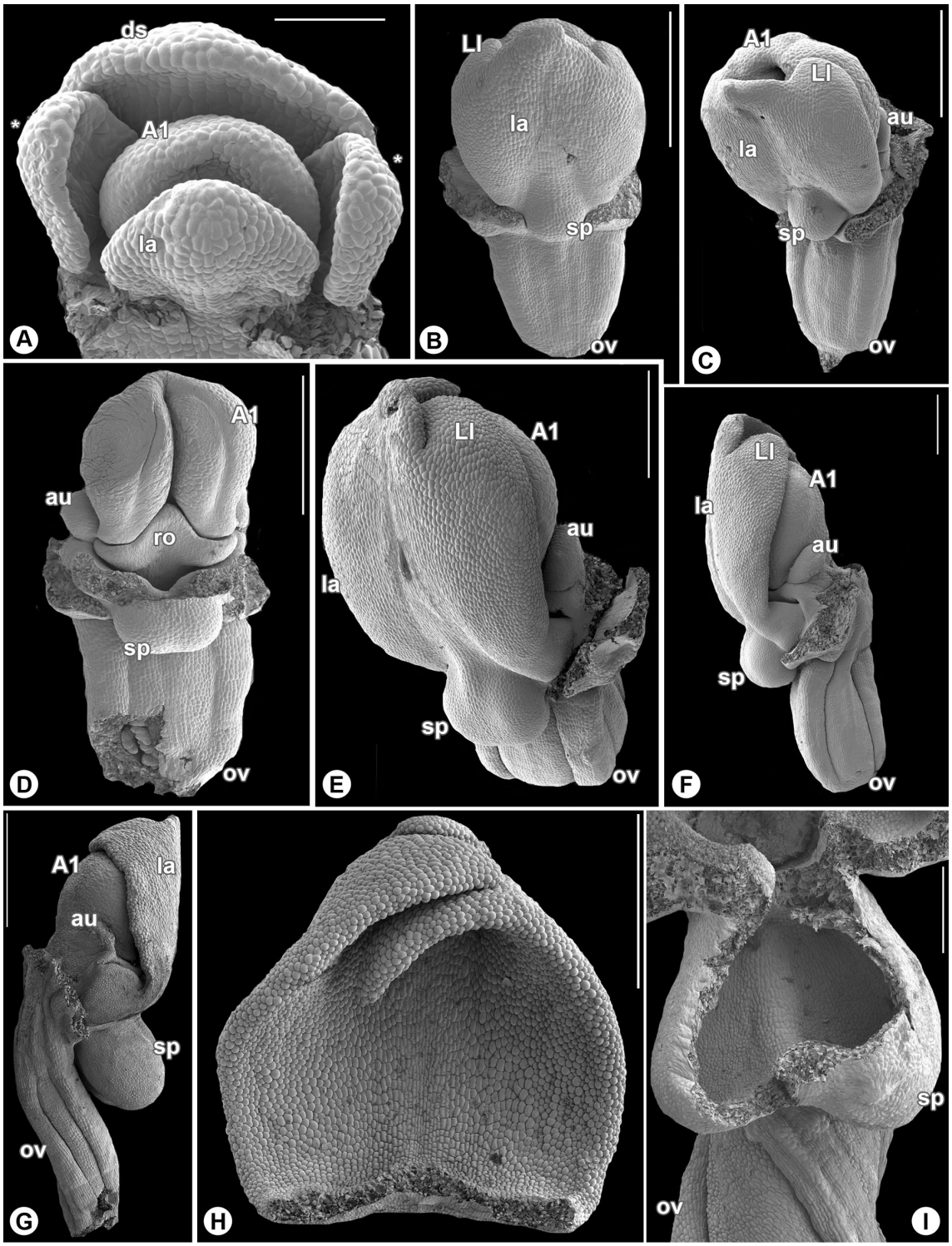


Figure 3.12: ***D.viridis*****, floral ontogeny (SEM) continued over the page...**

Figure 3.12: *D.viridis*, floral ontogeny (SEM) cont.
A. Early floral ontogeny. B. Spur initiation at base of labellum, coincident with lateral labellum lobe development. C. Elaboration of lateral labellum lobes, spur and ovary. D. Later elaboration and enlargement of spur into a broad saccate structure. E, F. The spur rapidly attains its mature size and shape and the lateral labellum lobes elongate. G. Semi-mature bud, the labellum forms a tight-fitting cap that sits on top of the anther, the spur has completed its development prior to 180° torsion of the ovary; the auricle and basal bulges are almost fully differentiated. H. Young labellum showing dome-shaped epidermal cells; the lateral lobes form a cap that fits neatly over the anther. I. Mature spur dissected to show the absence of papillae. Scale bars = 100 μmin A; 500 um in BF, I; 1 μmin G, H. *Figure abbreviations: * - lateral petals, A1 - median adaxial fertile stamen, au - auricle, la - labellum (lip), Ll - lateral labellum lobe, ov - ovary, ro - rostellum, sp - spur, th - anther thecae.*

3.3 Discussion

Patterns of floral ontogeny in *L. vulgaris* and the orchids *D. fuchsii* and *D. viridis* are discussed in relation to the evolution and development of spurs in each of these taxa and efforts to characterise the genetic determinants of floral nectar spurs.

3.3.1 Floral ontogeny in *L. vulgaris*

The pattern of floral ontogeny has been determined for *Linaria vulgaris* from initiation to maturity. Floral ontogeny in *L. vulgaris* is fundamentally similar to that of *Antirrhinum majus*, which has been described in detail by a number of authors, e.g. Awasthi *et al.* (1984); Singh & Jain (1979); Vincent & Coen (2004). In snapdragon, floral ontogeny is divided into 7 phases (Vincent & Coen, 2004). These phases are also adequate descriptors of the events that occur in floral ontogeny of *L. vulgaris* and have been adopted in this study. During phase A, the floral meristem is initiated in the axil of a bract. It is during this stage that expression of floral meristem identity genes such as *FLORICAULA* and *SQUAMOSA* has been demonstrated in *A. majus* (Coen *et al.*, 1990; Huijser *et al.*, 1992). Although phase A was not directly observed in *L. vulgaris* during this study, it is unlikely to differ significantly from the observations made in snapdragon. Floral symmetry is apparent very early in floral ontogeny, during phase B the floral meristem develops five-fold symmetry as the five sepals are initiated around the periphery of the floral meristem. In snapdragon, expression of floral identity genes such as *DEFICIENS* and *PLENA* have been demonstrated (Bradley *et al.*, 1993; Schwarz-Sommer *et al.*, 1992). In snapdragon and *L. vulgaris*, Almeida *et al.* (1997); Luo *et al.* (1999) and Cubas *et al.* (1999) have shown that expression of symmetry determining genes such as *CYCLOIDEA*, *DIVARICATA*, *DICHOTOMA* and *RADIALIS* precedes this stage of floral ontogeny and that such expression is maintained throughout much of the subsequent stages of floral ontogeny. During phase C, the remaining floral organ whorls develop in the sequence petals, stamens and carpels. In phase D, the petal primordia are fused congenitally at their bases forming a proximal corolla tube and the distal petal lobes that cover the inner whorls.

Up until phase D, ontogeny of *A. majus* and *L. vulgaris* flowers is indistinguishable. The subtle differences between snapdragon and toadflax flowers develop during subsequent phases of floral ontogeny. During phase E, critical common features of the *Antirrhinum* and *Linaria* flowers start to develop. The palate and lips start to become apparent as an in-growing ventral furrow starts to develop at the tube-lobe boundary. In addition, specialised cell types such as trichomes begin to differentiate and, crucially, a nectar spur primordium can be seen emerging as an abaxial bulge at the base of the ventral

petal of the growing corolla tube in *L. vulgaris*. Nectar spur initiation occurs in an identical position and at an identical time to a small swelling that forms at the base of the ventral petal of the snapdragon corolla tube, the gibba, strongly indicating that these structures are homologous. Some authors have claimed that the nectar spur of *Linaria* is a composite structure formed from the ventral and lateral petals (Endress & Matthews, 2006), however morphological and ontogenetic analyses presented here suggest that this is not the case. Such an assertion is strongly supported by the radially ventralised five-spurred actinomorphic *L. vulgaris cycloidea* mutants described by Cubas *et al.* (1999). In the last phase of *L. vulgaris* floral ontogeny, phase G, expansion of tissues, elaboration of the corolla and elongation of the stigma, style, corolla tube and spur, are the predominant features of late floral ontogeny.

Early floral ontogeny has been discussed previously by Cubas *et al.* (1999) in relation to understanding floral zygomorphy. However, Vincent & Coen (2004) suggested that many key floral features that characterise different species, such as organ shape and size, arise during later stages of floral ontogeny. The strong similarity of early floral ontogeny in *A. majus* and *L. vulgaris* suggests that, at least in part, this is true. Although the gross morphology of snapdragon and toadflax flowers is highly similar, a number of key morphological differences, such as the more elongate flower of *Linaria*, erect dorsal petal lobes and the long, narrow floral nectar spur that distinguish flowers of each species, develop very late in ontogeny and may result from subtle changes in the timing of organ development in each species, a phenomenon termed heterochrony (section 3.4.2).

Rolland-Lagan *et al.* (2003) used a combination of clonal analysis with estimates of cell sizes, to show that growth during phases D-F is largely associated with cell division in snapdragon, whilst growth during phase G is largely related to cell expansion. This indicates that many of the features that distinguish *Linaria* flowers from those of snapdragon may be due to cell elongation rather than cell division. Crude observations of cell size in this study, with particular emphasis on the floral nectar spur, appear to tell a similar story. Cells undergo significant increases in longitudinal dimensions during late stages of floral ontogeny, particularly in the corolla tube and spur. Interestingly, the floral nectar spur in *Linaria* is initiated during phase D/E and is characterised by a large number of small, isodiametric cells at initiation. Small, isodiametric cells strongly suggest that during early initiation at least, growth and development of the floral nectar spur in *Linaria* is related to cell division, consistent with the observations of Rolland-Lagan *et al.* (2003) during this stage of ontogeny in snapdragon. Later growth of the floral nectar spur seems not to be related to cell division but to elongation as, although cells at the proximal tip of the nectar spur remain relatively small compared to more distal cells, there appears to be no meristematic region at the tip, or base, of the elongating nectar spur. One may

expect such a region if growth of the nectar spur is driven predominantly by cell division however, diffuse cell divisions are also a common phenomenon in organ expansion during floral development and it is likely that the excessive length of many floral nectar spurs is formed by a combination of cell division and cell expansion.

3.4 Nectar spur evolution and development in Antirrhineae

The abaxial base of the Antirrhineae corolla tube is known to exhibit a more-or-less continuous range of structures that differ predominantly in proximal-distal length, variably known as entire, saccate, gibbous, or spurred (Sutton, 1988). By far the most commonly represented of these structures in *A. majus* and its close relatives is a small swelling located at the abaxial base of the corolla tube, known as a *gibba* (Sutton, 1988). Golz *et al.* (2002) interpreted the gibba of *A. majus* as a *vestigial* nectar spur. Use of the term vestigial is, in itself, problematic (below), however such a comparison also implies homology where none has been demonstrated.

Previous analyses of floral ontogeny in *L. vulgaris* were focused only on early events (Cubas *et al.*, 1999), providing no framework with which to compare to the complete ontogeny of snapdragon flowers (Vincent & Coen, 2004). In this chapter, floral ontogeny has been extended considerably to permit ontogenetic comparisons to be made with that of *A. majus*. Comparative floral ontogeny of *L. vulgaris* and *A. majus* clearly indicate that the nectar spur of *Linaria* and the gibba of *A. majus* are homologous. Criteria for reliable interpretation of *homology* have been extensively debated in recent years (see Kleisner (2007)) however, based on the fundamental criteria of position and similarity (topography, histology and ontogeny, Patterson (1982)), the gibba of snapdragon and the spur of toadflax may be regarded as homologous structures. Furthermore, both structures function to store nectar secreted from the base of the common gynoecial nectary to attract bee pollinators (Elisens & Freeman, 1988; Sutton, 1988). Firmly establishing the homology of these structures is a critical first step in exploring the evolution of the floral nectar spur in Antirrhineae.

Sutton (1988) and Golz *et al.* (2002) suggested that *A. majus* is unusual among Antirrhineae in that it does not possess nectar spurs. However, subsequent phylogenetic re-assessment of Antirrhineae has resulted in considerable taxonomic expansion of the tribe (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004). As such, the presence of nectar spurs in close relatives of *A. majus* appears to be relatively uncommon. The commonality of nectar spurs among other members of Antirrhineae, when combined with phylogenetic data, indicates the direction of character

evolution allowing one to infer whether spurs have evolved once and been lost on many occasions, or whether nectar spurs have evolved multiple times independently in Antirrhineae. Previous interpretations of nectar spur evolution in Antirrhineae (Golz *et al.*, 2002), have referred to the gibba of *A. majus* as a vestigial nectar spur. The term *vestigial* implies polarity, which can only be inferred from a phylogenetic tree. However, phylogenetic analyses of the family Plantaginaceae (Albach *et al.*, 2005), tribe Antirrhineae (Ghebrehiwet *et al.*, 2000; Vargas *et al.*, 2004), and the genus *Antirrhinum* (Oyama & Baum, 2004) using different taxon sampling, molecular and morphological markers, yield conflicting phylogenetic relationships, preventing a conclusive interpretation of polarity in the evolution of the nectar spur.

Currently the phylogenetic analysis with the most extensive taxon sampling of Antirrhineae is that of Ghebrehiwet *et al.* (2000), which uses combined morphological and plastid *ndhF* characters. According to this classification the Antirrhineae is divided into four principal clades, the *Antirrhinum*, *Maurandya*, *Gambelia* and *Annarhinum* clades. The spur bearing genera *Linaria* and *Chaenorhinum* are recovered as basal to *Antirrhinum* and other non-spurred genera such as *Misopates*, *Schweinfurthia*, *Howelliella* and *Mohaveae* in the *Antirrhinum* clade. The *Antirrhinum* clade, however, forms part of a polytomy with the *Maurandya* and *Gambelia* clades, both of which contain predominantly non-spur bearing genera (*Maurandya*, *Maurandella*, *Rhodochiton*, *Asarina*, *Gambelia* and *Galvezia*), with the exception of spur-bearing *Cymbalaria*. Ghebrehiwet *et al.* (2000) recover the spur-bearing genera *Annarrhinum* and *Kickxia* at the base of the Antirrhineae, which could suggest that nectar spurs are ancestral. However, the polytomy between the *Antirrhinum*, *Maurandya* and *Gambellia* clades makes it impossible to confidently determine whether nectar spurs have been lost in the genera *Maurandya*, *Maurandella*, *Rhodochiton*, *Asarina*, *Gambelia* and *Galvezia*, or gained in *Cymbalaria*, *Chaenorhinum* and *Linaria*.

Less well taxonomically sampled phylogenies using more molecular markers (Albach *et al.*, 2005), recover *Linaria* as sister genus to *Antirrhinum*, both of which form a sister clade to the genera *Cymbalaria* and *Maurandya*. Basal to this complex are the sister genera *Galvezia* and *Kickxia*. In each case, spurred and non-spurred genera are recovered as sister pairs, making it impossible to determine the order in which nectar spurs may have been lost or gained. A species level phylogeny of *Antirrhinum* based only on nuclear *ITS* data, including representatives from the other major Antirrhineae genera (Oyama & Baum, 2004), has excellent phylogenetic resolution but is not helpful in determining patterns of character evolution. Based on these phylogenetic data character reconstruction is equivocal. As a result, neither the gibba or the nectar spur can be resolved as ancestral or derived characters.

What is clear is that the nectar spur (and the related gibba) is a highly evolutionarily labile structure that has been lost or gained multiple times independently in Antirrhineae. Whilst it is not possible to resolve the evolutionary status of the gibba and nectar spur, these structures are homologous because they pass the similarity and positional tests. It is feasible therefore that the nectar spur and gibba are evolutionarily related, perhaps representative of a labile evolutionary transition that may result from relatively simple changes in the amount of growth of the gibba/spur. Such a step may be related to simple changes in the timing of development, heterochrony, and/or the loss of terminal stages of development.

3.4.1 Floral ontogeny in the orchids *D. fuchsii* and *D. viridis*

As is the case for *Linaria* and *Antirrhinum*, differences in gross floral morphology between *D. fuchsii* and *D. viridis* appear to be related to later stages of floral ontogeny, particularly the nectar spur. Early floral ontogeny in both *D. fuchsii* and *D. viridis* is identical to that described for closely related genera such as *Gymnadenia* (Box *et al.*, 2008) and other orchidoid orchids (Kurzweil, 1998). The outer perianth parts are initiated first of all, followed by the inner perianth parts, of which the ventral labellum petal grows considerably in relation to the other perianth parts. Subsequent development of the complex reproductive gynostemium follows, much of which is completed by the time the characteristic labellum lobes and floral nectar spurs develop in *D. fuchsii* and *D. viridis*.

The nectar spur is initiated as an abaxial bulge at the base of the labellum relatively late in floral ontogeny in both *D. fuchsii* and *D. viridis*. As is the case for *L. vulgaris*, at spur initiation the cells are small and isodiametric, indicating that the initial outgrowth of the spur primordium may be dependent, to some degree, on cell division. However, during later stages of ontogeny, cells of the nectar spur and labellum are extended longitudinally by up to 1,700% in *D. fuchsii*, strongly indicating that subsequent growth of the longer floral nectar spurs appears to be highly dependent upon cell elongation predominantly along their longitudinal axis and perhaps in combination with diffuse cell divisions. In addition, the ontogenetic observations presented here indicate that nectar spurs continue to elongate after anthesis. Similar observations have been made in *Gymnadenia* Box *et al.* (2008) and *Platanthera* Bateman & Sexton (2008), which may have major implications for hypotheses of adaptive optimisation of spur length.

3.4.2 Key differences in floral morphology are the result of heterochrony

A number of key morphological differences between the flowers of *A. majus* and *L. vulgaris*, and between flowers of the orchids *D. fuchsii* and *D. viridis*, respectively, result from subtle changes in the timing of organ development, termed heterochrony (Alberch *et al.*, 1979; Gould, 1977). In particular, they may result from a specific type of heterochrony called paedomorphosis.

Paedomorphosis results in the retention of juvenile traits in a derived species characteristic of earlier stages in the ontogeny of the ancestor (Alberch *et al.*, 1979; Gould, 1977). Although the final morphological result is an overall reduction in the amount of total growth of an organ, or group of organs, paedomorphosis can occur by one of three mechanisms, neoteny, progenesis and post-displacement. Neoteny refers to a reduction in the rate of development, as such development proceeds more slowly in the derived species compared to that of the ancestor. By contrast, progenesis results from early cessation of development in the descendant compared to the ancestor. Post-displacement involves a delay in the onset of development in the descendant relative to it's ancestor (Bateman & DiMichele, 1994; Fink, 1982). Heterochronies such as these can only be observed by comparing two closely related lineages for which there is strong phylogenetic evidence indicating the direction of character evolution, i.e. ancestral versus derived characters. Whilst differences in the structure of the homologous gibba and spur in Antirrhineae are likely to have arisen as a consequence of heterochronic shifts, the absence of strong phylogenetic data for, or against, the evolutionary status of either structure (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004), restricts further discussion of this topic in relation to Antirrhineae. By contrast, there is strong phylogenetic evidence for the ancestral status of the longer-spurred *D. fuchsii* flower relative to that of *D. viridis* (Bateman *et al.*, 2003). In light of this fact, discussions of heterochrony that follow will focus on the orchids *D. fuchsii* and *D. viridis*.

3.4.2.1 Paedomorphic features in orchid flowers: the floral nectar spur of *D. viridis*

Several distinctly paedomorphic features in the flowers of *D. viridis* have been identified in relation to the ancestral condition exemplified by *D. fuchsii*. Reduction of the labellar spur is the most significant paedomorphic feature that is likely to have played a key role in the evolution of *D. viridis* from longer-spurred ancestors. The ontogenetic series suggests that development of the floral nectar spur in *D. viridis* ceases prematurely relative to that of *D. fuchsii*. Thus, the development of the nectar spur in *D. viridis* appears to be

progenetically curtailed, generating the short spur characteristic of *D. viridis*. A similar phenomenon has been reported for nectar spur reductions apparent in the sister genus *Gymnadenia* (Box *et al.*, 2008).

3.4.2.2 Consequences of heterochronic shifts: spurs and speciation

Despite substantial differences in floral morphology and ontogeny, there is little variation in *ITS* nuclear rDNA sequence data among longer and shorter-spurred representatives in the *Dactylorhiza* and *Gymnadenia* clades (Bateman *et al.*, 2003). This suggests that heterochronic shifts such as nectar spur reduction may have been responsible for the recent radiation of short-spurred species amongst Orchidinae in general and within the *Dactylorhiza* clade in particular, as suggested for members of the *Gymnadenia* clade (Bateman *et al.*, 2006; Box *et al.*, 2008).

The morphology of the nectar spur has been intimately tied to reproduction in several angiosperm groups. Simple differences in the length, shape, orientation and colouration of spurs are commonly associated with different pollinators and affect reproductive isolation (Hodges, 1997; Hodges & Arnold, 1995). Phylogenetic evidence shows that nectar spurs have evolved on multiple occasions in a diverse range of angiosperms (Rudall *et al.*, 2003), especially within Orchidaceae, and spur evolution has frequently been linked to the unusually high species richness of these groups of flowering plants (see Chapter 1). Changes in spur length mediated by alterations in the timing of nectar spur development may have played a significant role in the recent evolutionary history of *Dactylorhiza* and its close relatives. Bateman & DiMichele (2002) proposed that a single mutation in a critical developmental gene might have instantaneously triggered the origin of the short-spurred floral phenotype from the ancestrally longer-spurred phenotype in the *Dactylorhiza* and *Gymnadenia* clades. Dramatic alterations in floral morphology that might have resulted from mutation in a single key developmental gene exemplify a possible case of post-saltational radiation (i.e. species radiation following a dramatic evolutionary leap in phenotype) among members of the subtribe Orchidinae. Our data are consistent with this suggestion, since the observed heterochrony is most likely the result of changes in timing of expression of a single master regulatory gene.

However, the ecological consequences of radical reduction of the spur in the orchids analysed here are unclear. Reduction in spur length among members of the sub-tribe Orchidinae, to which the genera *Dactylorhiza* and *Gymnadenia* belong, has been implicated in speciation by causing switches in pollinators or being rendered irrelevant by a coincident one-way transition from allogamy (cross-fertilisation) to autogamy (self-fertilisation) (Bateman & DiMichele, 2002). Some pollination studies exist for the orchids *D. fuchsii* and *D. viridis*, but such studies are of necessity geographically localised, and so cannot

capture the often radical shifts in the identity of the dominant pollinators that occur across the full geographical range of the species. In general, detailed pollination studies are lacking. There is need for further rigorous observations of pollination biology, particularly for short-spurred *D. viridis*. However, one must be careful when interpreting these data as almost all such previous studies are too parochial to be of great help in understanding orchid evolution, wherein the importance of co-evolution with pollinators has been hugely exaggerated. In the absence of concrete evidence for or against shifts in pollinators or switches from allogamy to autogamy, we cannot conclude whether this phenomenon explains the origin of the short-spurred floral phenotype in *Dactylorhiza*.

A further possibility is that nectar spur reduction may be associated with loss of nectar spur function, i.e. nectar secretion/storage. Papillae lining the interior of the nectar spur have been suggested to play a role in secretion and re-sorption of nectar (see Bell *et al.* (2009)), the absence of papillae in the short spur of *D. viridis* suggests that nectar secretion, i.e. function, may have been lost in relation to nectar spur reduction. However, the presence of papillae is not a guaranteed indicator of nectar secretion, as *D. fuchsii*, with prominent papillae-like cells lining the interior epidermis, does not secrete nectar. By contrast, *D. viridis*, whilst lacking any obvious nectar secreting structures inside the spur, is reported to secrete small amounts of nectar that may originate from alternative sources. This may well represent a shift from a deceptive to rewarding pollination strategy.

In animals, paedomorphosis has been implicated in the evolution of a more generalist organism able to exploit both adult and juvenile niches (Gould, 1977, 1992). Likewise progenetic paedomorphosis in the spur may improve generalism by increasing the range of animals able to access nectar, and thus reduce pollinator specificity. However, although many of these explanations seem plausible, seeking adaptive significance in spur reduction may be wholly unnecessary. At a single high altitude site in Austria, both *D. fuchsii* and *D. viridis* can be found flourishing, though with some differences regarding prevalence with respect to altitude, perhaps in relation to declining long-tongued pollinator abundance with altitude. In general however, both short and medium-spurred representatives occur alongside one another with no apparent evidence to suggest any difference in competitiveness or reproductive success. In the absence of field observations or genetic studies to suggest any competitive advantage of flowers with short versus longer nectar spurs, it is impossible to conclude that there are any adaptive advantages associated with nectar spur reduction. Consequently there may be no adaptive value to spur reduction in Orchidinae, rather this phenomenon might exemplify a case of making do with whatever mixed hand nature deals. Such a standpoint is firmly opposed to the predominantly adaptive viewpoint of Hodges (1997); Hodges & Arnold (1995); Hodges *et al.* (2002).

3.4.3 Using morphology to guide developmental genetic analyses of nectar spur development

Nectar spur development in *L. vulgaris* and the orchids *D. fuchsii* and *D. viridis* is similar to that of other taxa studied, such as *Aquilegia* (Gottlieb, 1984; Tepfer, 1953; Tucker & Hodges, 2005) and other members of the orchid sub-tribe Orchidinae (Kurzweil, 1998), e.g. *Gymnadenia* (Box *et al.*, 2008). Examination of nectar spur ontogeny in each of these plant groups demonstrates that, despite the significant morphological differences between the flowers of *Aquilegia* and spur bearing Orchidinae, nectar spurs are among the last floral organs to develop and are derived from a late forming abaxial bulge at the base of one or more petals that are themselves in a relatively advanced stage of ontogeny. These observations indicate that, despite often significant morphological variation between the nectar spurs of different angiosperm taxa and multiple independent origins of spurs, they may develop by the same underlying means.

Golz *et al.* (2002) suggested that class 1 *KNOX* genes, normally involved in maintaining indeterminacy at the shoot apical meristem, may be involved in the evolution and development of floral nectar spurs in Antirrhineae based on observations of ectopic *KNOX* gene expression in neomorphic spur-bearing snapdragon mutants. By carefully assessing floral ontogeny in *L. vulgaris*, *D. fuchsii* and *D. viridis*, ontogenetic stages have been identified that can be targeted for subsequent analyses aimed at understanding the genetic factors that may be important in driving morphological variation in nectar spurs. If *KNOX* genes do play a role in the development of the floral nectar spur, it appears likely that they will be expressed during early initiation of the nectar spur primordium, when cell division appears to be prevalent. Sampling later stages of nectar spur development are unlikely to identify key factors in the initiation of the nectar spur but may reveal genetic factors that are involved in nectar spur elongation. Factors identified during initiation and/or elongation of the nectar spur may be integral to the final length of the floral nectar spur and, as such, could prove to be important factors in the development of morphological differences in nectar spur length that may be important for speciation.

3.5 Conclusions

Nectar spur development in *Linaria vulgaris* and the orchids *D. fuchsii* and *D. viridis* occurs late in petal ontogeny. Despite the considerable phylogenetic distance between *L. vulgaris*, orchids and other published cases, e.g. *Aquilegia*, nectar spur ontogeny occurs by very similar means, suggestive of a common developmental pathway. The identification of spur-bearing mutants of snapdragon by Golz *et al.* (2002) has implicated class 1 *KNOX* genes as candidates for a key role in nectar spur development. Morphological and onto-

genetic analyses show that further genetic analyses, inspired by Golz *et al.* (2002), should focus on mid to late stages of floral ontogeny (but late petal ontogeny) during which cell division appears to characterise the initiation of the petal spur.

The snapdragon gibba and *Linaria* spur may be interpreted as homologous and highly evolutionarily labile structures that may be rapidly and easily interconverted, perhaps driven by heterochronic shifts in development related to changing patterns of *KNOX* gene expression? In the absence of firm phylogenetic data it is impossible to interpret the direction of character evolution in Antirrhineae and the potential heterochronies that may be operating on the gibba and spur. By contrast, differences in nectar spur length between the orchids *D. fuchsii* and *D. viridis* are caused by a specific form of heterochrony called progenesis, i.e. premature cessation of development of the floral nectar spur in short-spurred *D. viridis*. However, in the absence of detailed pollination studies the precise ecological consequences of spur reduction in *D. viridis* are difficult to ascertain. Investigating the genetic factors associated with nectar spur development in the orchids *D. fuchsii* and *D. viridis* may provide insights into how such morphological changes in nectar spur length occur at the genetic level.

Many cases of heterochrony can be observed among angiosperms, see Li & Johnston (2000), strongly suggesting that heterochrony is a popular mechanism employed in the evolution of morphological diversity. Some authors have even claimed that heterochrony **is** the basis of floral shape evolution (Kampny & Harris, 1998), however simple changes in size alone, allometry, are also likely to be important. The apparent importance of heterochrony in generating differences in spur length between *D. fuchsii* and *D. viridis*, and perhaps also the relationship between the gibba and the spur among members of the Antirrhineae, seem to confirm the central importance of heterochrony in the evolution of morphological diversity. In particular, the ontogenetic data presented in this chapter suggest that while nectar spur evolution may have involved a change in spatial location of gene expression (Golz *et al.*, 2002), heterochronic shifts in spur development may be an important factor in generating morphological diversity in spur characteristics.

Chapter 4

Isolation and characterisation of candidate spur-related *KNOX* genes

4.1 Introduction

Knotted1-like homeobox genes (*KNOX*), belong to the Three Amino acid Loop Extension (TALE) superfamily of homeodomain transcription factors (Bertolino *et al.*, 1995; Bürglin, 1997; Vollbrecht *et al.*, 1991). They are a plant specific group of homeodomain transcription factors characterised by the possession of evolutionarily conserved ELK, MEINOX and homeodomain that mitigate protein-protein and protein-DNA interactions (Bellaoui *et al.*, 2001; Chen *et al.*, 2003; Müller *et al.*, 2001; Smith *et al.*, 2002). KNOX transcription factors influence the transcription of a plethora of downstream target genes that have been shown to be integral to a variety of key developmental processes. The *KNOX* gene family is divided into two phylogenetically distinct and well-supported classes based on subtle differences in sequence homology, expression patterns, the positions of conserved introns and the function of the encoded proteins (Kerstetter *et al.*, 1994). The angiosperm class 1 *KNOX* superclade is divided into two subclades, class 1a and class 1b, each divided into two groups defined by the presence of a well-characterised *KNOX* gene. Based on phylogenetic data the class 1a subclade is divided into the *STM* and *KNAT1* groups, characterised by genes of the same name from *Arabidopsis*, while the class 1b subclade is divided into the *KNAT2* and *OSH6* groups, defined by the genes of the same name from *Arabidopsis* and rice, respectively (Bharathan *et al.*, 1999; Jouannic *et al.*, 2007; Reiser *et al.*, 2000).

The principal function of class 1 KNOX transcription factors is in the maintenance of the shoot apical meristem (SAM) by promoting indeterminate cell fate and suppressing differentiation (Endrizzi *et al.*, 1996; Kerstetter *et al.*, 1997; Long *et al.*, 1996; Scofield & Murray, 2006; Vollbrecht *et al.*, 2000). Although there is an increasingly large number

of notable exceptions, class 1 *KNOX* genes are predominantly expressed in specific subdomains of the SAM. Well-characterised *KNOX* genes such as *NTH1*, *NTH15*, *STM*, *OSH1* and *KN1* are expressed in the central or corpus zone (CZ) of the SAM where they function in maintaining the stem cell niche in the centre of the SAM, while *NTH20*, *OSH6*, *OSH15*, *RS1* and *KNAT1/BP* (*BREVIPEDICELLUS*) are expressed in the peripheral zone (PZ) of the SAM where they may function to maintain the organogenic potential in the region that gives rise to the stem and internodes (Hake *et al.*, 2004; Nishimura *et al.*, 1999; Reiser *et al.*, 2000; Sentoku *et al.*, 1999).

By contrast, the function of class 2 *KNOX* genes is relatively poorly understood due to considerably lower experimental investment, largely the result of a lack of available mutants and more ubiquitous expression patterns (Hake *et al.*, 2004; Reiser *et al.*, 2000; Scofield & Murray, 2006). Despite this, some authors have suggested that class 2 *KNOX* genes may function in late stages of plant organogenesis (Serikawa *et al.*, 1997) and root development (Truernit *et al.*, 2007).

In recent years the implication of *KNOX* gene involvement in a number of additional morphogenetic processes, such as complex leaf development (reviewed by Hay *et al.* (2009); Tsiantis & Hay (2003)), has attracted considerable attention to the *KNOX* gene family. Subsequent research strongly suggests that *KNOX* genes have a much broader significance in plant morphogenesis than first thought. The isolation of two independent class 1 *KNOX* gene mutants (*Hirz-d153* and *Ina-d1*) of *Antirrhinum majus*, with novel outgrowths on the petals resembling the floral nectar spurs of closely related species such as *Linaria vulgaris*, supports the assertion that *KNOX* genes may be of great importance to plant morphogenesis.

This provides a unique opportunity to investigate the evolution and development of one such trait (floral nectar spurs) among close relatives of snapdragon. The appearance of spur-like structures in mutants of snapdragon has been shown to result from unusual expression patterns of two class 1 *KNOX* genes, *HIRZINA* (*AmHIRZ*) and *INVAGINATA* (*AmINA*), outside of their normal expression domains in the SAM (Golz *et al.*, 2002). These mutants present the intriguing possibility that alterations in the spatio-temporal patterns of class 1 *KNOX* gene expression may play a key role in the evolution and development of floral nectar spurs, among close relatives of *A. majus* and more broadly among angiosperms.

4.1.1 Research aims and objectives

In order to further our understanding of the role that *KNOX* genes may play in nectar spur development a number of authors have suggested that the orthologous genes of *AmHIRZ* and *AmINA* should be isolated, and their functions characterised, in a close

relative of *A. majus* which bears a true floral nectar spur (Damerval & Nadot, 2007; Galego & Almeida, 2007; Whitney & Glover, 2007). *Linaria vulgaris* (common toadflax) is a close relative of snapdragon with floral nectar spurs that has been used in a number of molecular-developmental analyses, predominantly to supplement developmental-genetic work carried out previously in snapdragon (Almeida *et al.*, 1997; Cubas *et al.*, 1999; Galego & Almeida, 2007; Hileman & Baum, 2003). As such, *L. vulgaris* is an obvious choice of organism in which to develop an understanding of the role that *KNOX* genes may play in nectar spur development.

Whilst *L. vulgaris* provides an excellent developmental model in which to expand studies of nectar spur development, in order to understand the broader significance of *KNOX* genes in the evolution and development of floral nectar spurs additional angiosperm species must be used. Although orchids are by no means a good molecular model, they are a group renowned for considerable phenotypic variation and high species diversity and, as such, are exceptionally well characterised phylogenetically. Furthermore, recent phylogenies of the orchid sub-tribe Orchidinae (Bateman *et al.*, 2003) have implicated changes in nectar spur morphology as a potential agent for recent speciation events in this group. Phylogenetic and ontogenetic analyses clearly indicate that such speciation may be the result of spur reduction in this group of orchids (Bateman, 2005; Bateman & DiMichele, 2002; Box *et al.*, 2008), this study Chapter 3. Such differences may, in part, be related to changes in the timing and/or location of *KNOX* gene expression.

To address these questions the orthologous genes of *AmHIRZ* and *AmINA* from *L. vulgaris* have been isolated and the *KNOX* gene family has been broadly sampled in a number of closely related species of orchid (sub-tribe Orchidinae) with contrasting spur morphology. A combination of degenerate PCR and RACE-PCR techniques have been employed to isolate *KNOX* genes of interest from a mixture of vegetative and floral tissues at appropriate stages of development, as indicated by previous morphological and ontogenetic analyses (Chapter 3). The newly identified genes have been characterised in a phylogenetic context in order to identify those genes most likely to be involved in nectar spur development.

4.1.2 Inferring novel gene function: a phylogenetic approach

Although phylogeny is most commonly associated with reconstructing the past evolutionary relationships of taxa, phylogenetic techniques underlie many of the common bioinformatic tools that are used to characterise novel genes. As soon as a new gene sequence is obtained a number of applications that make use of phylogenetic methods are available to provide a basic understanding of what a novel protein sequence might do. Whether a simple BLAST query or a detailed phylogenetic reconstruction of genes/taxa, the use of

phylogenetic methods in modern evolutionary-developmental science is now a standard part of exploratory sequence analysis (Holder & Lewis, 2003). However, similarity of gene function determined by BLAST alone can be misleading in the absence of a phylogenetic framework. The establishment of phylogenetic relationships between members of gene families facilitates a comparative-developmental approach by providing such a framework in which to analyze the potential function of novel genes. When utilised in the correct phylogenetic context functional data highlighted by BLAST similarity searches can be an effective means by which plausible hypotheses of novel gene function may be determined.

Phylogenetics is necessary to resolve questions of homology both at the molecular and morphological level (Abouheif, 1997; Abouheif *et al.*, 1997). At the molecular level characters are considered to be homologous if they are derived from a common ancestor. Orthology and paralogy are distinct forms of homology. Paralogous genes arise from duplications within an organism whereas orthologous genes derive at the time of divergence between taxa (Fitch, 1970). Complex patterns of gene duplication in different lineages make it difficult to assign orthology in the absence of phylogenetic data, see Chapter 5. Through identification of gene homology or orthology crucial information regarding the ancestral conditions as well as trends in derived characters, such as patterns of gene expression, can be inferred. Such an approach is particularly useful when using non-model organisms, as this can relate novel genes of interest to well-characterised genes from established models such as *Arabidopsis*. A number of authors have used this approach with a high level of success. Particularly good examples include the phylogenetic analyses of novel gene members of the MADS-box family of transcription factors (Theißen *et al.*, 1996). However, phylogenetic techniques in no way replace bona-fide experimental data and are useful only in forming testable hypotheses.

4.1.3 Phylogenetic methods available

There are many different types of phylogenetic analysis that can be implemented (reviewed by Holder & Lewis (2003)). The type of analysis used is determined by a compromise between computational difficulty and the degree of rigour required. The main techniques are distance, parsimony and likelihood (including Bayesian Inference). While it is not appropriate to provide a detailed description and critique of different phylogenetic methods, it is necessary to present a brief appraisal in order to highlight the methodological limitations and to justify the choice of methods used.

4.1.3.1 Neighbour-joining (NJ)

NJ is a *distance* method of phylogenetic analysis. Distance methods such as NJ calculate pairwise distances between sequences and groups of sequences that are most similar to generate a distance matrix. The distance matrix can then be used to generate a tree topology using a minimum evolution (ME) criterion that favours the topology generating the minimum total branch length at each step of the algorithm (Saitou & Nei, 1987). The NJ algorithm iteratively selects a taxon pair, builds a new subtree, and agglomerates the pair of selected taxa to reduce the taxon set by one. While NJ has many benefits such as high reproducibility, low computational demand, high speed and the ability to handle large data sets, many authors prefer to use different methods to generate final tree topologies.

NJ has a number of conceptual flaws that make authors question the biological relevance of NJ tree topologies. The most critical of these is the fact that the observed differences between sequences may not necessarily be an accurate reflection of the evolutionary distances between them (Holder & Lewis, 2003). By compressing sequence information into distances, valuable evolutionary information can be lost, such as multiple substitutions at the same site in a DNA sequence. Such a failing is not so problematic when considering sequences that have diverged only recently but is compounded when inferring older relationships, such as those between *KNOX* genes. In addition, NJ only produces a single tree topology, making it impossible to examine conflicting tree topologies, and often the outcome of distance methods depends on the order in which entities are added to the tree topology. Although several recent modifications to the NJ algorithm have improved the rigour of NJ methods, see Gascuel & Steel (2006), many authors still consider phylogenetic reconstructions generated using NJ to be unrepresentative, favouring phylogenetic methods that are based on explicit criteria such as maximum parsimony (MP) or maximum likelihood (ML).

4.1.3.2 Maximum Parsimony (MP)

Unlike distance methods, MP and ML map the history of gene sequences onto a tree. Maximum parsimony (Fitch, 1971) assumes that shared characters in different entities are derived from common descent, groups of entities based on the possession of such characters are generated and those that require the simplest explanation are considered the correct, or *most parsimonious* one. As multiple different scenarios may be equally plausible many equally parsimonious trees can be generated and represented in a single strict consensus tree. The principal weakness of MP is that all events are considered equally likely. By failing to consider all possible mutational pathways, many equally parsimonious trees are generated that often cannot be resolved into a consensus topology. Unresolved groupings

are represented by multiple branches emanating from the same node of the tree, this is referred to as a polytomy and suggests that the data used for these entities may not be phylogenetically informative. Polytomies can also indicate hybridisation, or that multiple lineages are derived from the same ancestor. In addition, MP assumes that the number of changes between entities should be similar, i.e. that the rate of evolution is constant for all entities. Therefore MP is unable to distinguish two entities which have accumulated large numbers of changes by convergent evolution. This makes MP prone to artefactually group together what are in fact very distantly related entities, a phenomenon called *long-branch attraction* (Felsenstein, 1978).

4.1.3.3 Maximum Likelihood (ML)

ML approaches are commonly used to reconstruct the relationships between sequences that diverged long back in evolutionary time, or sequences that are evolving rapidly (Holder & Lewis, 2003). ML methods compute the probability that a data set fits a tree derived from a specified model of sequence evolution (Farris, 1973; Felsenstein, 1973, 1981; Hendy & Penny, 1989). The model of sequence evolution, i.e. the probability of various evolutionary events such as nucleotide transitions or transversions, can be specified by the user but is most often chosen using software that compares the data set to a number of established models of sequence evolution (section 2.4.2). In this way ML analyses are able to more fully explore all possible mutational pathways without the risk of long-branch attraction, resulting in multiple trees of which one will be most likely. However, ML analyses are computationally highly demanding. While it is true that in many cases technological limitations may hamper the implementation of ML analyses, modern computing clusters make even computationally challenging methods practicable. Despite high computational demands, ML is a robust and reliable method that is often considered the method of choice for phylogenetic reconstruction and has become the benchmark for any good phylogenetic analysis.

Bayesian Inference (BI) is a specific ML method that has recently gained in popularity. In BI a further set of assumptions, termed priors, are input into the original ML model of evolution and the branch swapping algorithms differ (Huelsenbeck & Ronquist, 2001). The field of Bayesian analysis is closely related to ML, the optimal tree is one that maximises the *posterior probability* (defined as the likelihood multiplied by the probability of a prior). BI methods are often faster than ML, generating tree topologies and support values (posterior probabilities) in a single analysis, and often more complex models of sequence evolution can be implemented (Nylander *et al.*, 2004).

Past phylogenetic analyses of spermatophyte *KNOX* genes have largely employed the vastly popular Neighbour-Joining (NJ) and/or Maximum Parsimony (MP) methods of

phylogeny reconstruction (Bharathan *et al.*, 1997, 1999; Bürglin, 1997; Champagne & Ashton, 2001; Chan *et al.*, 1998; Guillet-Claude *et al.*, 2004; Hake *et al.*, 2004; Harrison *et al.*, 2005a,b; Kerstetter *et al.*, 1994; Magnani & Hake, 2008; Reiser *et al.*, 2000; Vollbrecht *et al.*, 1991). Of these, few authors use phylogenetically more informative nucleotide data (Champagne & Ashton, 2001; Harrison *et al.*, 2005a,b) for phylogeny reconstruction; fewer still have employed the use of DNA data in combination with modern phylogenetic techniques such as Maximum Likelihood (ML) (Jouannic *et al.*, 2007; Mukherjee *et al.*, 2009; Sano *et al.*, 2005). Considering the relative sparsity of ML methods in past *KNOX* gene phylogenies, a BI analysis was chosen to characterise the novel *KNOX* genes identified in this work (section 4.2.4). In addition an NJ and MP analysis were also conducted (data not shown) but these trees had insufficient resolution to be informative.

4.2 Results

31 new *KNOX* genes have been identified from nine different species. The potential function and expression patterns of these novel genes were explored using bioinformatic tools in a phylogenetic context.

4.2.1 Degenerate PCR

A successful program of degenerate PCR was developed using hybrid oligonucleotide primers (CODEHOP - Consensus Degenerate Hybrid Oligonucleotide Primers; Rose *et al.* (1998)) designed to amplify distantly related *KNOX* gene sequences based on the highly conserved DNA binding and protein:protein interaction domains of KNOX proteins (see section 2.3.4).

The development of such a successful program of degenerate PCR facilitated the rapid isolation of 31 novel *KNOX* gene fragments from nine different species. A translated nucleotide BLASTX query for each of these novel *KNOX* gene fragments indicates that of the 31 novel fragments isolated, 23 are class 1, and 8 belong to the class 2 *KNOX* superclades (Table 4.1).

Table 4.1: Summary of novel *KNOX* genes identified by dPCR in nine different species; asterisks indicate presence and length of spur, (***) long spur, (**) medium spur, (*) short spur

Number of KNOX genes		
Species	*Class 1*	*Class 2*
*L. vulgaris****	2	1
*D. fuchsii***	4	-
*D. incarnata***	4	1
*D. viridis**	2	1
*G. conopsea****	5	1
*G. odoratissima***	3	1
*G. rhellicani**	1	1
*O. anthropophora**	2	1
*O. italica***	-	1
Total	23	8

4.2.2 Extension of candidate *KNOX* gene fragments by 5′ and 3′ RACE-PCR

Not all of the 31 *KNOX* gene fragments were extended to obtain full-length sequence information. Gene extension techniques were focused on the two class I *KNOX* genes from *L. vulgaris*, with high sequence similarity to *AmHIRZ* and *AmINA*, and the six class 1 *KNOX* gene fragments isolated from the medium and short-spurred orchid species pair of *D. fuchsii* and *D. viridis* to allow comparative expression/protein function analysis between longer and shorter-spurred orchid species. In all cases novel 5′ and 3′ sequence information was obtained using a combination of 5′ and 3′ RACE-PCR (section 2.3.12).

4.2.2.1 Extension of *L. vulgaris KNOX* gene fragments

3′ RACE-PCR was carried out using a combination of the RACE adapter primer B25 and two nested sense primers for each of the two *L. vulgaris* class I *KNOX* gene fragments arbitrarily named *LvKNOX1* (*LvKN1*) and *LvKNOX2* (*LvKN2*). A 10:1 ratio of B25 adapter to specific sense primer was used in each case with a standard PCR thermocycling profile to identify the remaining 3′ coding sequence and UTR for each gene (section 2.3.12.1). An RNA-Ligase Mediated (RLM) 5′ RACE-PCR protocol was employed for extension of the 5′ coding sequence using the GeneRacer® RACE-PCR kit (Invitrogen).

Like 3′ RACE a standard PCR thermocycle was used in combination with an adapter primer (provided with the kit) and two nested antisense primers for each of the two class I *KNOX* gene fragments identified (section 2.3.12.2).

Using this combination of techniques two full-length *KNOX* gene coding sequences were identified (Figure 4.1). Both genes included the full-length coding sequence, 5′ and 3′ UTRs. Each gene sequence was assembled into a contig using overlapping sequence data between each of the gene fragments isolated from the combination of dPCR, 5′ and 3′ RACE-PCR. Once assembled specific primers were designed at each end of the coding sequence and used to PCR amplify the full-length coding sequence using a proof-reading polymerase. Whilst genome walking techniques such as TAIL-PCR were attempted using gDNA as template to identify introns and putative promoter sequence for both *L. vulgaris* *KNOX* gene fragments, such attempts were consistently unsuccessful.

4.2.2.2 Extension of *D. fuchsii* and *D. viridis KNOX* gene fragments

An identical experimental programme was utilised to extend *KNOX* gene fragments from *D. fuchsii* and *D. viridis* (Figure 4.1). Fragment extension in *D. fuchsii* yielded three full-length *KNOX* genes including both 5′ and 3′ UTRs, *DfKNOX1, 2* and *4* (*DfKN1, 2, 4*) and one partially extended *KNOX* gene fragment consisting of the 5′ UTR and most of the coding sequence but lacking the 3′ end of the gene and its UTR, *DfKNOX3* (*DfKN3*). For *D. viridis* fragment extension was only successful at the 3′ end of each of the two *KNOX* gene fragments, *DvKNOX1* (*DvKN1*) and *DvKNOX2* (*DvKN2*). Genome walking techniques such as TAIL-PCR were attempted using gDNA as template for both *D. fuchsii* and *D. viridis KNOX* gene fragments to identify introns and potential upstream promoter sequences but these were consistently unsuccessful.

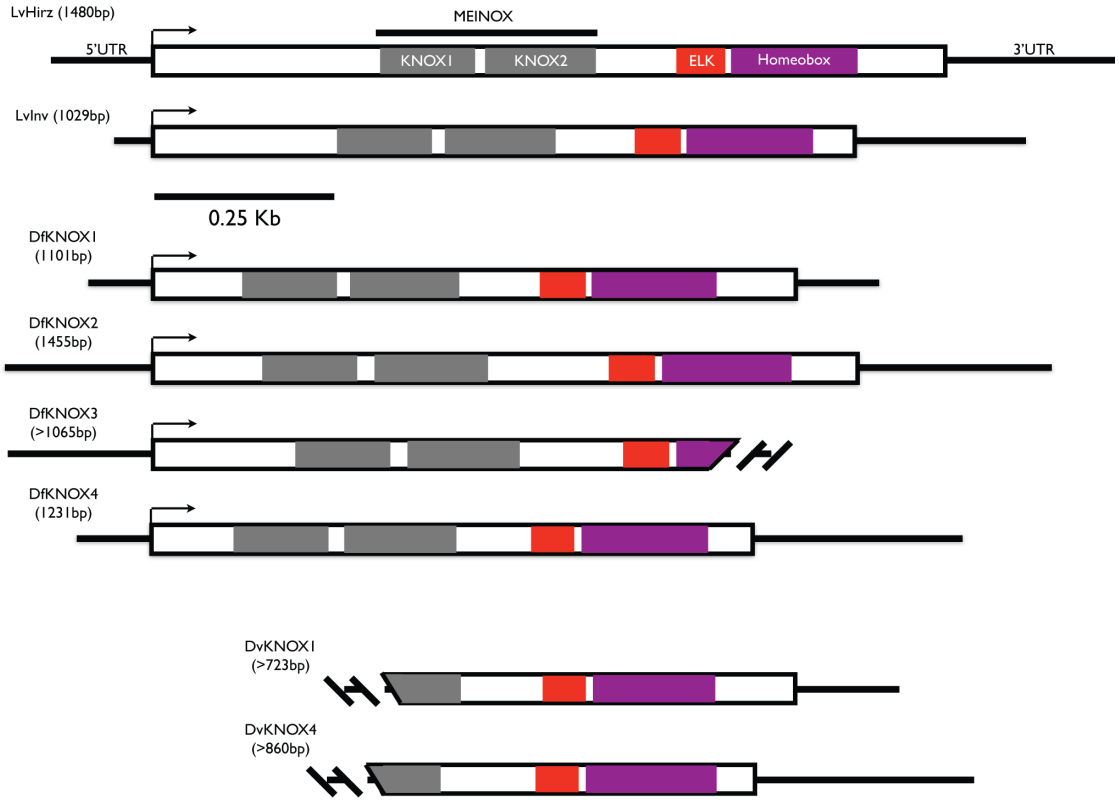


Figure 4.1: Schematic representation of the novel full-length *KNOX* genes identified from *L. vulgaris* *D. fuchsii* and *D. viridis*

Scale representation of the newly identified full-length *KNOX* genes isolated from *L. vulgaris*, *D. fuchsii* and *D. viridis* showing the conserved domain structure, 5′ and 3′ UTRs. The KNOX1 and KNOX2 domains form the characteristic protein binding MEINOX domain (silver), downstream of this are the ELK (red) and Homeobox (purple) DNA binding domains. Missing 5′ and 3′ sequence information is indicated by double hashed lines. The total size of the clone is indicated in brackets.

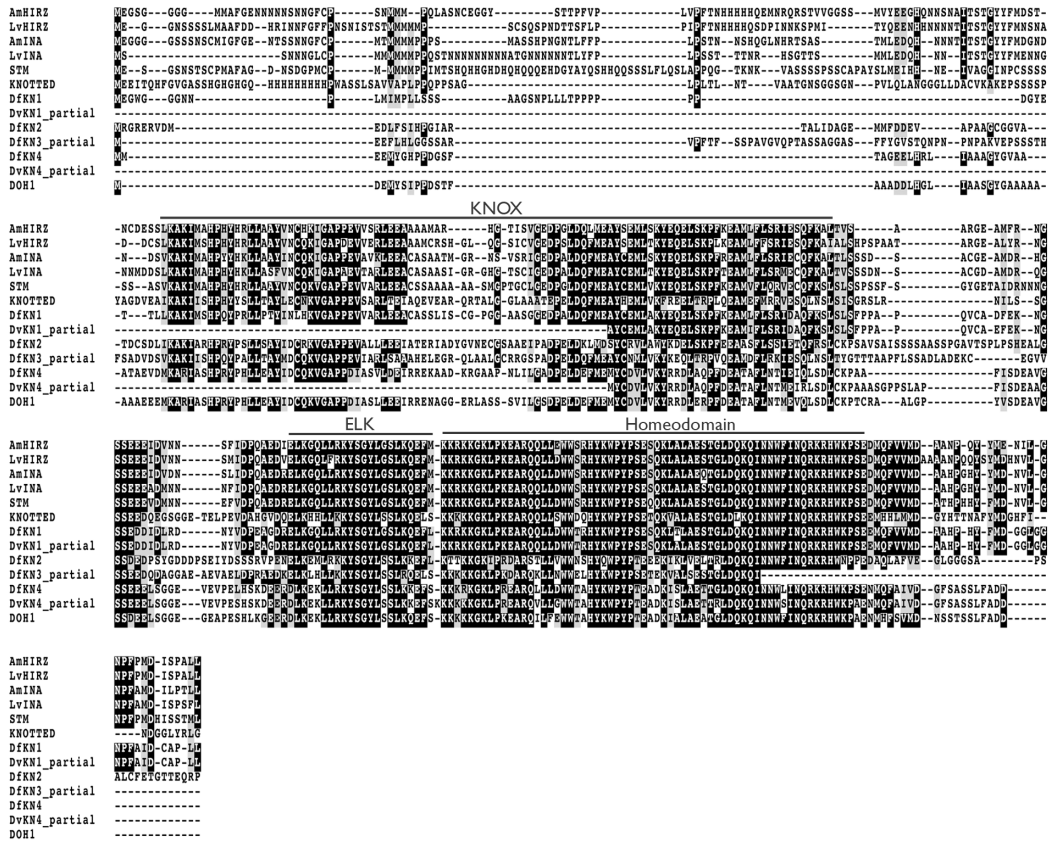


Figure 4.2: ClustalW protein alignment of the novel KNOX sequences identified from *L. vulgaris*, *D. fuchsii* and *D. viridis*

Comparison of the amino acid sequences of the novel KNOX proteins identified from *L. vulgaris*, *D. fuchsii* and *D. viridis* with the landmark KNOX sequences, AmHIRZ (AAL67666.1), AmINA (AAL67665.1), STM (NP 176426.1), Knotted1 (CAA43605.1) and DOH1 (CAB88029.1). Identical amino acids are boxed black; conservative substitutions are boxed grey. Only partial sequence information was recovered for DvKN1, DfKN3 and DvKN4 however, where sequence is available all expected protein domains are represented in each of the novel proteins identified.

4.2.3 Sequence analysis of six novel *KNOX* genes from *L. vulgaris*, *D. fuchsii* and *D. viridis*

Two novel *KNOX* genes were isolated from *L. vulgaris*, four from *D. fuchsii* and two from *D.viridis*. The sequences of each of the genes was characterised using bioinformatic techniques in order to identify potential candidate genes involved in nectar spur development based on their nucleotide and protein level similarity with published *KNOX* gene sequences deposited in GenBank. The composition of each newly identified gene is summarised in table 4.2.

Table 4.2: *KNOX* gene statistics

KNOX gene statistics								
Statistic	*LvKN1*	*LvKN2*	*DfKN1*	*DfKN2*	*DfKN3*	*DfKN4*	*DvKN1*	*DvKN2*
Total (nt)	1449	1249	1081	1436	1065	1204	708	821
5′ UTR (nt)	141	52	89	205	203	103	-	-
CDS (nt)	1101	1029	894	981	862	837	576	558
3′ UTR (nt)	207	168	98	250	-	264	132	263
Protein (aa)	367	343	298	327	287	279	191	185
Complete?	Y	Y	Y	Y	N	Y	N	N

4.2.3.1 Sequence analysis of two novel *L. vulgaris KNOX* genes

The two novel full-length genes isolated from *L. vulgaris* were 1,449 and 1,249bp for *LvKN1* and *LvKN2*, respectively. Each gene included the full-length coding sequence, 5′ and 3′ UTRs. Translating each coding sequence generated proteins of 367 and 343 amino acids in length (Table 4.2) encoding all of the vital domains and motifs typical of other published KNOX transcription factors (Figures 4.1, 4.2). Statistics related to nucleotide identity (identical nucleotides), coverage (the percentage of the queried sequence over which there is similarity) and the protein level identity (identical amino acids) and similarity (identical or substitution with a similar amino acid) were obtained by querying these gene sequences against the NCBI sequence database.

A nucleotide BLAST query of *LvKN1* showed that it is a class 1 *KNOX* gene most similar to *AmHIRZ* (Golz *et al.*, 2002) (82% identity, 77% coverage) the *Antirrhinum* ortholog of *Arabidopsis STM* with which it shares similar expression patterns and function in the SAM. Ectopic expression of *AmHIRZ* generates spur-like petal tubes (Golz *et al.*, 2002). The next most similar well characterised gene is *NTH15* from *Nicotiana*

tabacum which is expressed in the central zone (CZ) of the SAM where it plays a role homologous to that of *STM* (Tamaoki *et al.*, 1997) (78% identity, 62% coverage). A more recently identified monocot *KNOX* gene, *EgKNOX1* from the African oil palm *Elaeis guineensis* (76% identity, 48% coverage), expressed in the CZ of all above ground meristems (Jouannic *et al.*, 2007), also shares high similarity to *LvKN1* at the nucleotide level, particularly in the DNA coding regions that generate the highly conserved ELK domain and homeodomain. A translated BLASTp query shows that at the protein level LvKN1 is also most similar to AmHIRZ (73% identity, 79% similarity) and the well characterised KNOX protein ARBORKNOX1 (Groover *et al.*, 2006) from *Populus alba x tremula* (63% identity, 71% similarity), which is expressed in the SAM and vascular cambium where it functions in the same manner as *STM*.

A nucleotide BLAST query of *LvKN2* shows that it is most similar to other class 1 *KNOX* genes including *AmINA* (Golz *et al.*, 2002) (83% identity, 83% coverage). The next most similar gene is *LeT6/TKN2* (Janssen *et al.*, 1998b), a well characterised ortholog of *STM* from *Solanum lycopersicum* (77% identity, 71% coverage and 79% identity, 55% coverage respectively), which is expressed in the SAM and in developing leaves where it is associated with compound leaf morphogenesis (Chen *et al.*, 1997; Janssen *et al.*, 1998a; Parnis *et al.*, 1997). A translated BLASTp query of LvKN2 shows that at the protein level it is also most similar to AmINA (82% identity, 89% similarity). The next most similar protein is ARBORKNOX1 (73% identity, 83% similarity).

4.2.3.2 Sequence analysis of four novel *D. fuchsii KNOX* genes

Three novel full-length genes isolated from *D. fuchsii* were 1,081, 1,436, 1,204bp in length, *DfKN1, 2* and *4*, including the full-length coding sequences, 5′ and 3′ UTRs. The remaining incomplete gene, *DfKN3*, lacking its 3′ end, was almost complete with a length of 1,065bp. Translating each coding sequence generated complete proteins of 298, 327 and 279 amino acids in length for *DfKN1, 2* and *4*, respectively (Table 4.2). These full-length KNOX proteins contain all of the vital domains and motifs typical of other published KNOX transcription factors (Figures 4.1, 4.2). The truncated *DfKN3 KNOX* gene fragment, missing its 3′ end, translates into an incomplete protein of 287 amino acids. Although this gene is incomplete it contains all of the vital N-terminal domains and approximately 50% of the homeodomain. It is fully expected to encode a typical KNOX protein (Figure 4.1). The nucleotide and amino acid sequences of these four new genes were queried against the NCBI sequence database.

A nucleotide BLAST query of *DfKN1* showed that it is most similar to other class 1 *KNOX* genes including *EgKNOX1* (78% identity, 83% coverage). *DfKN1* is also highly similar to the *STM*-like gene *RaSTM* from the monocot *Ruscus aculeatus* (Asparagaceae)

(Hirayama *et al.*, 2007) (76% identity, 77% coverage), which is expressed strongly in the CZ of the SAM and down regulated in lateral organ primordia. A high level of similarity to the well characterised *KNOX* gene *ROUGH SHEATH1* (*RS1*) from *Zea mays* (Schneeberger *et al.*, 1995) (83% identity, 28% coverage) is also apparent. *RS1* has a subtly different pattern of expression in the SAM and is expressed in the peripheral zone (PZ), subtending lateral organ sites where it is thought to pattern the placement of lateral organs along the axis of the shoot. The similarity between *DfKN1* and *RS1* is primarily focused in the homeodomain encoding region. Another highly characterised gene to which *DfKN1* is similar is *OSH1* from *Oryza sativa* (Matsuoka *et al.*, 1993) (82% identity, 28% coverage), constituting only the ELK and homeodomain encoding regions of the gene. *OSH1* is primarily expressed in the SAM and CZ of the floral meristem where it is thought to function in meristem maintenance. However, *OSH1* transcripts are also detectable in glumes, the bract-like structure associated with the flowers of grasses, which suggests it may also play a role in flower development (Matsuoka *et al.*, 1993). A translated BLASTp query shows that the DfKN1 protein also has a high similarity to the STM-like protein SrSTM1 (Harrison *et al.*, 2005b) from *Streptocarpus rexii* and EgKNOX1 (72% identity, 85% similarity and 72% identity, 79% similarity respectively). *Streptocarpus rexii* has an unusual body plan that lacks a conventional SAM, despite this *SrSTM1* is expressed in meristematic tissues that give rise to new leaf-like structures therefore playing an equivalent role to other *STM*-like genes in plants with a conventional SAM structure (Harrison *et al.*, 2005b).

Analysis of *DfKN2* using a nucleotide BLAST query demonstrated that this gene is most similar to other class 1 *KNOX* genes, in particular a number of STM-like genes from *Oryza sativa* including *Os05g0129700* (Ohyanagi *et al.*, 2006) (65% identity, 79% coverage), *OSH71* (Sentoku *et al.*, 1999) (65% identity, 79% coverage), *HOS9* (Ito *et al.*, 1999) (65% identity, 79% coverage), *Oskn2*, (Postma-Haarsma *et al.*, 1999) (65% identity, 79% coverage) and *OSH6* (Sentoku *et al.*, 1999) (65% identity, 87% coverage). These genes tend to be expressed in the CZ of the SAM and inflorescence/floral meristems where they maintain the meristematic cells. Interestingly, *OHS71* and *OSH6* are expressed between developing lateral organs suggesting they may also play a role in early patterning events (Sentoku *et al.*, 1999). A translated BLASTp query also reflects the high similarity of DfKN2 to KNOX proteins isolated from *O. sativa*, most notably OSH6 (56% identity, 70% similarity). In addition, the DfKN2 protein is also highly similar to LIGULELESS3 (Muehlbauer *et al.*, 1999) from *Z. mays* (53% identity, 67% similarity), which is expressed in all above ground vegetative meristems and developing floral structures.

The incomplete sequence for *DfKN3* was analysed using a nucleotide BLAST query that showed it to be highly similar to other class 1 *KNOX* genes including the *Hordeum*

vulgare knotted class 1 homeobox gene (Krusell *et al.*, 1997) (71% identity, 78% coverage), the barley homolog of maize *KN1*, *RS1* from *Z. mays* (72% identity, 77% coverage) and *Triticum aestivum KNOTTED1*-like homeobox gene b (*Taknox1b*) (Takumi *et al.*, 2000) (70% identity, 78% coverage). A translated BLASTp query of DfKN3 also shows high similarity with the knotted-1-like 2 (KNAP2) (Watillon *et al.*, 1997) protein from *Malus x domestica* (69% identity, 83% similarity), expressed in elongated parts of the stem and the internodes, and the class 1 KNOTTED1-like homeobox protein (IBKN3) (Tanaka *et al.*, 2008) of *Ipomoea batatas* (72% identity, 83% similarity), which is expressed almost constitutively in root tissues and is thought to be involved in storage root development of sweetpotato, perhaps in relation to elevated cytokinin levels. However, it is possible that *IBKN3* is a pseudogene and that *IBKN1* and *IBKN2*, expressed predominantly in the vascular cambium, may be more important in this function (Tanaka *et al.*, 2008).

The remaining full-length 1,204bp *KNOX* gene identified from *D. fuchsii*, *DfKN4*, was analysed using a nucleotide BLAST query that showed it to be highly similar to other class 1 *KNOX* genes. *DfKN4* is most similar to the relatively well characterised class 1 *KNOX* gene *DnSTM1* from the orchid *Dendrobium nobile* (Leng, Ye and Liu, unpublished) (79% identity, 99% coverage), *DOH1* from *Dendrobium* grex Madame Thong-IN (Yu *et al.*, 2000) (79% identity, 99% coverage) and the KNOTTED1-like gene *PKn2* from *Ipomoea nil* (Kobayashi, Suzuki and Yamaguchi, unpublished) (71% identity, 80% coverage). The results of a translated BLASTp query of DfKN4 further demonstrate the particularly high sequence similarity of this protein to other characterised orchid class 1 KNOX transcription factors, sharing high similarity to DOH1 and DnSTM1 (76% identity, 86% similarity for both genes). Both *DOH1* and *DnSTM1* are proposed to function in the maintenance of the SAM where they are both predominantly expressed. However, *DOH1* has very broad patterns of gene expression (Yu *et al.*, 2000), indicating potential involvement in a broad range of morphogenetic processes.

4.2.3.3 Sequence analysis of two novel *D. viridis KNOX* genes

Partial sequences of two *KNOX* gene fragments were isolated from *D. viridis* both lacking 5′ UTRs and partial 5′ coding sequences, *DvKN1* and *DvKN2*. The gene fragments isolated were 708 and 821bp in length and consisted of most of the 3′ coding sequence for each gene and 3′ UTRs. Translating each coding sequence generated incomplete proteins of 191 and 185 amino acids in length (Table 4.2) containing all of the typical C-terminal domains but having only approximately 50% of the MEINOX domain due to the incomplete nature of the sequences (Figures 4.1, 4.2). These genes are fully expected to encode typical KNOX proteins. Despite the incomplete nature of both sequences the nucleotide and amino acid sequences of these two new genes were queried against the

NCBI sequence database.

A nucleotide BLAST query of *DvKN1* showed that it is most similar to other monocot class 1 *KNOX* genes such as *EgKNOX1* (82% identity, 73% coverage), *RS1* from *Z. mays* (82% identity, 39% coverage) and *Taknox1b* from *Triticum aestivum* (83% identity, 36% coverage), the most similar regions of these genes include the ELK domain and homeodomain encoding regions. A translated BLASTp query also reflects the high similarity of this protein to EgKNOX1 (84% identity, 89% similarity). The DvKN1 protein also has high similarity to an STM-like protein from *Kalanchoe daigremontiana* (Garcês *et al.*, 2007) (75% identity, 87% similarity), which appears to be involved in vegetative reproduction via the formation of plantlets along leaf margins. In general BLAST results for *DvKN1* and *DfKN1* are very similar, suggesting that these genes could be orthologous.

The *DvKN2* gene was found to be most similar to other class 1 *KNOX* genes using a nucleotide BLAST query. *DvKN2* is most highly similar to orchid *KNOX* genes from the genus *Dendrobium* including *DOH1* from *Dendrobium* grex Madame Thong-IN (80% identity, 75% coverage), *DnSTM1* from *Dendrobium nobile* (80% identity, 75% coverage) and *OVG2* from *Dendrobium* grex Madame Thong-IN (Yu and Goh, unpublished) (81% identity, 45% coverage over the ELK and homeodomain). A translated nucleotide BLASTp query reflects the high similarity of DvKN2 to other identified orchid KNOX proteins with high similarity to DOH1 (77% identity, 85% similarity) and DnSTM1 (76% identity, 84% similarity). In general the BLAST results for *DvKN2* are very similar to those of *DfKN4*, suggesting that these genes could be orthologous.

4.2.4 Candidate *KNOX* gene alignment and phylogenetic reconstruction

BLAST similarity searches alone are not necessarily informative in the absence of a phylogenetic context that permits the identification of homology. The 31 newly identified *KNOX* gene fragments from all eight species of orchid and *L. vulgaris* were placed in a phylogenetic background of 86 well characterised class 1 and class 2 *KNOX* gene sequences included in the recent phylogenetic analyses of Bharathan *et al.* (1999); Jouannic *et al.* (2007); Reiser *et al.* (2000); Sano *et al.* (2005) (Appendix C). A series of multiple alignments using nucleotide sequence data were generated (section 2.4.1) to identify orthologous groups of *KNOX* genes. The alignment of *KNOX* genes at the nucleotide level was problematic, due largely to considerable sequence variation in the upstream region of *KNOX* sequences and also the fact that the sequences used reflect a considerable amount of taxonomic variation, including 48 sequences from eudicots, 51 from monocots (mostly grasses, Poales), 9 from gymnosperms, 2 from pteridophytes, 3 from lycophytes, 3 from bryophytes and one from the green alga *Acetabularia acetabulum* (*AaKNOX1*), which was

used as the out-group.

To improve confidence in nucleotide alignment four separate alignments were generated using nucleotide data from only the most conserved MD and HD encoding regions of the genes. The four alignments were analysed phylogenetically using Bayesian Inference (section 2.4.2). The parameters for phylogenetic analyses are discussed in section 2.4, links to the precise Bayesian code implemented and raw alignments are included in Appendix C. In the following analyses of the resulting trees, support values at each node are given as posterior probabilities (pp) between 0 and 1.00.

Class 1 and 2 *KNOX* genes are quite different to one another and can be problematic to align even over the highly conserved MD and HD domains. For this reason many published *KNOX* gene phylogenies have focused on the relationships between class 1 *KNOX* genes only. To check the robustness of phylogenetic relationships generated using the complete class 1 and class 2 *KNOX* gene datasets, two HD versus overlapping HD/MD domain alignments were generated using only class 1 *KNOX* genes identified from the literature and confirmed by phylogenetic analyses of the total class 1 and class 2 *KNOX* gene dataset.

4.2.4.1 Phylogenetic reconstruction using the total class 1 and class 2 *KNOX* gene data set

Two alignments were generated using the total gene data set containing both class 1 and class 2 *KNOX* gene sequences (Appendix C). In each case an alignment was generated using data from the highly conserved homeodomain (HD) region of the genes and from an overlapping dataset consisting of sequence information from both the highly conserved MEINOX (MD) and homeodomain regions.

The HD only dataset consists of 496 characters (nucleotides) from representative nucleotide sequences from both class 1 and class 2 *KNOX* genes. BI generated 30,566 trees after 3,000,000 iterations of the BI algorithm. 50% of these trees (15,283) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 4.3). The resulting tree has a highly resolved tree topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 75% of nodes having very high statistical support with posterior probability values (pp) $\geq 0.95(95\%)$.

The overlapping HD/MD dataset consists of 823 characters from representative nucleotide sequences from both class 1 and class 2 *KNOX* genes. BI generated 11,976 trees after 1,370,000 iterations of the BI algorithm. 50% of these trees (5,988) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 4.4). The resulting tree has a highly resolved tree topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 85% of nodes having very high statistical

support with posterior probability values $\geq 0.95(95\%)$.

4.2.4.2 Phylogenetic reconstruction using the class 1 *KNOX* gene data set

The class 1 *KNOX* HD only dataset consists of 505 characters from representative nucleotide sequences from only class 1 *KNOX* gene sequences. BI generated 12,782 trees after 1,130,000 iterations of the BI algorithm. 50% of these trees (6,391) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 4.5). The resulting tree has a highly resolved tree topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 78% of nodes having very high statistical support with posterior probability values $\geq 0.95(95\%)$.

The overlapping class 1 KNOX MD/HD dataset consists of 813 characters from representative nucleotide sequences from only class 1 *KNOX* gene sequences. BI generated 12,172 trees after 1,510,000 iterations of the Bayesian inference algorithm. 50% of these trees (6,086) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 4.6). The resulting tree has a highly resolved tree topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 79% of nodes having very high statistical support with posterior probability values $\geq 0.95(95\%)$.

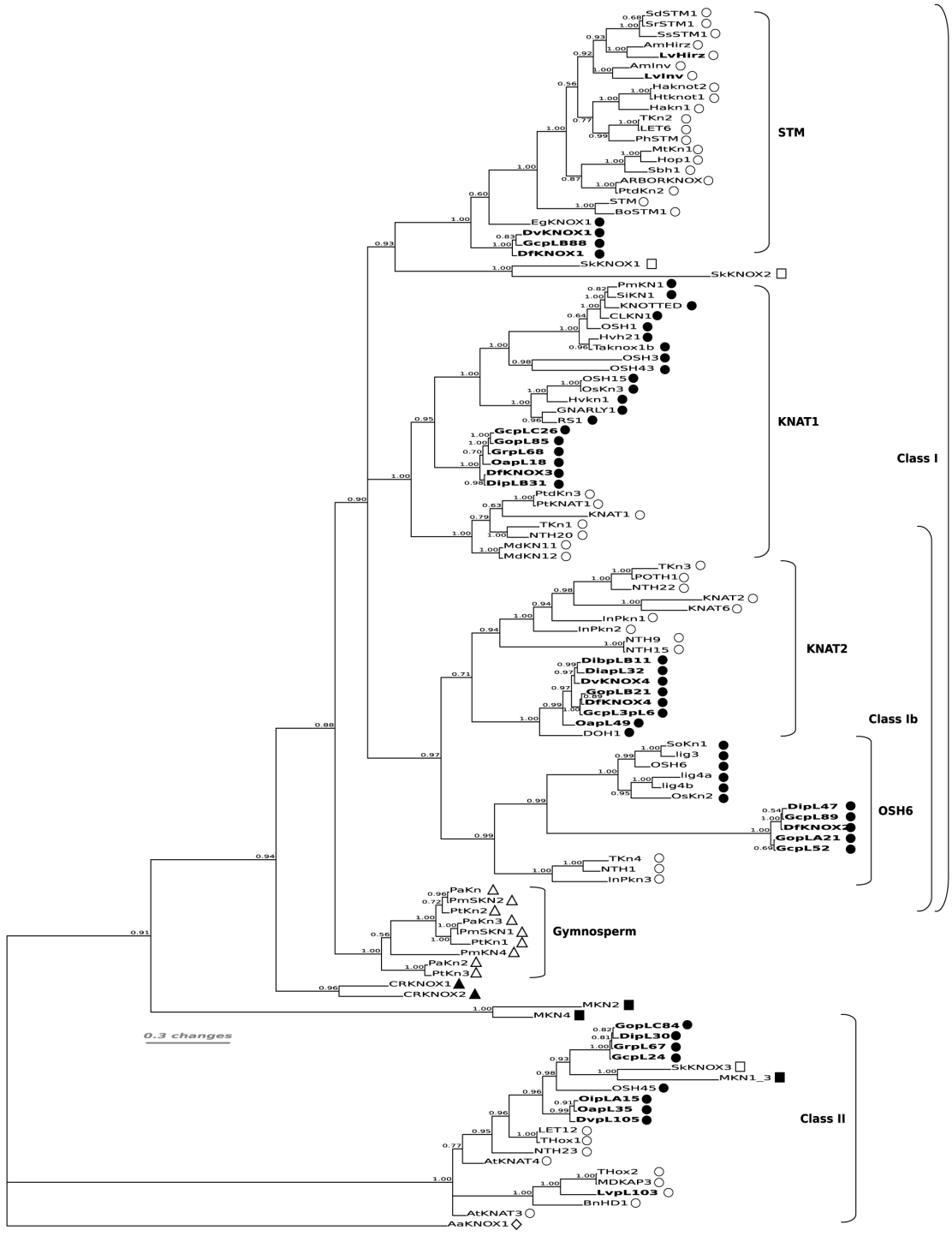


Figure 4.3: 50% majority consensus Bayesian phylogram based on the conserved HD region of 117 class 1 and class 2 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

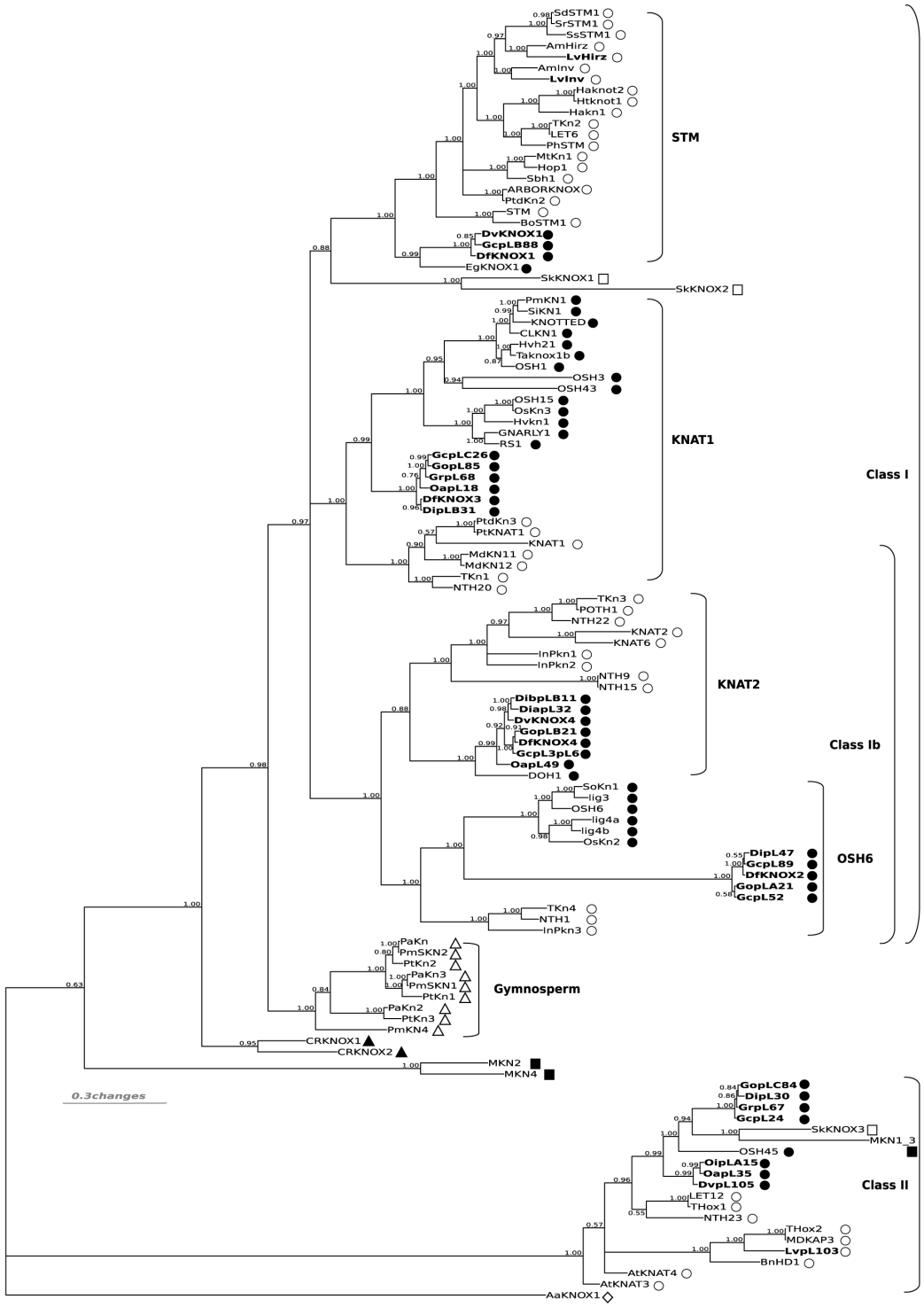


Figure 4.4: 50% majority rule consensus Bayesian phylogram based on the conserved MD and HD region of 117 class 1 and class 2 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

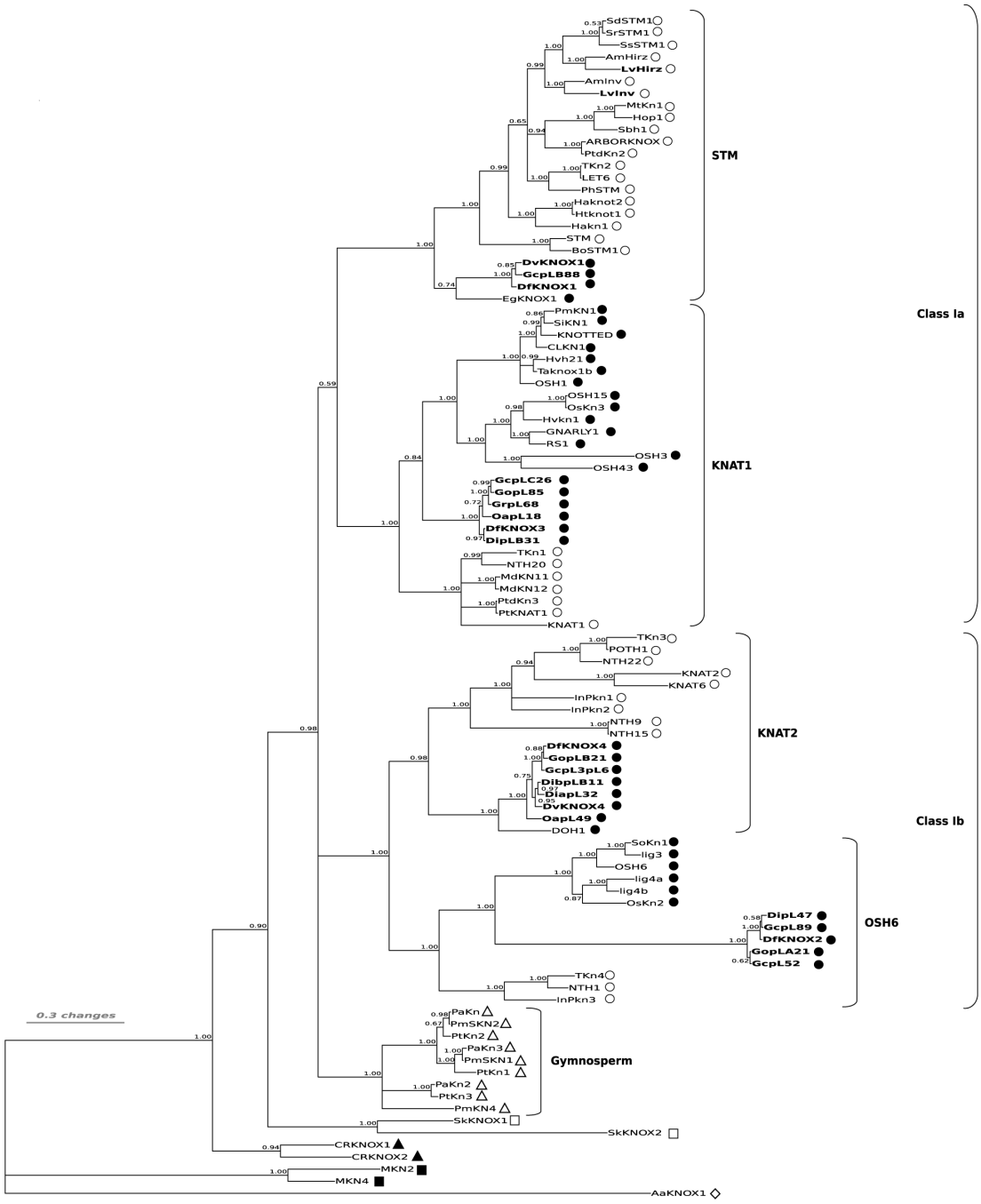


Figure 4.5: 50% majority rule consensus Bayesian phylogram based on the conserved HD region of 98 class 1 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

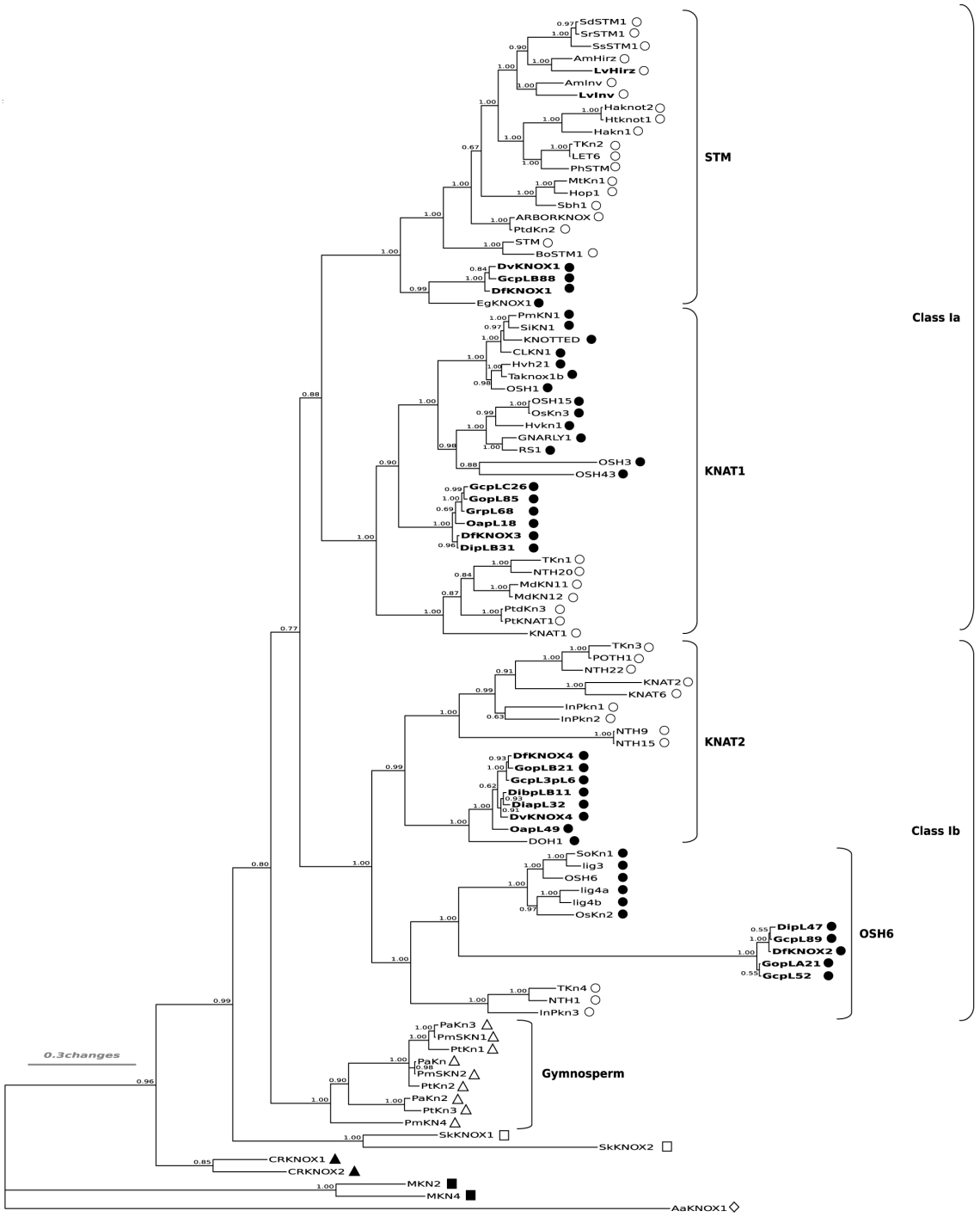


Figure 4.6: Bayesian phylogram based on the conserved MD and HD region of 98 class 1 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

4.2.4.3 Phylogenetic assessment of *LvKN1* and *LvKN2* from *L. vulgaris*

Phylogenetic analyses using different taxon sampling and different regions of the class 1 only or class 1/class 2 combined *KNOX* gene datasets consistently recover *LvKN1* and *LvKN2* from *L. vulgaris* as class 1 *KNOX* genes in the *STM* group. In accordance with the aims, *LvKN1* and *LvKN2* were recovered as orthologs of *AmHIRZ* and *AmINA* respectively. In light of phylogenetic data *LvKN1* and *LvKN2* are subsequently referred to as *LvHIRZINA* (*LvHIRZ*) and *LvINVAGINATA* (*LvINA*) respectively. This name change is reflected in the tree topologies presented earlier. The phylogeny confirms the results of bioinformatic assessment of the two *L. vulgaris* genes and the high degree of similarity to other *STM*-like genes recovered from BLAST analyses such as *EgKNOX1*, *ARBORKNOX1* and *LeT6/TKN2* (section 4.2.3.1). However, phylogenetic analysis does not support similarity of *LvHIRZ/LvINA* to *NTH15* which is recovered in a different clade.

4.2.4.4 Phylogenetic assessment of *DfKN1-4* and *DvKN1, 2* from *D. fuchsii* and *D. viridis*

Of the four *KNOX* genes isolated from *D. fuchsii* all are confirmed as belonging to the class 1 *KNOX* superclade. *DfKN1* and *DvKN1* are orthologous genes confidently recovered in the *STM* group which also contains well characterised sequences such as *AmHIRZ*, *AmINA* from *A. majus*, *STM* from *Arabidopsis*, *LeT6/TKN2* from tomato and *EgKNOX1* from the African oil palm, the only other monocot sequence in this group. These data validate some of the BLAST results obtained for *DfKN1/DvKN1* (sections 4.2.3.2, 4.2.3.3), in particular supporting high similarity to *EgKNOX1* and *SrSTM1* however, other similar genes recovered from BLAST, such as *RS1* and *OSH1*, are recovered in a different clade. This is not surprising as high similarity of *RS1* and *OSH1* is restricted to the highly conserved HD domain.

DfKN2 is recovered in the *OSH6* group alongside well characterised genes such as *OSH6* from rice, *LIGULELESS3* from maize and *NTH1* from *Nicotiana tabacum*. These data strongly confirm the results obtained from BLAST (section 4.2.3.2), indicating high similarity to class 1 *KNOX* genes from rice, such as *Oskn2* and *OSH6*, plus the maize *KNOX* gene *LIGULELESS3*.

DfKN3 is recovered in the *KNAT1* group with many well characterised monocotyledonous *KNOX* genes including *RS1* and *KN1* from maize, plus *OSH1* from rice and *Taknox1b* from *Triticum aestivum*. These relationships strongly support the results of BLAST similarity searches (section 4.2.3.2).

DfKN4 and *DvKN2* are orthologous genes confidently recovered in the *KNAT2* group alongside the well-characterised orchid *KNOX* gene *DOH1*, the only other monocot se-

quence represented in the *KNAT2* group. To reflect the orthology of *DfKN4* and *DvKN2* the *D. viridis* ortholog has been renamed *DvKNOX4* (*DvKN4*), this name change is reflected in the tree topologies presented earlier. The orthology of *DfKN4* and *DvKN2* is supported by the results of the BLAST similarity searches for both genes, which are very similar (section 4.2.3.2, 4.2.3.3).

4.2.5 Estimating the size of the *KNOX* gene family in *D. fuchsii*, *D. viridis* and *L. vulgaris*

Degenerate PCR and subsequent phylogenetic analysis clearly indicates that in orchids such as *D. fuchsii*, *D. incarnata* and *G. conopsea* there are between four and six class 1 *KNOX* genes (Figure 4.6, Table 4.1). In several of the other related orchid species such as *D. viridis*, *G. odoratissima* and *O. anthropophora* fewer class 1 *KNOX* genes were recovered. Rather than indicating that these species have fewer class 1 *KNOX* genes it is more likely that degenerate PCR was merely less successful in these groups. In addition to the class 1 *KNOX* genes isolated from each of the orchid species analysed, with the exception of *D. fuchsii*, a single class 2 *KNOX* gene has also been identified. The results of the degenerate PCR and subsequent phylogenetic analyses (section 4.2.4) strongly suggest that the *KNOX* gene family in these orchids consists of a minimum of five class 1 KNOX genes and one class 2 KNOX gene. The presence of six class 1 KNOX genes in *G. conopsea* suggests that this is likely to be an underestimate.

Exhaustive degenerate PCR performed on *L. vulgaris* however, recovered fewer *KNOX* genes. Only two class 1 *KNOX* genes and a single class 2 *KNOX* gene were recovered from *L. vulgaris* and subsequently shown to be distinct in phylogenetic analyses (section 5.2.3). To estimate the size of the *KNOX* gene family in *L. vulgaris* Southern blotting was used. Four different restriction enzymes (*ApaI, SmaI, SspI* and *XbaI*) were used to digest *L. vulgaris* gDNA extracted from fully expanded seedlings. Using an $[\alpha^{32}P]$-dCTP radio-labelled probe consisting of the highly conserved HD coding sequence of *LvHIRZ*, multiple bands were detected for two of the restriction digests (Figure 4.7). These data suggest the presence of several *KNOX* gene family members in *L. vulgaris*. When digested with *ApaI* two clear bands are visible, consistent with the identification of *LvHIRZ* and *LvINA* by dPCR. Two clear bands are also apparent when gDNA is digested with *SspI* but two additional bands are also evident. These results indicate that in *L. vulgaris* there may be a minimum of four class 1 *KNOX* genes, although it is also possible that *SspI* has cut within the *LvHIRZ* or *LvINA* genes generating a false impression of 4 genes.

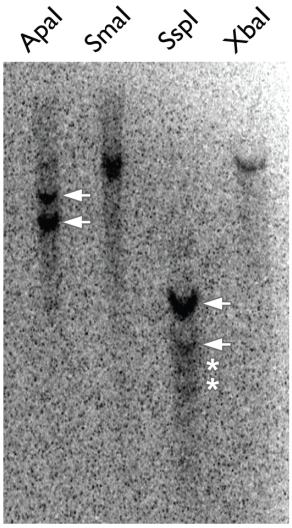


Figure 4.7: Southern-blot analysis of KNOX genes in *L. vulgaris*.
Each lane contains 10μg of digested total gDNA using the restriction enzymes *ApaI*, *SmaI*, *SspI* and *XbaI*. Genomic DNA was prepared from fully expanded seedlings. The digested DNA samples were probed using a [αP^{32}]-dCTP radio-labelled probe of the highly conserved HD encoding region of *LvHIRZ*. Hybridisations and washes were performed under high-stringency conditions. Arrows indicate clearly identifiable bands, less confidently identifiable bands are indicated by an asterisk (*).

4.3 Discussion

In this chapter the isolation and characterisation of 31 novel *KNOX* genes from nine different angiosperm species is described, including *Linaria vulgaris* and several closely related species of orchid in the subtribe Orchidinae (*Dactylorhiza fuchsii*, *D. incarnata*, and *D. viridis*; *Gymnadenia conopsea*, *G. odoratissima*, and *G. rhellicanii*; *Orchis anthropophora* and *O. italica*), with particular emphasis on the *Dactylorhiza fuchsii*/*D. viridis* species pair. The probable function/expression patterns of these novel *KNOX* gene fragments was explored in a phylogenetic background of well-characterised class 1 and class 2 *KNOX* genes.

4.3.1 Degenerate PCR

The use of hybrid oligonucleotide degenerate primers (Rose *et al.*, 1998) targeted to the nucleotide sequences encoding the highly conserved protein-protein and DNA binding domains of KNOX proteins (the DQFM and WFIN motifs in particular), has contributed considerably to the rapid and extensive isolation of *KNOX* gene fragments from a wide range of distantly related angiosperm species. These primers have been shown to be effective on a diverse range of taxa including *Linaria vulgaris* and several closely related species of orchid. Furthermore, these primers have been trialled extensively on additional taxa and have been found to be highly effective at isolating *KNOX* genes from Asterids, Legumes and *Brassica* species (data not shown). The use of degenerate primers targeted to the nucleotide sequences encoding the highly conserved DQFM and WFIN amino acid motifs also permits these primers to amplify both class 1 and class 2 *KNOX* gene fragments with ease. As such these primers represent an extremely useful commodity in the identification and characterisation of novel members of the *KNOX* gene family. The result of such a successful program of degenerate PCR is the identification of 31 new *KNOX* gene fragments from nine different species (section 4.2.1). Subsequently two *KNOX* genes from *L. vulgaris*, four *KNOX* gene fragments from the orchid *D. fuchsii* and two from *D. viridis* were extended to full-length (section 2.3.12) for subsequent expression analyses (Chapters 6, 8).

4.3.2 *LvKN1* and *LvKN2* are the *L. vulgaris* orthologs of *AmHIRZ* and *AmINA*

One of the principle aims of this research was to identify the *L. vulgaris* orthologs of *AmHIRZ* and *AmINA*. Using a combination of phylogenetic reconstruction and bioinformatic tools, the *LvKN1* and *LvKN2* genes isolated from *L. vulgaris* were recovered

as orthologs of *Arabidopsis STM* and *A. majus AmHIRZ* and *AmINA*, isolated by Golz *et al.* (2002). As a result *LvKN1* and *LvKN2* are subsequently referred to as *LvHIRZINA* (*LvHIRZ*) and *LvINVAGINATA* (*LvINA*). Support for these positions is high in each of the phylogenetic reconstructions generated in this work. BLAST analyses of the *LvHIRZ* and *LvINA* nucleotide and protein sequences strongly support their orthology to *AmHIRZ* and *AmINA*. High nucleotide similarity of *LvHIRZ* and *LvINA* is also notable to other *STM* orthologs such as *EgKNOX1* from the African oil palm *Elaeis guineensis* (Jouannic *et al.*, 2007), *ARBORKNOX1* from *Populus alba x Populus tremula* (Groover *et al.*, 2006) and *LeT6/TKN2* from tomato (Janssen *et al.*, 1998b), all of which are expressed in the CZ of the SAM where they function in meristem maintenance. In addition, both *LvHIRZ* and *LvINA* also have high similarity to *NTH15* from tobacco, which, despite similar expression patterns in the CZ of the SAM, is a member of the *KNAT2* group in this analysis but the *STM* group in others, for example Tamaoki *et al.* (1997). Analysis of the inferred amino acid sequence of both the LvHIRZ and LvINA proteins suggest that both genes encode typical class 1 KNOX proteins, 367 and 343 amino acid residues in length (Table 4.2). The polypeptides encoded by each of the genes possess the key conserved domains of class 1 KNOX transcription factors: the MEINOX, GSE/PEST, ELK and HD domains. It is therefore reasonable to assume that both genes encode functional proteins, capable of binding DNA via the HD and forming dimers via the MEINOX domain.

Many of the pitfalls of using phylogenetic data to infer novel gene function are compounded when examining genes from distantly related taxa. However, the relatively close phylogenetic relationships between *L. vulgaris* and *A. majus* (Oyama & Baum, 2004) improve the likelihood that the paralogous *LvHIRZ* and *LvINA* gene pair may have similar patterns of expression and protein function to their *A. majus* orthologs. *Antirrhinum majus HIRZ* and *INA* are paralogs thought to be the result of a relatively recent gene duplication event in the Antirrhineae (Golz *et al.*, 2002). Although, the precise phylogenetic relationship between *Linaria* and *Antirrhinum* is uncertain (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004), they are consistently recovered in the same clade and are, as such, closely related species. In light of phylogenetic data and the presence of *AmHIRZ* and *AmINA* orthologs in *L. vulgaris*, it is highly likely that the duplication event that generated *AmHIRZ* and *AmINA* occurred early in the evolution of Antirrhineae, implying that most, if not all, members of the *Antirrhinum* clade in Antirrhineae possess orthologs of *AmHIRZ* and *AmINA*. Orthologs may also be present in other, less closely related genera in Antirrhineae. Identifying the point at which this duplication event occurred may be useful in understanding the evolutionary origin of floral nectar spurs in this group of plants.

Waites *et al.* (1998) and Golz *et al.* (2002) have shown that *AmHIRZ* and *AmINA*

are predominantly expressed in the shoot apical meristem and provascular cells of internodes. AmHirz and AmINA appear to act redundantly to maintain the SAM, playing a functionally equivalent role to other eudicot orthologs of STM from *Arabidopsis*, tomato (LeT6/TKN2) and soybean (SBH1). Despite being paralogs, *AmHIRZ* and *AmINA* show no strong patterns of neo/sub-functionalisation. Both *AmHIRZ* and *AmINA* are expressed in the central zone (CZ) of the meristem. However, *AmHIRZ* is also expressed in the outermost layer (L1) (Golz *et al.*, 2002), suggesting that there may be some functional differences between the two genes. At least with regard to the SAM, the orthologous genes *LvHIRZ* and *LvINA* isolated in this research are likely to be expressed in a similar pattern to *AmHIRZ*, *AmINA*, *STM*, *LeT6/TKN2* and *SBH1*.

Despite similar expression patterns in the SAM, orthologs of *STM* have evolved additional patterns of expression that appear to be lineage specific (Chapter 5). An excellent example of this is the tomato ortholog *LeT6/TKN2* which is **not** down-regulated in incipient leaf primordia but continues to be expressed in immature leaves where it has been implicated in the development of compound leaves (Chen *et al.*, 1997; Janssen *et al.*, 1998a; Parnis *et al.*, 1997). Golz *et al.* (2002) also demonstrated that in the *Hirz-d153* and *Ina-d1* mutants of snapdragon, *AmHIRZ* and *AmINA* are capable of being missexpressed when the non-coding sequences of the genes are disrupted by transposon insertions. Ectopic expression of *AmHIRZ* and *AmINA* in the *Hirz-d153* and *Ina-d1* mutants is detectable in leaves and petals at a late stage of development. This was correlated with the appearance of a number of pleiotropic phenotypes such as unusually small and rounded leaves with ectopic trichomes and midribs, a buckled lamina as a result of excessive cell proliferation in the central and proximal regions of the leaf and ectopic petal tubes that strongly resemble the nectar spurs of closely related taxa. Golz *et al.* (2002) suggested that the nectar spurs of closely related taxa may have evolved via expansion of *KNOX* gene expression beyond the SAM and into the developing petals. Such extensive patterns of expression that characterise the *Hirz-d153* and *Ina-d1* mutants are unlikely in close relatives of *A. majus* such as *L. vulgaris*, not least because the leaves of *L. vulgaris* are simple and lanceolate. However, it is possible that the expression of *LvHIRZ* and *LvINA* may have extended beyond the confines of the SAM and become expressed in late stages of floral organogenesis, perhaps in relation to the development of floral nectar spurs, in a similar way to *LeT6/TKN2* in the compound leaves of tomato. Detailed expression patterns of *LvHIRZ* and *LvINA* are presented in Chapter 6.

4.3.3 Phylogenetic position and potential functions of novel *D. fuchsii* and *D. viridis KNOX* genes

Whilst the snapdragon mutants *Hirz-d153* and *Ina-d1* provided clear candidates for subsequent isolation and characterisation, the investigation of *KNOX* gene involvement in orchid nectar spur development required a less focused approach. In accordance with this aim, a range of orchid *KNOX* genes were identified from *D. fuchsii* and *D. viridis* that varied significantly in nucleotide sequence. Phylogenetic reconstruction of *D. fuchsii KNOX* genes in particular, recovered a widespread distribution of *KNOX* genes strongly suggesting that most, if not all, class 1 *KNOX* genes have been isolated from *D. fuchsii* and approximately 50% for *D. viridis*. For a more detailed discussion on the size of the *KNOX* gene family in angiosperms see Chapter 5.

Of the four novel *DfKNOX* genes isolated, one is represented in each of the major groups of angiosperm class 1 *KNOX* genes. Both *DfKN1* and *DfKN3* are consistently recovered in the *STM* and *KNAT1* groups of class 1a *KNOX* genes, whilst *DfKN2* and *DfKN4* are confidently recovered in the *OSH6* and *KNAT2* groups of class 1b *KNOX* genes. Phylogenetic analysis of the two *KNOX* genes isolated from *D. viridis* (*DvKN1* and *DvKN2* respectively) strongly supports their positions as orthologs of *DfKN1* and *DfKN4*. As such these genes have been named *DvKN1* and *DvKN4*. Analysis of the inferred amino acid sequences of the proteins encoded by all of the *KNOX* genes isolated from *D. fuchsii* and *D. viridis* strongly suggest that all are likely to encode functional KNOX proteins, despite the truncated nature of the *DfKN3*, *DvKN1* and *DvKN4* gene sequences.

The results of both BLAST and phylogenetic methods were highly congruent, strongly supporting the phylogenetic position of the *D. fuchsii* and *D. viridis KNOX* genes. *DfKN1* and *DvKN1* were both found to be highly similar to other monocot *STM*-like class 1 *KNOX* genes such as *EgKNOX1* and *RaSTM* and more distantly related class 1 *KNOX* genes such as the *KNAT1*-like gene *RS1* from maize (Schneeberger *et al.*, 1995). Both monocot *STM*-like genes *EgKNOX1* and *RaSTM* have similar patterns of expression to *Arabidopsis STM* in the SAM where they are believed to play a role that is functionally homologous to *STM* in meristem maintenance (Hirayama *et al.*, 2007; Jouannic *et al.*, 2007), although *EgKNOX1* may also be involved in complex leaf morphogenesis (Jouannic *et al.*, 2007). Previous monocot *KNOX* genes that have been identified with a function homologous to that of *Arabidopsis STM*, e.g. *KN1* from maize, are not true genetic orthologs and are recovered in a separate clade to *STM*, i.e. the *KNAT1* group. Although monocot *KNOX* genes have been sampled predominantly from grasses, the *STM*-like function of *KN1* prompted theories that monocots in general may have lost orthologs of *STM* (Jouannic *et al.*, 2007). The identification of *EgKNOX1*, *RaSTM* and *DfKN1*/*DvKN1* as

orthologs to *STM* suggests that monocot specific loss of *STM* orthologs is unlikely. In addition, the existence of monocot *KNOX* genes with *STM*-like patterns of expression and function suggest that the *DfKN1* and *DvKN1* genes may function in a similar manner and have similar patterns of SAM specific expression to *STM* from *Arabidopsis*. However, such a function for *DfKN1/DvKN1* remains to be demonstrated.

Given the disparity between gene duplication and functional evolution in the *KNOX* gene family (see Chapter 5), the ancestral role of SAM maintenance may have been retained by the orthologs *DfKN4* and *DvKN4*. Both genes share a high degree of similarity to the orchid *KNOX* genes *DOH1* and *DnSTM1* from *Dendrobium*. Yu *et al.* (2000) proposed that, like all other class 1 *KNOX* genes, *DOH1* is predominantly involved in maintenance of the SAM. However, Yu *et al.* (2000) defined a broad pattern of expression for *DOH1* outside of the SAM, including expression in vegetative apices, transitional buds, floral buds, provascular strands of leaf primordia, inflorescence meristems and floral primordia. Such a broad pattern of expression seems unusual for a single *KNOX* gene, even though extra-meristematic functions have been suggested for a number of more well-characterised genes such as *TKN2* in tomato. It is likely therefore that *DfKN4* and *DvKN4* may have similar patterns of expression to *DOH1* in the SAM but it is difficult to determine if they are also likely to be expressed in such a broad range of tissues, given the stochastic diversification in different angiosperm lineages (see Chapter 5).

DfKN2 is recovered in the *OSH6* group and is highly similar only to other *OSH6*-like sequences such as *OSH6* from rice itself, rice *OsKn2* (Sentoku *et al.*, 1999) and *LIGULELESS3* from maize (Bauer *et al.*, 2004). All are expressed in the PZ of the SAM. Conservation of this pattern of expression seems to be high among closely related members of this group, strongly indicating that *DfKN2* may also share similar patterns of expression. However, all other monocot representatives in this group are grasses, as such expression in the PZ of the SAM may be Poales specific. Eudicot *KNOX* sequences in the *OSH6* group have a very different pattern of expression in the CZ of the SAM (Nishimura *et al.*, 1999).

The phylogenetic position of *DfKN3* suggests it has a high similarity to other monocot *KNOX* genes in the *KNAT1* group such as the maize paralogs *RS1* and *GNARLY* (Jackson *et al.*, 1994; Moore *et al.*, 1995) and the eudicot ortholog *TaKnox1b* from tomato (Takumi *et al.*, 2000). Transcripts of *RS1* and *GNARLY* tend to be excluded from the CZ of the SAM and are restricted to the PZ and between lateral organs where *RS1* expression corresponds with the internode in maize (Jackson *et al.*, 1994; Schneeberger *et al.*, 1995). As discussed earlier, some monocot members of the *KNAT1* group such as *KN1* are expressed in a similar manner to *STM* from *Arabidopsis* but this is likely to be specific to Poales. If *STM*-like expression for *KNAT1*-like genes is specific to Poales it is likely that

DfKN3 may be expressed in a similar manner to *RS1* and its orthologs in the PZ of the SAM in relation to stem and internode development (Hake *et al.*, 2004; Reiser *et al.*, 2000; Schneeberger *et al.*, 1995), rather than in the CZ as is the case for *KN1*. Expression in the PZ of the SAM is also a characteristic of eudicot *KNAT1*-like KNOX genes including *Arabidopsis KNAT1* (Lincoln *et al.*, 1994), tobacco *NTH20* (Nishimura *et al.*, 1999) and tomato *TKN1* (Hareven *et al.*, 1996).

4.3.4 Novel gene function inferred from phylogeny should be treated with caution

Although evolutionary patterns in the *KNOX* gene family are discussed in Chapter 5, it should be noted that hypotheses generated from phylogenetic data relating to novel gene function must be treated with caution. Patterns of gene evolution can be highly complex, and as such, not all phylogenetic groupings necessarily reflect a common function between the gene members. What is evident from the examination of gene function presented in this chapter is that there is no clear correspondence between phylogenetic position and gene function. For example, *Arabidopsis STM* and maize *KN1* are implicated in the maintenance of the SAM and have specific patterns of expression that are functionally equivalent (Jackson *et al.*, 1994; Long *et al.*, 1996). Based on similarity of protein sequence and gene expression patterns a number of authors have suggested that *STM* and *KN1* are orthologous genes. However, phylogenetic analyses (including this one) resolve *STM* and *KN1* into separate clades (Bharathan *et al.*, 1999; Jouannic *et al.*, 2007; Reiser *et al.*, 2000; Sano *et al.*, 2005), suggesting that they are not simply genetic orthologs but represent paralogous lineages derived from an earlier gene duplication event.

Pervasive gene duplication events in the *KNOX* gene family have resulted in complex patterns of functional evolution (see Chapter 5). As such, no clear concept of gene function can be identified from sequence analysis alone. This problem is compounded when comparing evolutionarily distantly related taxa which may have undergone significant diversification in *KNOX* gene function e.g. leaf morphogenesis and nectar spur development. The case of *STM* and *KN1* in particular, highlights the importance of evaluating genetic orthology and functional equivalence as separate entities (Theißen *et al.*, 2002). Bearing this in mind, it is impossible to accurately determine likely candidate *KNOX* genes involved in nectar spur development based on sequence similarity alone.

4.3.5 *KNOX* genes are members of multi-gene families in *L. vulgaris*, *D. fuchsii* and *D. viridis*

Most plant genes are members of multigene families (Wendel *et al.*, 2002). This reflects the importance of gene duplications in plant evolution. In *Arabidopsis thaliana* 65% of all genes belong to gene families, 37.4% of which are members of families composed of at least five genes (Yuan *et al.*, 2002). Given the ancient nature of *KNOX* genes and the apparent commonality of gene duplication events, it is reasonable to assume that most, if not all, angiosperms have a complement of several *KNOX* genes, consisting of at least 4 class 1 and a single class 2 *KNOX* gene. This assertion is strongly supported by the identification of five class 1 *KNOX* genes from polyploid tobacco (although there may be many more) (Nishimura *et al.*, 1999), nine from the polyploid maize (Reiser *et al.*, 2000), at least five from rice (Sentoku *et al.*, 1999), four from *Arabidopsis* (Dean *et al.*, 2004), four from tomato (Reiser *et al.*, 2000) and five from snapdragon (Golz *et al.*, 2002). Further sampling in light of the recent availability of genome sequences for many model plants, has identified eight class 1 *KNOX* genes in *Arabidopsis*, 15 in Poplar, 12 in rice (plus one pseudogene), and 13 in maize, (Mukherjee *et al.*, 2009); in each case *KNOX* gene representatives are found in most, if not all, of the major *STM*, *KNAT1*, *KNAT2* and *OSH6 KNOX* gene groupings. This likely reflects evolutionarily ancient gene/genome duplication events.

Prior to this work few *KNOX* genes had been identified and characterised in orchids. Only single *KNOX* genes have been isolated from the orchids *Dendrobium nobile* (*DnSTM1*; Leng, Ye and Liu, unpublished) and *Dendrobium* grex Madame Thong-IN (*DOH1*; Yu *et al.* (2000)). (Yu *et al.*, 2000) suggested that *Dendrobium* may have only a single *KNOX* gene, implying that specific lineages of orchid, or orchids in general, may be distinct from other angiosperms which tend to have multiple *KNOX* genes. However, the identification of multiple *KNOX* genes from orchid species in the subtribe Orchidinae, disagrees strongly with this hypothesis. Multiple *KNOX* genes have been isolated from all eight species of orchid sampled in this work. Most notable are *D. fuchsii* and *D. incarnata* from which 4 class 1 *KNOX* genes were isolated and *G. conopsea*, which has at least 5 class 1 *KNOX* genes, being represented twice in the *OSH6 KNOX* gene grouping. In each of these three species at least one *KNOX* family member is represented in each of the major *KNOX* gene subgroups. Although sampling was not equally successful for all of the orchid species, the close phylogenetic relationships between the orchid taxa sampled, suggest that they all possess a minimum of four class 1 *KNOX* genes and a single class 2 *KNOX* gene. As a practical consideration, these results also indicate that the *KNOX* gene family has been sampled extensively in *D. fuchsii* which has been used in subsequent experiments related to nectar spur development. Based on these data it is likely that the

Dendrobium genome also carries multiple *KNOX* gene orthologs and paralogs.

Sampling of *KNOX* genes from *L. vulgaris* was particularly poor using degenerate PCR alone. Only two class 1 *KNOX* genes were identified, along with a single class 2 *KNOX* gene. To date, five class 1 *KNOX* genes have been identified from *A. majus* (Golz *et al.*, 2002). As *A. majus* is a reasonably close relative of *L. vulgaris*, it seems likely that there are also five class 1 *KNOX* genes in *L. vulgaris*. This assertion is supported by Southern blot analysis, which suggests the existence of at least four class 1 *KNOX* genes in *L. vulgaris*. The Southern blot was probed with the highly conserved HD region of the *LvHIRZ* gene under high-stringency conditions. Due to significant nucleotide sequence variation between class 1 and class 2 *KNOX* genes, even in the HD region, this Southern blot most likely represents only class 1 *KNOX* genes and so cannot confirm the presence of any additional class 2 *KNOX* genes in *L. vulgaris*.

4.3.6 Conclusions

A successful program of degenerate PCR has resulted in the isolation of the *L. vulgaris* orthologs of *AmHIRZ* and *AmINA* and a representative group of *KNOX* gene candidates from the medium-spurred orchid *D. fuchsii* and the short-spurred orchid *D. viridis*. Placing these genes in a phylogenetic context has resulted in a highly supported tree topology that corroborates novel gene identities as determined by bioinformatic techniques. While phylogenetic techniques are a useful tool for the assessment of novel gene function, the pervasive gene duplication events among *KNOX* genes, their apparently evolutionarily labile nature, poor taxonomic sampling, high levels of functional redundancy and co-option in other developmental processes, makes it particularly difficult to infer novel *KNOX* gene function from phylogeny. This situation is likely to improve with broader taxonomic sampling. However, this is unlikely to resolve the problem completely, especially between phylogenetically distant taxa that may have lineage specific patterns of *KNOX* gene evolution. Predictions of novel gene function from sequence analysis is therefore no substitute for experimental analyses of gene function.

Chapter 5

A phylogenetic re-appraisal of the *KNOX* gene family

5.1 Introduction

The phylogenetic trees of the *KNOX* gene family recovered in Chapter 4 were generated with the intention of providing a phylogenetic context in which hypotheses concerning the probable functions of novel *KNOX* gene candidates could be assessed. However, in addition to resolving questions of homology, these phylogenetic trees allow us to explore the evolution of gene families. A number of previous broad scale analyses of the *KNOX* gene family have been published in the last 15 years, many as an accompaniment to a newly isolated *KNOX* gene, for example Harrison *et al.* (2005a,b); Hirayama *et al.* (2007); Jouannic *et al.* (2007). However, phylogenetic trees have also been generated with the intention of exploring larger evolutionary patterns in the *KNOX*/homeobox gene family as a whole (Bharathan *et al.*, 1999; Champagne & Ashton, 2001; Mukherjee *et al.*, 2009; Reiser *et al.*, 2000; Sano *et al.*, 2005). Whilst such trees have proven reasonably robust in their determination of some broad patterns of evolution in the *KNOX* gene family, many of these analyses have suffered from narrow taxon sampling, particularly among monocots, and the use of unsatisfactory phylogenetic techniques and less informative protein data (Bharathan *et al.*, 1997, 1999; Bürglin, 1997; Champagne & Ashton, 2001; Chan *et al.*, 1998; Guillet-Claude *et al.*, 2004; Hake *et al.*, 2004; Harrison *et al.*, 2005a,b; Kerstetter *et al.*, 1994; Magnani & Hake, 2008; Reiser *et al.*, 2000; Vollbrecht *et al.*, 1991). Few authors use phylogenetically more informative nucleotide data (Champagne & Ashton, 2001; Harrison *et al.*, 2005a,b) for phylogeny reconstruction and fewer still have employed the use of DNA data in combination with modern phylogenetic techniques such as Maximum Likelihood (ML) (Jouannic *et al.*, 2007; Mukherjee *et al.*, 2009; Sano *et al.*, 2005). As such some of the theories pertaining to broad scale patterns of evolution in the

KNOX gene family remain contentious, or are only poorly supported. The *KNOX* gene tree topologies presented in this work represent the most robust exploration of *KNOX* gene phylogeny to date and provide an improved understanding of evolutionary patterns in the *KNOX* gene family.

5.1.1 The present understanding of *KNOX* gene evolution

In plants, *KNOX* genes fall into two classes, each distinguished by the similarity of residues within the homeodomain and in the positioning of introns (Kerstetter *et al.*, 1994). Many previous phylogenies of *KNOX* genes in plants strongly support these broad phylogenetic groupings (Bharathan *et al.*, 1999; Champagne & Ashton, 2001; Jouannic *et al.*, 2007; Magnani & Hake, 2008; Mukherjee *et al.*, 2009; Reiser *et al.*, 2000; Sano *et al.*, 2005; Serikawa & Mandoli, 1999). Since both clades contain representatives from angiosperms, gymnosperms, ferns and bryophytes, these phylogenetic analyses strongly suggest that the class 1 and class 2 *KNOX* genes have resulted from an evolutionarily ancient gene duplication that occurred approximately 400 MYA, before the evolution of seed plants (Champagne & Ashton, 2001; Sano *et al.*, 2005). Furthermore, the identification of an individual *KNOX* gene with both class 1 and class 2 *KNOX* gene characteristics in green algae such as *Acetabularia* (Serikawa & Mandoli, 1999) and *Chlamydomonas* (Mukherjee *et al.*, 2009), presents the possibility that this duplication event may be even older, occurring as early as 500 MYA, concomitant with the evolution of multicellular plant bodies (Mukherjee *et al.*, 2009; Sano *et al.*, 2005).

The evolutionary history of plant *KNOX* genes reflects the complex patterns of genome duplication and gene loss that have characterised the evolutionary history of land plants, sharing a particularly close association with increasing complexity of plant body plans (Mukherjee *et al.*, 2009). Several recent phylogenetic analyses suggest that, in addition to the ancient duplication event that has given rise to class 1 and class 2 *KNOX* genes, at least a further three additional duplication events have occurred during angiosperm *KNOX* gene evolution (Bharathan *et al.*, 1999; Jouannic *et al.*, 2007; Reiser *et al.*, 2000; Sano *et al.*, 2005) generating the strongly supported classes 1a and 1b *KNOX* gene groups and their subsequent *STM, KNAT1* and *KNAT2, OSH6* subgroupings, respectively. Three duplication events have also been identified among gymnosperm *KNOX* genes and have been implicated in functional divergence (Guillet-Claude *et al.*, 2004). In addition to these major gene duplication events it seems that in some angiosperm lineages, particularly among the grasses maize and rice, there have been additional and extensive lineage specific gene duplication events and considerable functional divergence (Jouannic *et al.*, 2007), most likely related to whole genome duplications.

5.1.2 Possible consequences of *KNOX* gene duplications

Gene duplications serve as a mechanism to increase diversity at the molecular level (Averof & Akam, 1995; Ohno, 1970) and may lead to the evolution of innovative developmental programs via changes in protein sequence or, more often, patterns of gene expression (Meyer, 1996). Duplication events such as these are commonplace among plants, where they occur at an unusually high rate compared to other divisions of life (Lawton-Rauh, 2003). The majority of new genes originate through duplication, chromosomal rearrangement, and the subsequent divergence of pre-existing genes (Bennetzen, 2002). The presence of additional copies of a gene alleviates constraints on evolution from natural selection, which in combination with positive/negative directional selection, may result in pseudogenization, functional redundancy, subfunctionalization or neofunctionalization (Lawton-Rauh, 2003; Li, 1983; Ohno, 1970; Zhang *et al.*, 2002). Many of the alterations that result from gene duplication events are assumed to change amino acid sequence. However, gene duplication may also result in functional diversification by retention of ancestral function but expansion of spatio-temporal patterns of gene expression (Averof & Akam, 1995; Doebley & Lukens, 1998; Meyer, 1996). A number of authors have suggested that the pervasive gene duplication events in the *KNOX* gene family may have been important in the diversification of seed plant body plans (Mukherjee *et al.*, 2009; Sano *et al.*, 2005), particularly in angiosperms, in the same way that MADS-box gene duplications have been integral to the evolution of diverse floral morphologies (Theißen *et al.*, 1996). How *KNOX* gene evolution has contributed to morphological and functional diversification is relatively poorly understood.

5.1.3 Research aims and objectives

The principal aim of this chapter is to examine the patterns of *KNOX* gene evolution using the phylogenetic trees generated in Chapter 4 and to explore the probable roles of *KNOX* gene duplication events in the evolution of morphological diversity.

5.2 Results

The 31 newly identified *KNOX* gene fragments from all eight species of orchid and *L. vulgaris* were placed in a phylogenetic background of 86 well characterised class 1 and class 2 *KNOX* gene sequences included in the recent phylogenetic analyses of Bharathan *et al.* (1999); Jouannic *et al.* (2007); Reiser *et al.* (2000); Sano *et al.* (2005) (Appendix C). A series of multiple alignments using nucleotide sequence data were generated (section 2.4.1) to identify orthologous groups of *KNOX* genes. The alignment of *KNOX* genes

at the nucleotide level was problematic, due largely to considerable sequence variation in the upstream region of *KNOX* sequences and also the fact that the sequences used reflect a considerable amount of taxonomic variation, including 48 sequences from eudicots, 51 from monocots (mostly grasses, Poales), 9 from gymnosperms, 2 from pteridophytes, 3 from lycophytes, 3 from bryophytes and one from the green alga *Acetabularia acetabulum* (*AaKNOX1*), which was used as the out-group (see section 4.2.4). For ease of reading the phylogenetic trees generated in Chapter 4 are repeated in this chapter.

5.2.1 Assessing class 1 and class 2 *KNOX* gene relationships

5.2.1.1 Total *KNOX* phylogeny using only HD sequence data

The HD only dataset consists of 496 characters (nucleotides) from representative nucleotide sequences from both class 1 and class 2 *KNOX* genes. BI generated 30,566 trees after 3,000,000 iterations of the BI algorithm. 50% of these trees (15,283) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 5.1). The resulting tree has a highly resolved tree topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 75% of nodes having very high statistical support with posterior probability values (pp) $\geq$ 0.95(95%).

The tree supports the monophyly of class 1 angiosperm *KNOX* genes (pp=0.90), as well as gymnosperm (pp=1.00) and pteridophyte (pp=0.96) *KNOX* genes which are recovered below class 1 *KNOX* genes from angiosperms. The tree also supports strong monophyly of the class 2 *KNOX* genes (pp=1.00). However, the class 2 *KNOX* gene superclade includes representative sequences from eudicots, monocots, lycophytes and bryophytes.

Angiosperm class 1 *KNOX* genes are divided into several distinct groups identified in previous phylogenetic analyses, with each class 1 *KNOX* group including representatives from both the eudicots and monocots. The four major class 1 *KNOX* groups include the *STM* group, containing sequences similar to *STM* from *Arabidopsis* (pp=1.00); the *KNAT1* group, with sequences similar to *KNAT1* from *Arabidopsis* (pp=1.00), *OSH1* from rice and *KN1* from maize; the *KNAT2* group, containing sequences similar to *KNAT2* from *Arabidopsis* (pp=0.71), and the *OSH6* group, containing sequences similar to *OSH6* from rice (pp=0.99). The specific relationships between these class 1 groups cannot be resolved with this tree due to polytomy but it is obvious that the *OSH6* and *KNAT2* groups are more closely related to one another than to the *STM* and *KNAT1* groups. The *OSH6* and *KNAT2* groups form part of a larger monophyletic class 1b *KNOX* gene clade (p=0.97) as described in previous phylogenetic frameworks.

5.2.1.2 Total *KNOX* phylogeny using MD/HD sequence data

The overlapping HD/MD dataset consists of 823 characters from representative nucleotide sequences from both class 1 and class 2 *KNOX* genes. BI generated 11,976 trees after 1,370,000 iterations of the BI algorithm. 50% of these trees (5,988) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 5.2). The resulting tree has a highly resolved tree topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 85% of nodes having very high statistical support with posterior probability values $\geq 0.95(95\%)$.

This tree topology is similar to that generated using nucleotide data only from the HD encoding region but with improved support for the monophyly of class 1 angiosperm *KNOX* genes (pp=0.98), as well as gymnosperm (pp=1.00) and pteridophyte (pp=0.95) *KNOX* genes which are placed basally to class 1 *KNOX* genes from angiosperms. The tree supports strong monophyly of the class 2 *KNOX* genes (pp=1.00). However, the class 2 *KNOX* superclade also includes representative sequences from eudicots, monocots, lycophytes and bryophytes, as was the case for the tree generated using the HD only dataset.

As before, angiosperm class 1 *KNOX* genes are divided into several distinct groups containing representatives from both the eudicots and monocots. Statistical support for monophyly of the *KNAT2* and *OSH6* groups is marginally improved (pp=0.88 and 1.00 respectively). Precise relationships between the class 1 groups are still unresolvable due to polytomy. However, the subdivision of *KNAT2* and *OSH6* groups to form the class 1b *KNOX* gene clade is still highly supported (p=1.00). The odd position of the lycophyte genes *SkKNOX1* and *SkKNOX2* as sister to the *STM* class 1 *KNOX* group is less well supported than when using only the HD dataset and remains anomalous (pp=0.88). The overall topology of the tree is maintained but with some minor alterations such as the position of *OSH1* relative to *Taknox1b/Hvh21*; the *TKN1/NTH20* and *MdKN11/MdKN12* pairing, plus the relationship between *MKN4/MKN2* and *AtKNAT4/AtKNAT3*.

5.2.2 Assessing class 1 *KNOX* gene relationships

5.2.2.1 Class 1 *KNOX* phylogeny using only HD sequence data

The class 1 *KNOX* HD only dataset consists of 505 characters from representative nucleotide sequences from only class 1 *KNOX* gene sequences. BI generated 12,782 trees after 1,130,000 iterations of the BI algorithm. 50% of these trees (6,391) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 4.5). The resulting tree has a highly resolved tree topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 78% of nodes having very high statistical

support with posterior probability values $\geq 0.95(95\%)$.

The tree is unable to support the strict monophyly of class 1 angiosperm *KNOX* genes which form a polytomy with a highly supported monophyletic gymnosperm *KNOX* gene clade (pp=1.00). This angiosperm/gymnosperm polytomy is however distinct from the *KNOX* genes of non-seed plants (pp=0.98). The lycophyte sequences *SkKNOX1* and *SkKNOX2* are placed outside the angiosperm/gymnosperm class 1 *KNOX* polytomy (pp=0.90) in a well supported clade of their own (pp=1.00). Located outside this are the pteridophyte *KNOX* genes which are recovered basally to class 1 *KNOX* genes from angiosperms, gymnosperms and lycophytes (pp=0.94). The bryophyte genes *MKN2* and *MKN4* form the basal most clade of class 1 *KNOX* genes (pp=1.00). The overall topology of this tree is similar to those generated using the total *KNOX* gene datasets with a few minor alterations to the topology at the very tips of the branches. Despite minor topological alterations at the tips, the major branching topology is well maintained.

As was the case for the trees generated using the combined class 1 and 2 *KNOX* gene datasets, class 1 *KNOX* genes from angiosperms are divided into several distinct groups containing representatives from both the eudicots and monocots. Statistical support for monophyly of these groups is very high, *STM* (pp=1.00), *KNAT1* (pp=1.00), *KNAT2* (pp=0.98) and *OSH6* (pp=1.00). As described in previous analyses of class 1 *KNOX* genes the four major class 1 *KNOX* gene groupings can be divided between two subclades. The *STM* and *KNAT1*-like *KNOX* gene sequences form a large class 1 *KNOX* gene subclade referred to as the class 1a *KNOX* genes, which is only weakly supported in this analysis (pp=0.59). The *KNAT2* and *OSH6*-like class 1 *KNOX* genes also form a subclade of genes referred to as the class 1b subclade (pp=1.00) which was also apparent when using the total *KNOX* gene datasets.

5.2.2.2 Class 1 *KNOX* phylogeny using MD/HD sequence data

The overlapping class 1 KNOX MD/HD dataset consists of 813 characters from representative nucleotide sequences from only class 1 *KNOX* gene sequences. BI generated 12,172 trees after 1,510,000 iterations of the Bayesian inference algorithm. 50% of these trees (6,086) were used to generate a consensus tree rooted with the algal *KNOX* gene *AaKNOX1* (Figure 4.6). The resulting tree has a highly resolved topology. All nodes of the tree have a minimum posterior probability value of 0.5 (50%) with 79% of nodes having very high statistical support with posterior probability values $\geq 0.95(95\%)$.

The tree supports the monophyly of class 1 angiosperm *KNOX* genes (pp=0.77) and the monophyletic gymnosperm *KNOX* gene clade (pp=1.00). The lycophyte sequences *SkKNOX1* and *SkKNOX2* are placed outside of the angiosperm and gymnosperm class 1 *KNOX* clades with slightly lower support than in the tree derived from the HD only

dataset (pp=0.80) but in a well supported clade of their own (pp=1.00). Recovered outside of this are the pteridophyte *KNOX* genes which are basal to class 1 KNOX genes from angiosperms, gymnosperms and lycophytes (pp=1.00). The bryophyte genes *MKN2* and *MKN4* form the basal most clade of class 1 *KNOX* genes (pp=1.00) near the root of the tree. The overall topology of this is largely identical to that generated using only the class 1 *KNOX* HD dataset with a few minor alterations to the topology and resolution at the very tips of the branches. Despite minor topological alterations at the tips the major branching topology is well maintained with improved support.

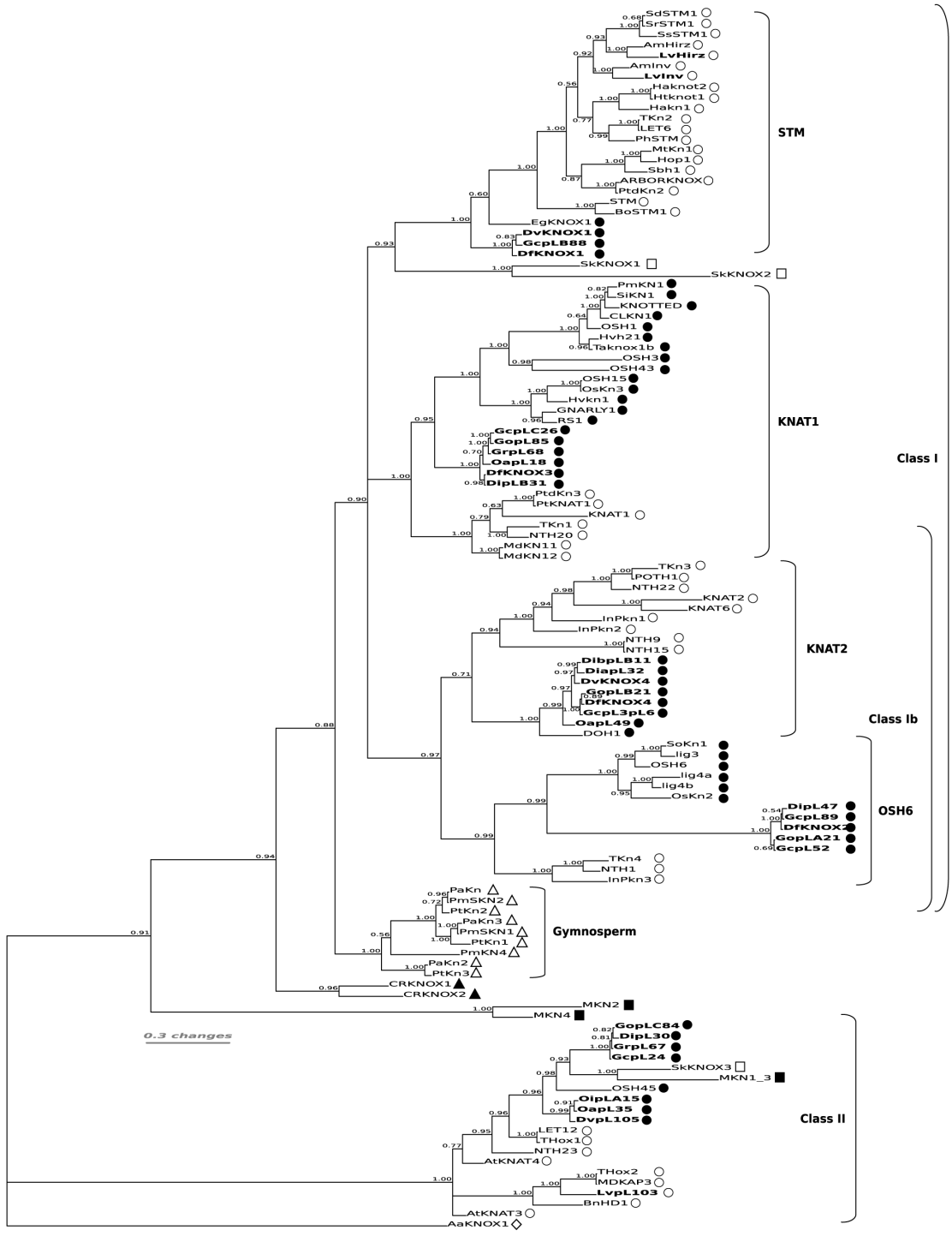


Figure 5.1: 50% majority consensus Bayesian phylogram based on the conserved HD region of 117 class 1 and class 2 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

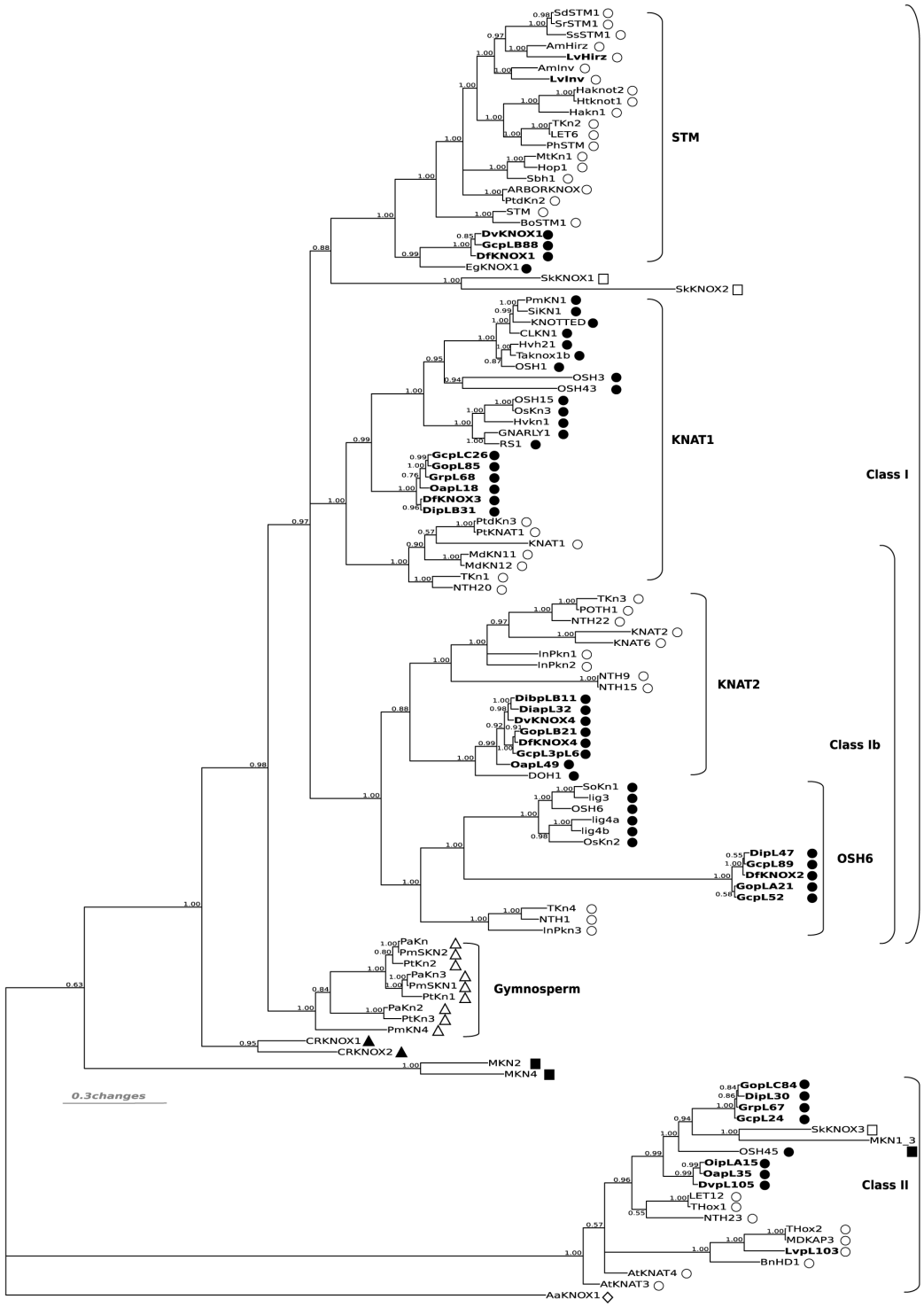


Figure 5.2: 50% majority rule consensus Bayesian phylogram based on the conserved MD and HD region of 117 class 1 and class 2 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

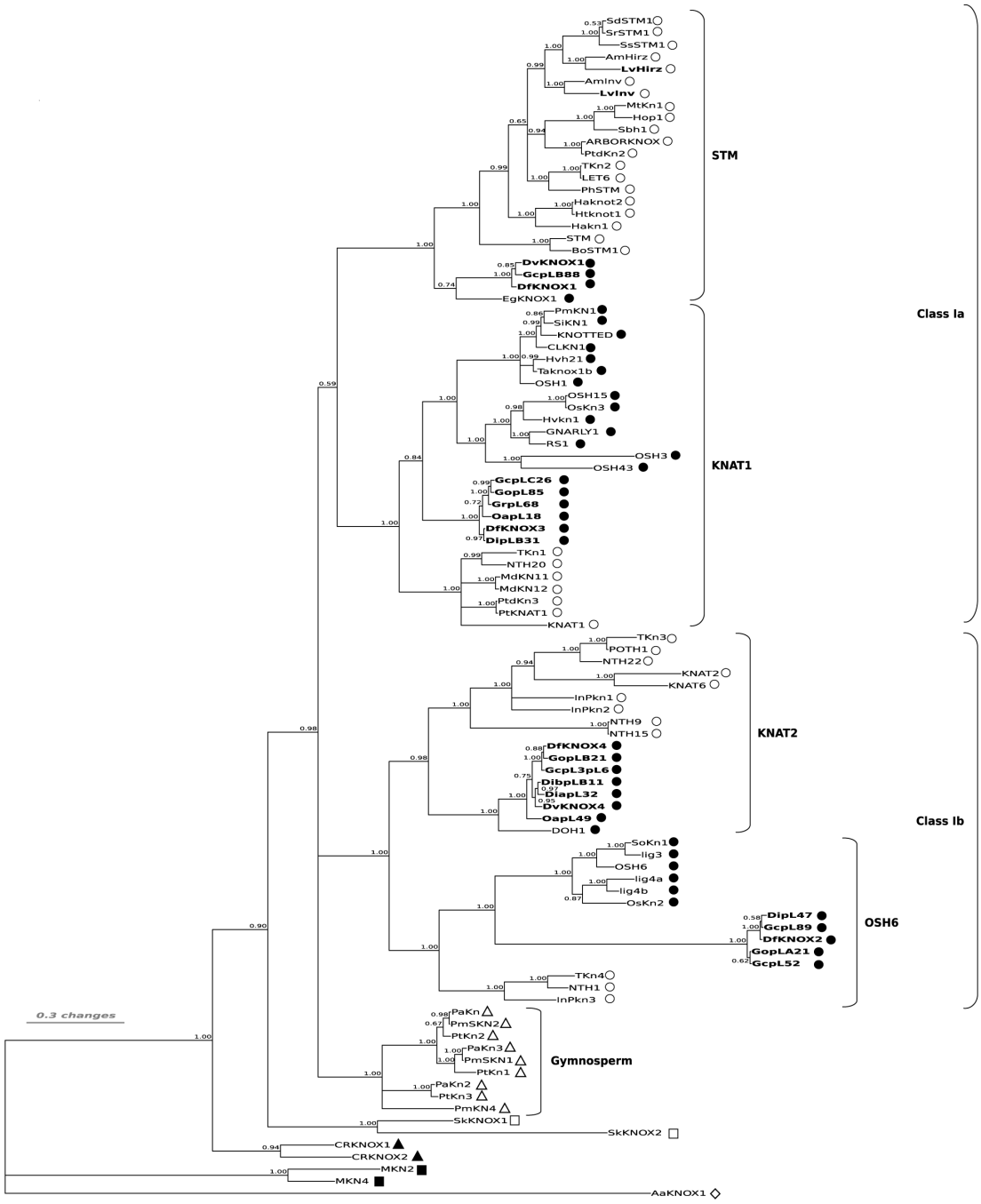


Figure 5.3: 50% majority rule consensus Bayesian phylogram based on the conserved HD region of 98 class 1 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

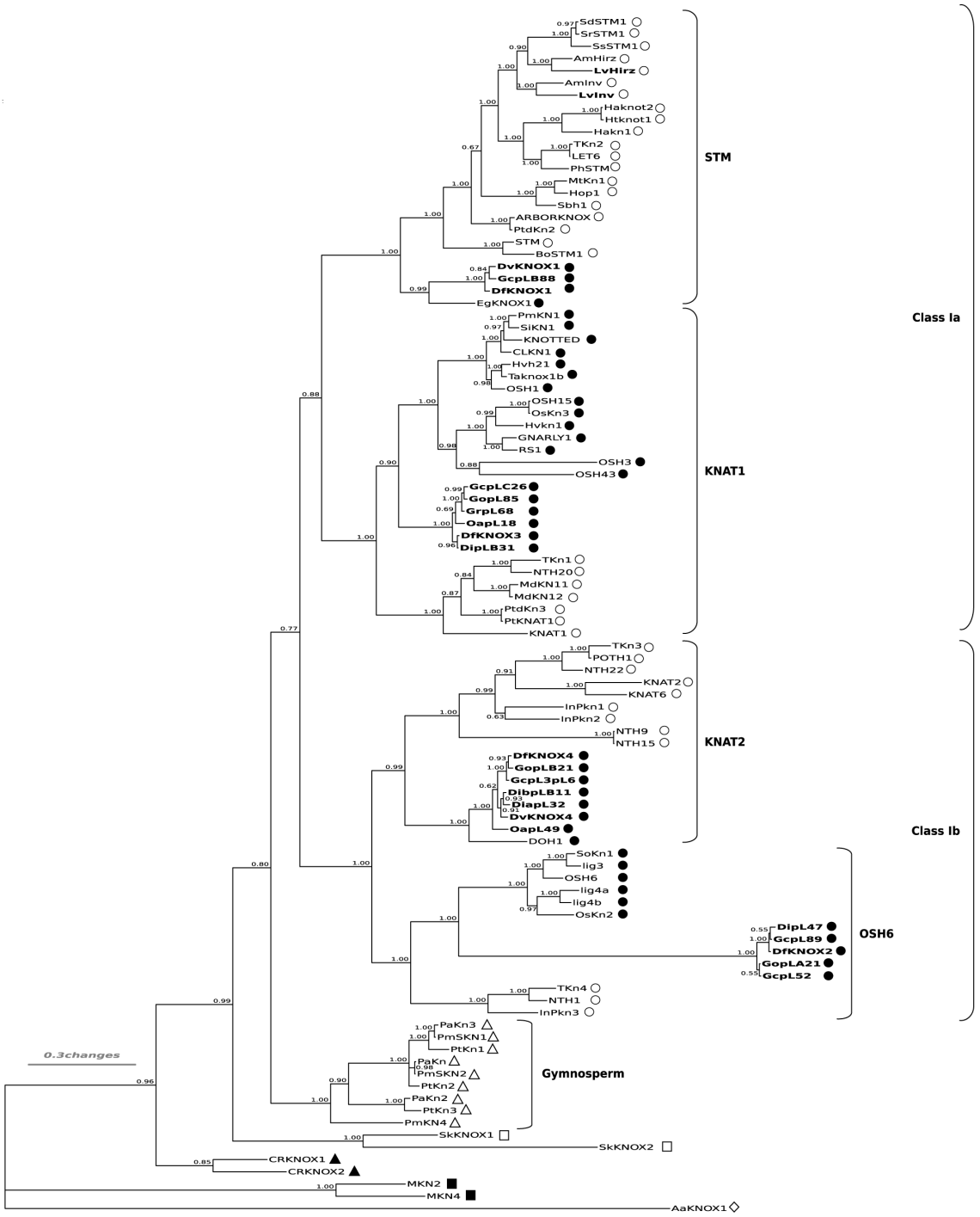


Figure 5.4: Bayesian phylogram based on the conserved MD and HD region of 98 class 1 *KNOX* gene DNA sequences.

Newly identified *KNOX* gene sequences are highlighted in bold. Symbols indicate plant groups; eudicots, open circles; monocots, closed circles; gymnosperms, open triangles; pteridophytes, closed triangles; lycophytes, open squares; bryophytes, closed squares; green algae, open rhombus (used as out-group).

5.2.3 Summary of combined phylogenetic results

Whilst there are numerous minor differences in topology and statistical support, analysis of the four *KNOX* gene trees generated by Bayesian Inference demonstrates several key points. Class 1 and class 2 *KNOX* genes are distinct, forming two well supported superclades (Figures 5.1 , 5.2). The angiosperm class 1 *KNOX* genes form a well supported monophyletic superclade distinct from gymnosperm, lycophyte, pteridophyte and bryophyte *KNOX* gene sequences. However, lycophyte and bryophyte class 2 *KNOX* gene representatives are not distinct from angiosperm class 2 *KNOX* genes.

Class 1 *KNOX* genes from angiosperms are divided into two distinct subclades, class 1a and class 1b (Figure 5.4). Class 1a *KNOX* genes consist of two distinct groups containing sequences with high similarity to *STM* and *KNAT1* from *Arabidopsis*. Class 1b *KNOX* genes also consist of two distinct groups with high similarity to *KNAT2* from *Arabidopsis* and *OSH6* from rice. Class 1 *KNOX* gene representatives from both monocots and eudicots are present in each of these four major angiosperm class 1 *KNOX* gene groups. However, although monocot and eudicot sequences are represented in each of the four groupings they form distinct clades within the larger class 1 *KNOX* gene groupings (Figures 5.1, 5.2, 5.3, 5.4).

5.3 Discussion

In this chapter the large scale patterns of evolution in the *KNOX* gene family were assessed using the phylogenetic trees originally developed to assess the homology of novel *KNOX* genes isolated in this work (see Chapter 4).

5.3.1 An improved *KNOX* gene phylogeny

The addition of 31 novel *KNOX* genes from nine different species (Table 4.2), including a large number of novel sequences from a vastly under represented group of plants, the orchids, adds considerable taxonomic breadth to current phylogenies of *KNOX* genes that largely represent monocot *KNOX*-related accessions from grass species (Poaceae). The addition of so many new sequences has vastly improved the confidence in current accepted *KNOX* gene phylogenies in terms of both statistical support and sensible changes in tree topology. Previously published tree topologies have been subject to some confusion, particularly with regard to the position of *KNOX*-related sequences from monocot accessions (Harrison *et al.*, 2005a,b; Jouannic *et al.*, 2007; Mukherjee *et al.*, 2009; Reiser *et al.*, 2000; Sano *et al.*, 2005).

The methods utilised in this analysis have also contributed to improvements in tree

topology and statistical support. Past phylogenetic analyses have favoured the popular Neighbour-Joining (NJ) and/or Maximum Parsimony (MP) methods of phylogeny reconstruction (Bharathan *et al.*, 1997, 1999; Bürglin, 1997; Champagne & Ashton, 2001; Chan *et al.*, 1998; Guillet-Claude *et al.*, 2004; Hake *et al.*, 2004; Harrison *et al.*, 2005a,b; Kerstetter *et al.*, 1994; Magnani & Hake, 2008; Reiser *et al.*, 2000; Vollbrecht *et al.*, 1991). Of these, few authors have used phylogenetically more informative nucleotide data (Champagne & Ashton, 2001; Harrison *et al.*, 2005a,b) for phylogeny reconstruction; fewer still have employed the use of modern phylogenetic techniques such as Maximum Likelihood (ML) in combination with DNA data (Jouannic *et al.*, 2007; Mukherjee *et al.*, 2009; Sano *et al.*, 2005). The phylogenetic re-constructions presented in this work have benefited from methodological improvements using BI in combination with nucleotide data. This is largely regarded as the most effective means to assess phylogenetic relationships. Combined with the increased representation of monocot *KNOX* gene sequences, the *KNOX* gene tree topologies presented in this work represent the most robust exploration of *KNOX* gene phylogeny to date.

Modern deep sequencing and genomic resources will soon eclipse the analysis presented in this work and help to plug the holes in our understanding of *KNOX* gene evolution. As of February 2010, there are 18 publically available green plant (Viridiplantae) genomes that can be effectively mined for sequence information, see www.phytozome.net. However, these genomes are model organisms and food crops, including Arabidopsis, Papaya, Soybean, Cucumber, Rice and Maize (Table 5.1). The availability of these resources will inevitably improve our understanding of *KNOX* gene evolution by providing critical non-coding sequence information, which could greatly improve the concordance between phylogeny and *KNOX* gene function, for example by providing regulatory sequence information including promoters and introns. However, current genomic resources for monocot taxa are still only represented by grasses. As the cost of deep sequencing becomes more affordable and the field of comparative genomics continues to expand I am confident that undersampling of critical angiosperm lineages, such as monocots, will improve.

Table 5.1: Summary of the 18 currently available Viridiplantae genome sequences (www.phytozome.net)

Organism	*Common name*	*Source*
Arabidopsis thaliana	Mouse-ear cress	TAIR release 9 acquired from TAIR
Arabidopsis lyrata	Lyre-leaved rock cress	JGI release 1.0
Carica papaya	Papaya	ASGPB release of 2007
Populus trichocarpa	Poplar	JGI v2.0 annotation of the v2 assembly
Medicago truncatula	Barrel medic	Release Mt3.0 from the Medicago Genome Sequence Consortium
Glycine max	Soybean	JGI Glyma1.0 annotation of the chromosome-based Glyma1 assembly
Ricinus communis	Castor bean	TIGR release 0.1
Manihot esculenta	Cassava	JGI/Roche v1.1 assembly and annotation
Cucumis sativus	Cucumber	Roche 454-XLR assembly and JGI v1 annotation
Vitis vinifera	Grape	Sept 2007 annotation from Genoscope
Sorghum bicolor	Sweet Sorghum	Sbi1.4 models from MIPS/PASA on v1.0 assembly
Zea mays	Maize	Protein coding models from Maizesequence.org release 4a.53
Oryza sativa	Rice	MSU Release 6.0 of the Rice Genome Annotation
Brachypodium distachyon	Purple false brome	JGI v1.0 8x assembly of strain Bd21 with JGI/MIPS PASA annotation
Mimulus guttatus	Monkey flower	JGI 7x assembly of strain IM62, annotation v1.0
Selaginella moellendorffii	Spikemoss	JGI v1.0 assembly and annotation
Phycomitrella patens	Moss	JGI v1.1 assembly and annotation
Chlamydomonas reinhardtii	Green algae	Augustus u9 annotation of JGI v4 assembly

5.3.2 Gymnosperm and pteridophyte *KNOX* genes are distinct from those of angiosperms

Notable topological improvements over previous phylogenies include the position of pteridophyte and gymnosperm *KNOX* sequences (Figure 5.5, and see sections 5.2.1, 5.2.2). The phylogenetic placement of pteridophyte and gymnosperm *KNOX* sequences are considerably different in this current phylogenetic analysis when compared to previously published *KNOX* gene tree topologies. The phylogenetic analyses of Reiser *et al.* (2000), Sano *et al.* (2005) and Jouannic *et al.* (2007) all recover gymnosperm and pteridophyte *KNOX* genes as distinct clades within the larger angiosperm class 1 *KNOX* gene superclade. The position of gymnosperm and pteridophyte *KNOX* genes amongst angiosperm class 1 *KNOX* gene sequences appears to be inconsistent with the almost basal position of bryophyte and lycophyte *KNOX* sequences in these analyses. The location of pteridophyte and gymnosperm *KNOX* sequences within the angiosperm class 1 *KNOX* clade therefore seems unlikely.

The tree topologies presented in this work show that bryophyte, pteridophyte, gymnosperm and lycophyte sequences form a group distinct from angiosperm class 1 *KNOX* genes in an order that recapitulates that of seed plant evolution. This pattern is well supported in all four trees generated in this study. The variable position of lycophyte *KNOX* gene sequences in class 1 versus total (class 1 and class 2) *KNOX* gene data sets, suggests that their position may be highly sensitive to the analysis methods implemented. This may also be true of pteridophyte and gymnosperm *KNOX* sequences and may explain the extraneous position recovered in previous phylogenies, see Sano *et al.* (2005) and Jouannic *et al.* (2007). Whilst it is possible that this pattern may be artifactual, the basal

location of non-angiospermous class 1 *KNOX* genes is more consistent with the current hypotheses on the pattern of *KNOX* gene evolution in seed plants. Previous phylogenies that place pteridophyte and gymnosperm *KNOX* genes within the larger angiosperm class 1 grouping most likely represent methodological artefacts.

5.3.3 Multiple gene duplication events characterise *KNOX* gene evolution

The phylogenetic trees presented in this work support those of Jouannic *et al.* (2007); Mukherjee *et al.* (2009); Reiser *et al.* (2000); Sano *et al.* (2005), strongly suggesting that *KNOX* gene duplication events are both ancient and modern and appear to have been a pervasive feature of evolution in the *KNOX* gene family (Figure 5.5), particularly among angiosperms. Such pervasive gene duplications may be related to the extreme morphological diversity of flowering plants.

5.3.3.1 An ancient evolutionary gene duplication in spermatophytes

The broad topology of spermatophyte *KNOX* genes is strongly supported by a large number of phylogenetic analyses of *KNOX* gene accessions (including this one) constructed using a range of methods including NJ, MP and ML (Bharathan *et al.*, 1999; Champagne & Ashton, 2001; Jouannic *et al.*, 2007; Magnani & Hake, 2008; Mukherjee *et al.*, 2009; Reiser *et al.*, 2000; Sano *et al.*, 2005; Serikawa & Mandoli, 1999). The phylogenetic reconstructions generated in this work strongly support the subdivision of class 1 and class 2 *KNOX* genes, thereby providing support for the theory of an ancient gene duplication event that may have occurred between 400 and 500 MYA (Figure 5.5, purple arrow), possibly before the evolution of land plants and concomitant with the evolution of complex multicellular plant bodies (Mukherjee *et al.*, 2009; Sano *et al.*, 2005).

5.3.3.2 Additional gene duplications/losses have occurred in a lineage specific manner

Given the extreme age of the *KNOX* gene family it is not surprising that *KNOX* gene phylogenies have highlighted multiple duplication events that have occurred subsequent to the ancient evolutionary duplication event that divided the *KNOX* gene family into the two fundamental *KNOX* gene classes. Phylogenetic reconstructions presented in this work strongly support the division of angiosperm class 1 *KNOX* genes into two larger monophyletic clades, which Jouannic *et al.* (2007) described as classes 1a and 1b. Both classes 1a and 1b *KNOX* genes are further subdivided into two monophyletic subgroups

(*STM, KNAT1* and *KNAT2, OSH6*) as previously reported in earlier phylogenetic reconstructions generated using different methods (Bharathan *et al.*, 1999; Jouannic *et al.*, 2007; Reiser *et al.*, 2000).

In each of the four phylogenetic trees reconstructed in this work, both monocot and eudicot sequences are represented in each of the class 1a and 1b clades as proposed by the tree topologies of Sano *et al.* (2005) and Jouannic *et al.* (2007). This contrasts directly with the tree topologies recovered by Harrison *et al.* (2005a,b) in which class 1 *KNOX* sequences from monocots were found to be monophyletic. The phylogenetic data presented here, support the theory that there have been at least two large-scale gene duplication events among angiosperm *KNOX* genes that occurred before the divergence of eudicot and monocot lineages (Figure 5.5, red arrows). Subsequent to the evolution of class 1 *KNOX* genes an ancestral class 1 *KNOX* gene has been duplicated in angiosperms leading to the divergence of the class 1a and 1b groups, this duplication event has subsequently been followed by the duplication of an ancestral class 1a gene to generate the *STM* and *KNAT1* groups and, an ancestral class 1b gene duplication to form the *KNAT2* and *OSH6* groups (Bharathan *et al.*, 1999; Jouannic *et al.*, 2007; Reiser *et al.*, 2000). This hypothesis was weakly supported in the phylogenetic analyses of Jouannic *et al.* (2007) and Hirayama *et al.* (2007), both of which include only a single monocot representative in the *STM* group (*EgKNOX1* and *RaSTM* respectively), and just a few monocot representatives in the *KNAT2* group. The addition of 3 novel monocot genes to the *STM* group and seven to the *KNAT2* group in this phylogenetic analysis, adds substantial support to these proposed duplication events.

Closer examination of the four principal angiosperm class 1 *KNOX* gene groups suggest that additional, lineage specific, duplications, or gene duplicate losses, have occurred in both monocots and eudicots (Figure 5.5, black and grey arrows respectively). Notably, monocot sequences have been regarded as absent from the *STM* group by some authors, e.g. Mukherjee *et al.* (2009), and are shown to be uncommon in the *STM* group in the trees presented in this study. By contrast monocot *KNOX* sequences are highly represented amongst the *KNAT1* group. Phylogenetic analyses presented in this work clearly support those of Jouannic *et al.* (2007) and Hirayama *et al.* (2007), indicating that rather than being an exclusively eudicot *KNOX* clade, monocot sequences appear to be less common in the *STM* group compared to *KNAT1*. Whilst it is true that the breadth of taxon sampling among monocot species is relatively poor and has focused largely on grasses (Poales), perhaps explaining the absence of *STM*-like monocot *KNOX* genes even in recent phylogenetic analyses (Mukherjee *et al.*, 2009), it is possible that the relative scarcity of monocot sequences in the *STM* group in particular, may be due to significant loss of *STM*-like paralogs in monocots (Tioni *et al.*, 2003). By contrast, many monocot

sequences are represented in the *KNAT1* group with relatively fewer eudicot representatives. Such observations may be biased by insufficient taxon sampling, or exaggerated by the polyploid genomes of the monocot species sampled, however these observations indicate that subsequent to initial common duplication events, monocot and eudicot *KNOX* gene lineages have evolved along distinct trajectories in a lineage specific manner. In some cases, such lineage specific duplications may have resulted in subfunctionalization and neofunctionalization, or changes in their spatio-temporal expression patterns and could be associated with morphological diversification of these groups. Functional diversification of *KNOX* gene duplicates has also been suggested for other seed plant groups e.g. Gymnosperms (Guillet-Claude *et al.*, 2004).

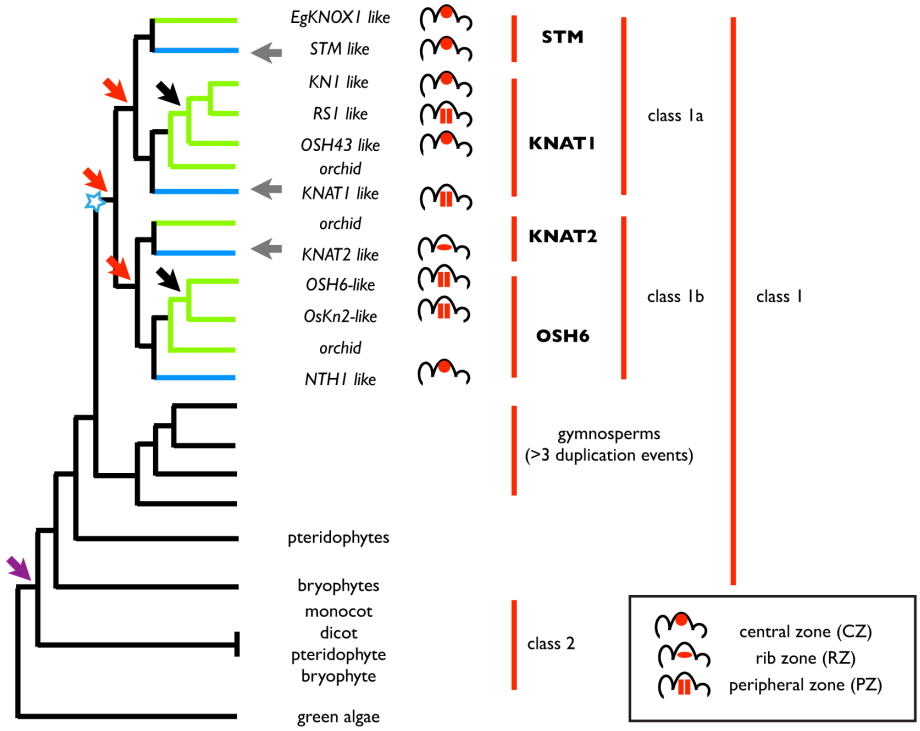


Figure 5.5: Schematic summarising the pervasive gene duplication events among angiosperm class 1 *KNOX* genes and stochastic distribution of SAM specific expression patterns

Schematic representation of the phylogenetic relationships between major *KNOX* gene representatives established in this work. Named *KNOX* genes represent lineages of closely related *KNOX* genes, not single genes themselves. Branches coloured in blue represent eudicots, while green branches indicate monocot lineages. Red braces indicate the class 1 *KNOX* gene subclades as defined by Bharathan *et al.* (1999); Jouannic *et al.* (2007); Reiser *et al.* (2000). An ancient gene duplication event, 400-500MYA, generated class 1 and class 2 *KNOX* genes (purple arrow). Among the monophyletic angiosperm class 1 *KNOX* genes (blue star), there have been at least three gene duplication events prior to the monocot-eudicot split (red arrows). In specific monocot (black arrows) and eudicot (grey arrows) lineages, evidence of subsequent lineage specific gene duplication events and/or losses, are commonplace. Multiple gene duplication events have also been recorded for gymnosperms (Guillet-Claude *et al.*, 2004). Patterns of *KNOX* gene expression in the SAM are presented for each lineage, revealing that even for the SAM, there is no overall correspondence between phylogenetic relationship and expression patterns for class I KNOX proteins, i.e. particular expression patterns are not constrained to certain gene lineages.

5.3.3.3 *KNOX* gene function is determined stochastically in different plant lineages

KNOX gene duplication events appear to be associated with stochastic retention of ancestral gene function and neo-functalisation, resulting in phylogenetic clustering of genes that may not be functionally equivalent (Figure 5.5), making functional assessment via sequence homology highly unreliable (see section 4.3.4). This suggests that the mechanisms that control the expression and subsequent neo-functionalization of gene paralogs evolve independently, supporting similar assertions by Jouannic *et al.* (2007).

A good example of stochastic retention of ancestral gene function is provided by the *Arabidopsis* gene *STM* and its functional homolog *KN1* from maize. Both *STM* and *KN1* are implicated in the maintenance of the SAM and have specific patterns of expression that are functionally equivalent (Jackson *et al.*, 1994; Long *et al.*, 1996). Based on similarity of protein sequence and gene expression patterns a number of authors have suggested that *STM* and *KN1* are orthologous genes. However, phylogenetic analyses (including this one) resolve *STM* and *KN1* into separate clades (Bharathan *et al.*, 1999; Jouannic *et al.*, 2007; Reiser *et al.*, 2000; Sano *et al.*, 2005), suggesting that they are not simply genetic orthologs but represent paralogous lineages derived from an earlier gene duplication event. This is also true of rice, in which the *OSH1* gene is orthologous to *KN1* and functions to maintain the SAM (Matsuoka *et al.*, 1993). However, in the orchid *Dendrobium*, the *KNAT2*-like gene (*DOH1*) seems to have the ancestral function equivalent to *STM* (Yu *et al.*, 2000). Stochastic retention of ancestral function following gene duplication is not uncommon. For example, discordance between patterns of gene duplication and functional evolution have also been demonstrated between the C-function genes *AGAMOUS* (*AG, Arabidopsis*) and *PLENA* (*PLE, Antirrhinum*) (Kramer *et al.*, 2004). Although both *AG* and *PLE* play functionally homologous roles in determining stamen and carpel identity, recent phylogenetic analyses resolve *AG* and *PLE* into separate paralogous lineages (Davies *et al.*, 1999; Kramer *et al.*, 2004; Svensson & Engstrom, 2002).

Though incomplete sampling among monocots is an obvious consideration, few *STM* orthologs have been identified from monocots and likewise in eudicots, few orthologs of *KN1* have been recovered (Figure 5.5). The absence of *STM* orthologs, at least among Poales, suggests that the retention of ancestral gene function by *KN1*, rather than an *STM* ortholog in maize, may be the result of loss of *STM* orthologs in monocots. In the absence of more complete sampling of monocot *KNOX* genes it is difficult to determine whether the use of *KN1*, rather than an *STM* ortholog, is a general characteristic of monocots as a whole, or is in fact peculiar to grasses. The recent discovery of two monocot orthologs of *Arabidopsis STM* in *Ruscus aculeatus* (Asparagales) and *Elaeis guineensis* (Arecales), which use an *STM* ortholog rather than *KN1* for ancestral SAM function (Hirayama

et al., 2007; Jouannic *et al.*, 2007), in combination with the identification here of two *STM*-like *KNOX* genes from Orchids (Asparagales), *DfKN1/DvKN1*, strongly suggest that the apparent loss of *STM* orthologs in grasses and subsequent function of *KN1*, may be specific to grasses. Taking into account phylogeny, Asparagales and Arecales diverged earlier than Poales (Chase, 2004; Stevens, 2001). This evidence supports the idea that *STM* homologs were lost some time during diversification of Poales and their function assumed by the *KN1* homolog.

In addition to stochastic retention of ancestral *KNOX* gene function, diversification of *KNOX* gene function is also stochastic. In *Arabidopsis STM* is expressed in the SAM where it functions in meristem maintenance, subsequently being down-regulated in developing leaf primordia (Jackson *et al.*, 1994; Smith *et al.*, 1992). However, in closely related *Cardamine hirsuta*, which has dissected leaves, *STM* expression in leaf primordia is sufficient to initiate leaflet formation (Hay & Tsiantis, 2006). Such a role for *KNOX* genes in complex leaf morphogenesis is also implicated in the evolution of compound leaves in tomato, in which the tomato *STM* ortholog *TKN2* is also expressed in developing leaf primordia (Chen *et al.*, 1997; Janssen *et al.*, 1998a; Parnis *et al.*, 1997). By contrast, other orthologs of *Arabidopsis STM* have been implicated in quite different developmental processes. The most pertinent example would be that of *STM* orthologs *AmHIRZ* and *AmINA* in the development of floral nectar spurs in Antirrhineae (Golz *et al.* (2002), this study). At the protein level *STM*, *TKN2*, *AmHIRZ* and *AmINA* are very similar with no obvious features that would seem to favour their involvement in any one process over another. The involvement of similar orthologs in different developmental processes indicates firstly that the function of *KNOX* genes may be common to both complex leaves and nectar spurs (see Chapter 6), but also that diversification of *KNOX* gene function is, to some degree, unpredictable.

Despite distinct roles in apparently quite different developmental processes, *KNOX* genes are thought to carry out a common core role, namely the maintenance of developmental potential by delaying the onset of cellular differentiation (see Chapter 6). Given such a common developmental role in the SAM and other morphogenetic processes, functional redundancy observed between paralogous class 1 *KNOX* genes that have resulted from gene duplication events, e.g. *STM* and *KNAT1* (Byrne *et al.*, 2002), may explain the stochastic retention and diversification of *KNOX* gene function apparent in the *KNOX* gene family. Functional redundancy may reduce phylogenetic constraint on which gene duplicate may be selected for retention of ancestral gene function or neo-functionalization, thereby explaining the stochastic evolutionary patterns of ancestral and novel functions apparent in the *KNOX* gene family.

5.3.4 Conclusions

The addition of 31 new *KNOX* genes in combination with the use of modern phylogenetic techniques, has resulted in highly supported tree topologies that complement more recent phylogenetic analyses of the *KNOX* gene family and previously published theories of *KNOX* gene evolution. The tree topologies presented in this chapter have resolved the extraneous position of gymnosperm and pteridophyte *KNOX* gene sequences recovered in previous analyses and represent the most robust exploration of *KNOX* gene phylogeny to date.

KNOX genes are an evolutionarily ancient gene family, and, consistent with this age the *KNOX* gene family has undergone a multitude of gene duplication events that are both evolutionarily ancient, generating a common complement of *KNOX* genes at least among angiosperms, in combination with more recent lineage specific gene duplication events and/or gene losses associated with stochastic retention of ancestral gene function and the acquisition of new roles. Functional redundancy between gene duplicates is a likely explaination for the stochastic evolutionary pattern of ancestral versus novel gene function evident in the *KNOX* gene family when reconstructed using protein coding DNA sequence information. However, incorporating regulatory sequence information for *KNOX* genes may reveal a pattern that is presently masked by the limited DNA sequence resources currently available.

The tree topologies presented in this chapter mark a considerable improvement over previous published phylogenetic analyses. However many taxa, particularly among monocots, remain severely under-represented. Many of the monocot sequences represented in past and present *KNOX* gene phylogenies are almost exclusively derived from grasses that have undergone vast amounts of gene duplication and neo/subfunctionalization by polyploidy. It appears highly likely that, rather than reflecting general trends in monocot *KNOX* gene evolution, current published phylogenies show trends that are peculiar to Poales. As such, critical information regarding the pattern of *KNOX* gene evolution in monocot lineages is not currently available. In the fast approaching era of affordable plant genomics, limitations such as insufficent sequence and taxon representation are likely to be overcome. Such advances are almost here and will greatly advance our understanding of *KNOX* gene evolution.

Chapter 6

Characterisation of *LvHIRZ* and *LvINA* expression patterns

6.1 Introduction

The traditional view that class 1 *KNOX* genes are involved exclusively in maintenance of the SAM and are actively excluded from lateral organ development (Endrizzi *et al.*, 1996; Jackson *et al.*, 1994; Kerstetter *et al.*, 1997; Long *et al.*, 1996; Smith *et al.*, 1992) has been challenged in recent years. This challenge is the result of a dramatic increase in the breadth of taxa from which *KNOX* genes have been isolated and their expression patterns studied. Additional roles for *KNOX* genes have been suggested in the development of carpels (Scofield *et al.*, 2007, 2008), lateral roots (Dean *et al.*, 2004), tubers (Chen *et al.*, 2003; Rosin *et al.*, 2003) and in the development of novel axes of leaf and petal growth (Golz *et al.*, 2002; Hay & Tsiantis, 2006).

Whilst it still holds that plants with relatively simple lateral organ morphologies exclude *KNOX* gene expression from lateral organ primordia, e.g. in the morphogenesis of the simple leaves of *Arabidopsis*, a number of authors now recognise the importance of *KNOX* genes in the development of complex lateral organs (Golz *et al.*, 2002; Hareven *et al.*, 1996; Hay & Tsiantis, 2006; Shani *et al.*, 2009). Despite initial down-regulation of *KNOX* expression, a broad range of angiosperm taxa with dissected leaves or complex petal morphologies re-activate *KNOX* gene expression in late stages of lateral organ morphogenesis (reviewed by Hay *et al.* (2009); Tsiantis & Hay (2003)). The requirement for indeterminacy factors in determinate organ morphogenesis is an apparent contradiction. However, the development of complex lateral organs appears to require a sustained indeterminate environment in late stages of morphogenesis (Hareven *et al.*, 1996; Hay & Tsiantis, 2006; Shani *et al.*, 2009).

Additional roles for *KNOX* genes have also been proposed in the initiation of novel

axes of petal growth as a requirement for the development of floral nectar spurs in close relatives of *Antirrhinum majus* (Golz *et al.*, 2002). Despite lacking nectar spurs of its own, snapdragon mutants (*Hirz-d153* and *Ina-d1*) ectopically expressing class 1 *KNOX* genes (*AmHIRZ* and *AmINA*, respectively) outside of the SAM in petals and leaves, develop ectopic petal tubes that resemble the nectar spurs of closely related taxa such as *Linaria vulgaris* (common toadflax). The nectar spurs of closely related taxa in Antirrhineae are formed from growth in a novel axis after initiation of the normal petal tube, have parallel venation, show mirror image symmetry of cell types and orientation about their long axes and occur only ventrally (Sutton, 1988). Like the natural spurs of other members of Antirrhineae, the ectopic petal tubes of snapdragon mutants share these properties, strongly suggesting that floral nectar spurs in Antirrhineae may have evolved as additional petal tubes (Golz *et al.*, 2002).

In wild-type *A. majus*, Waites *et al.* (1998) and Golz *et al.* (2002) noted expression of *AmHIRZ* and *AmINA* in the SAM and provascular cells of internodes. However, transcripts of both genes were absent from lateral organ primordia, consistent with typical *KNOX* gene function. Both *AmHIRZ* and *AmINA* have similar patterns of expression in the SAM with one exception, *AmHIRZ* accumulates in the L1 layer of the SAM, from which *AmINA* is absent (Golz *et al.*, 2002). In floral meristems wild-type *AmHIRZ* and *AmINA* expression is equivalent to that in the SAM and transcripts are absent from floral organ primordia. However, expression persisted in the receptacle, proximal to developing organs, with particularly strong expression in the cells that mark the boundary between the stamen and carpel. The wild-type expression of *AmHIRZ* and *AmINA* in the receptacle prompted Golz *et al.* (2002) to suggest that these genes *organise* proximo-distal patterning of floral organs. The ectopic expression of these genes in snapdragon mutants is thought to generate an additional ectopic organiser of petal growth, resulting in the repeated petal tube on the ventral part of the corolla in *Hirz-d153* and *Ina-d1* snapdragon mutants.

The identification of these mutants strongly suggests that elaboration of *KNOX* gene expression patterns have contributed to the evolution of nectar spurs in Antirrhineae by ectopic expression of a proximo-distal organiser of petal development (Golz *et al.*, 2002). Despite the feasibility of this model, it is based only on observations of snapdragon mutants and not natural spur-bearing genera.

6.1.1 Research aims and objectives

A number of authors have suggested that the role of *KNOX* genes in nectar spur development should be assessed in a close relative of *A. majus* with natural spurs, such as *L. vulgaris* (Damerval & Nadot, 2007; Galego & Almeida, 2007; Whitney & Glover, 2007).

In accordance with this suggestion the orthologous genes of *AmHIRZ* and *AmINA* were identified from *L. vulgaris* (Chapter 4). Bioinformatic and phylogenetic assessment of these genes strongly supports their orthology to those from *A. majus*. The closely related nature of *A. majus* and *L. vulgaris* (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000) suggests that these genes are likely to play a fundamental role in maintenance of the SAM, as is the case for wild-type *AmHIRZ* and *AmINA* (Golz *et al.*, 2002). However, in the absence of expression data from close relatives of *L. vulgaris* with naturally occurring nectar spurs, phylogenetic data alone are insufficient to predict whether the orthologous genes from *L. vulgaris* are likely to have similarly broad patterns of *KNOX* expression to *AmHIRZ* and *AmINA* in the *Hirz-d153* and *Ina-d1* snapdragon mutants.

The principal aim of this part of the project was to explore the pattern of *LvHIRZ* and *LvINA* transcript accumulation in meristematic, vegetative and floral tissues at appropriate stages of development, in order to establish a wild-type role for class 1 *KNOX* genes in floral nectar spur development in Antirrhineae. A combination of techniques was used to explore the expression patterns of *LvHIRZ* and *LvINA* in relation to the development of floral nectar spurs. Patterns of *LvHIRZ* and *LvINA* transcript accumulation in different tissues and floral organs were assessed using a combination of gene-specific reverse transcription PCR (RT-PCR) and quantitative real time PCR (QRT-PCR). In addition, *in situ* hybridisation (ISH) was attempted in order to precisely identify the spatio-temporal patterns of *LvHIRZ* and *LvINA* expression in particular parts of developing organs not amenable to further dissection and sampling using PCR methods. Unfortunately ISH was not successful in revealing patterns of *KNOX* gene expression and could not be optimised for *L. vulgaris* in the time available.

6.2 Results

6.2.1 Determining the pattern of *LvHIRZ* and *LvINA* transcript accumulation

Transcript accumulation of the two newly identified *KNOX* genes *LvHIRZ* and *LvINA* was determined in a number of different tissues and organs using a combination of gene-specific RT-PCR and QRT-PCR. Expression in each case was determined relative to the housekeeping gene *Tubulin alpha-5* (*LvTUA5*). For each gene, a minimum of three biological replicates was carried out for PCR based expression analyses (each consisting of three technical replicates) and identical stages of vegetative and floral development were sampled for both RT and QRT-PCR. Floral transcript accumulation was determined from flowers at the earliest stages of floral nectar spur development (Figure 6.1).

6.2.1.1 Gene-specific reverse transcription PCR (RT-PCR)

RT-PCR revealed that *LvHIRZ* and *LvINA* transcripts accumulate to readily detectable levels in a variety of floral and vegetative tissues. *LvHIRZ* transcripts were detected in developing floral buds with early initiating nectar spur primordia (Figure 6.1) but, unusually, were not present at detectable levels in the vegetative shoot apex containing the SAM. Despite high reproducibility, absence of *LvHIRZ* expression in the SAM is likely to be artifactual, as expression of *LvHIRZ* in the SAM is readily detectable by QRT-PCR (section 6.2.1.2).

Developing flowers at the earliest stages of nectar spur development were dissected into their component floral organs. These consisted of the calyx, reproductive organs (androecium plus gynoecium), the dorsal petal, lateral petal and ventral petal (which bears the nectar spur primordium; Figure 6.1A). Total RNA was extracted from each tissue type to determine precisely where *LvHIRZ* transcripts accumulate in early developing floral buds. From this analysis *LvHIRZ* transcripts were detected in the calyx, the combined reproductive units of the flower and in the dorsal and ventral petals of the fused corolla tube, but not in the lateral petals (Figure 6.1B). As expected, *LvHIRZ* does not appear to be expressed in leaves. DNA sequencing analysis confirmed specific amplification of *LvHIRZ* in all tissues and gave no indication that *LvHIRZ* transcripts differed in sequence in the broad range of tissues sampled, i.e. that alternatively spliced transcripts may be present in different tissues.

An identical analysis was conducted to assess transcript accumulation of *LvINA* using the same cDNA samples as those utilised in the *LvHIRZ* expression analysis described above (Figure 6.1). Transcript accumulation of *LvINA* is apparent in the vegetative shoot apex (SAM) and in developing floral buds. More specifically RT-PCR suggests that in the developing floral buds *LvINA* transcripts accumulate in the calyx, the combined reproductive units of the flower (androecium and gynoecium) and in the dorsal and ventral petals of the fused corolla tube. However, transcript accumulation is barely detectable in the lateral petals. As expected, *LvINA* is also not expressed in leaves. DNA sequencing analysis confirms specific amplification of *LvINA* in all tissues and does not indicate any difference in the sequence of *LvINA* transcripts between tissues sampled.

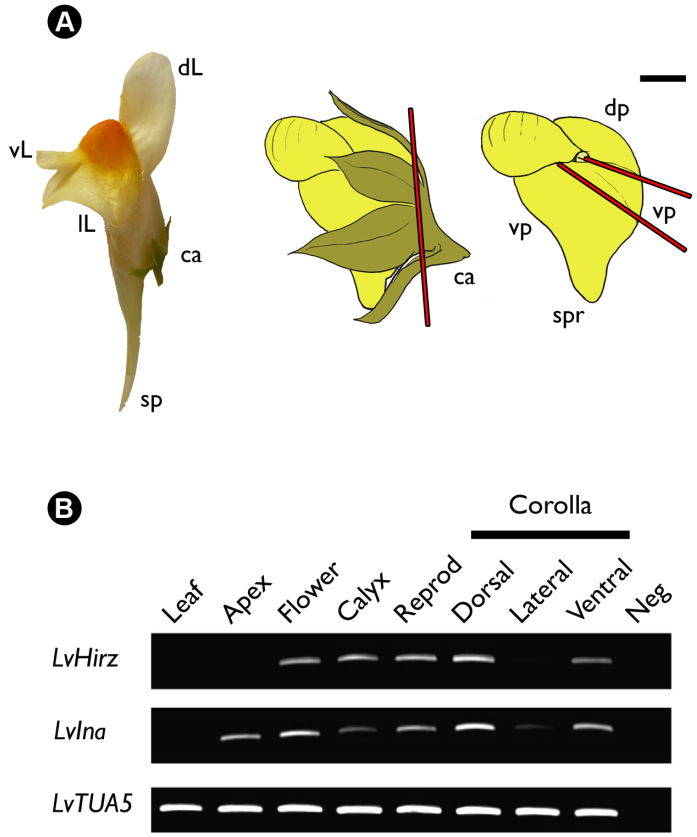


Figure 6.1: Gene-specific RT-PCR analysis of *LvHIRZ* and *LvINA* transcript accumulation in various tissues of the developing floral bud.
Schematic representation of *L. vulgaris* floral dissection and the tissues used in subsequent gene-specific RT-PCR analyses of *LvHIRZ* and *LvINA* relative to the housekeeping gene *LvTUA5*. A. Floral dissection plan of an *L. vulgaris* flower. Left, mature *L. vulgaris* flower; middle, line drawing of a developing *L. vulgaris* flower bud; right, line drawing of a developing *L. vulgaris* flower bud from which the calyx has been removed. dL, lL and vL denote dorsal, lateral and ventral petals/petal lobes, ca calyx, sp nectar spur, spr spur primordium, red lines indicate cuts for dissection, scale bar for line drawings = 1mm. Buds represented in line drawings are representative of the developmental stage sampled for gene expression analyses. B. Gene-specific RT-PCR of *LvHIRZ* (305bp) and *LvINA* (301bp) relative to *LvTUA5* (358bp).

6.2.1.2 Quantitative real time PCR (QRT-PCR)

Quantitative differences in gene expression may be an important predictor of biological activity. While gene-specific RT-PCR analyses indicate the presence or absence of gene expression, they provide no reliable indication of quantitative differences in transcript accumulation between tissues and/or between genes of interest.

In order to quantify *LvHIRZ* and *LvINA* transcript accumulation QRT-PCR was conducted using an identical experimental rationale to the gene-specific RT-PCR experiment described in section 6.2.1.1. Accumulation of *LvHIRZ* and *LvINA* transcripts was assessed in the leaf, apex (SAM), flower and floral organs relative to the housekeeping gene *LvTUA5*. In order to accurately compare *LvHIRZ* and *LvINA* transcript accumulation, differences in the reaction efficiencies of the QRT-PCR primer sets for each gene were efficiency corrected using the method outlined in section 2.6.3. The statistical significance of the resultant data was assessed using a students' t-test.

QRT-PCR of *LvHIRZ* and *LvINA* strongly supports the expression data generated by non-quantitative gene-specific RT-PCR (section 6.2.1.1). Neither *LvHIRZ* or *LvINA* are expressed in the leaf, as defined by consistent failure to amplify detectable product in independent QRT-PCR reactions. For each gene, expression in the vegetative shoot apex containing the SAM was used as a benchmark for assessing transcript accumulation in different floral tissues.

Compared to transcript levels in the apex (Figure 6.2), *LvHIRZ* expression is increased 10-fold in developing floral buds with early initiating nectar spur primordia. Dissected floral buds at an identical stage of development show that expression of *LvHIRZ* is 9-fold greater in the calyx than the SAM, 5-fold greater in the combined reproductive units of the flower (androecium plus gynoecium) and approximately 12-fold greater in the dorsal and ventral petals. However, expression of *LvHIRZ* remains equivalent to the apex in the lateral petals (Table 6.1, Figure 6.2). A students' t-test was performed on these data. Elevated *LvHIRZ* expression in the calyx, dorsal and ventral petals was found to be highly statistically significant ($p \leq 0.05$). Despite a 10-fold increase in expression of *LvHIRZ* in developing floral buds relative to the SAM, this increase was not found to be statistically significant. However, data for floral bud expression had the greatest variation and would become statistically significant if sampled further.

In contrast, *LvINA* is expressed at a relatively consistent level in the apex, developing floral bud and dissected floral tissues. There is a marked reduction in *LvINA* transcript accumulation in the calyx and statistically significantly reduced expression in the lateral petal ($p \leq 0.05$) in which *LvINA* expression is less than 20% that of the SAM (Table 6.1, Figure 6.2).

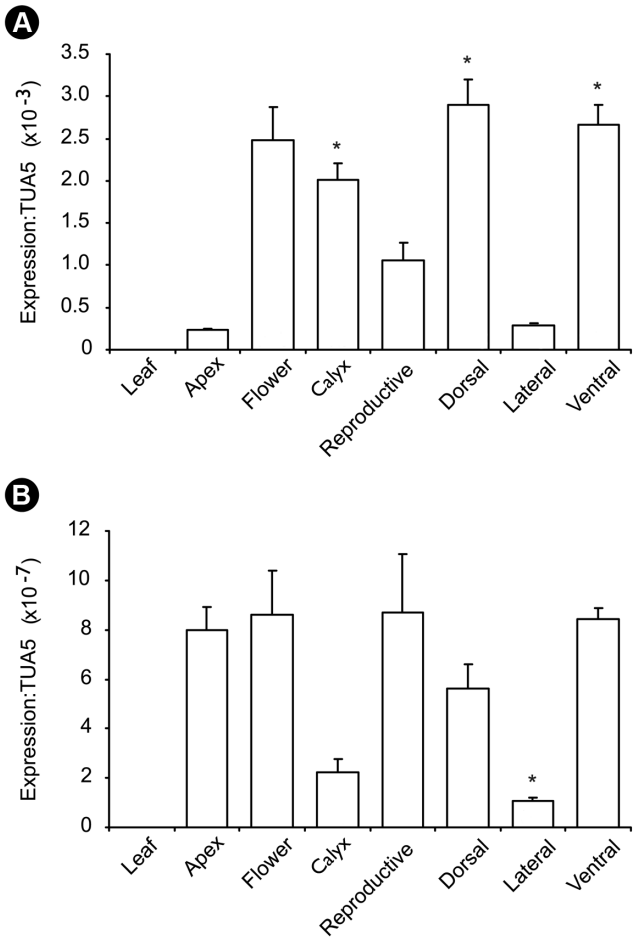


Figure 6.2: Quantitative RT-PCR analysis of *LvHIRZ* and *LvINA* transcript accumulation in various tissues of the developing floral bud

Histograms of transcript accumulation of *LvHIRZ* and *LvINA* from a variety of different tissues of the developing floral bud measured relative to the housekeeping gene *LvTUA5*. Each bar represents the geometric mean from three biological replicates, each consisting of a minimum of three technical replicates. Error bars were calculated from the log (base 2) of the data. Expression of *LvHIRZ* and *LvINA* in different tissues is directly comparable as the QRT-PCR data has been efficiency corrected. Statistical significance ($p \leq 0.05$) of expression relative to that in the SAM (apex) is indicated by an asterisk (*). A. Relative expression of *LvHIRZ*. B. Relative expression of *LvINA*.

Table 6.1: Summary of *LvHIRZ*/*LvINA* QRT-PCR

Students' t-test - p $\leq 0.05^{*}$, p $\leq 0.02^{**}$, p $\leq 0.01^{***}$

Fold-difference in expression relative to SAM			
Tissue	*LvHIRZ*	*LvINA*	*HIRZ vs INA*
Leaf	0	0	-
Apex	1	1	290^{***}
Flower	10.7	1.1	2880
Calyx	8.7^{*}	0.3	$8{,}990^{***}$
Reproductive	4.5	1.1	1,210
Dorsal	12.4^{*}	0.7	$5{,}130^{**}$
Lateral	1.25	0.1^{*}	$2{,}670^{***}$
Ventral	11.5^{*}	1.1	$3{,}160^{***}$

The expression of *LvINA* is considerably lower than *LvHIRZ* in all tissues. Expression of *LvHIRZ* is approximately 290-fold greater than *LvINA* in the SAM (p ≤ 0.01), and between 2,670 and 8,990-fold greater (p ≤ 0.01) in dissected floral organs (Table 6.1). These data strongly suggest that expression of *LvHIRZ* is significantly greater in floral tissues while *LvINA* shows a broad, but consistently low-level pattern of gene expression. Based on this evidence *LvHIRZ* is a likely candidate for a role in nectar spur development.

6.3 Discussion

Expression patterns of the *KNOX* genes *LvHIRZ* and *LvINA* isolated from *Linaria vulgaris* are discussed in relation to those of their orthologs *AmHIRZ* and *AmINA* from *Antirrhinum majus*, with particular emphasis on their potential role in the evolution and development of nectar spurs in Antirrhineae (cf. Golz *et al.* (2002)).

6.3.1 *LvHIRZ* and *LvINA* are likely to be involved in maintenance of the SAM

LvHIRZ and *LvINA* are a paralogous gene pair orthologous to the *STM*-like class 1 *KNOX* genes *AmHIRZ* and *AmINA* from *Antirrhinum majus* (Section 4.3.2). Golz *et al.* (2002) described typical class 1 *KNOX* gene expression patterns for *AmHIRZ* and *AmINA* in the outermost cell layer of the SAM (L1) and the internodes/pedicel region. In wild-type *A. majus* both genes are absent from leaves, organ initials and primordia. As such, wild-type *AmHIRZ* and *AmINA* are likely to function redundantly in maintaining the SAM (Golz *et al.*, 2002), as is typical for other class 1 *KNOX* genes (Endrizzi *et al.*, 1996; Kerstetter *et al.*, 1997; Long *et al.*, 1996; Scofield & Murray, 2006; Vollbrecht *et al.*, 2000).

In the absence of *in situ* hybridisation data it is difficult to determine the cell-specific expression patterns of *LvHIRZ* and *LvINA*, however, transcripts of both genes are readily detectable in the SAM and are absent from vegetative tissues such as leaves. Given the close phylogenetic relatedness of *L. vulgaris* and *A. majus* (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004), it is highly probable that wild-type transcripts of *LvHIRZ* and *LvINA* may have a functionally redundant role in maintaining the SAM, equivalent to that of *AmHIRZ* and *AmINA*. Such a role may be common for orthologs of these genes among other genera in the Antirrhineae.

6.3.2 Extra-meristematic patterns of *LvHIRZ* and *LvINA* gene expression resemble those of *AmHIRZ* and *AmINA* in snapdragon mutants

Despite a probable wild-type role in maintaining the SAM, wild-type transcripts of *LvHIRZ* and *LvINA*, identical in sequence to those expressed in the SAM (data not shown), have much broader patterns of expression than their wild-type orthologs from *A. majus*. Both *LvHIRZ* and *LvINA* are detectable in late developmental stages of floral organs such as the calyx, androecium, gynoecium and the dorsal and ventral petals of the corolla (Figures 6.1, 6.2).

While both *LvHIRZ* and *LvINA* are clearly expressed outside of the SAM, QRT-PCR shows that *LvHIRZ* is predominantly a floral-expressed class 1 *KNOX* gene with 10-fold higher expression in floral tissues when compared to the SAM. In particular, *LvHIRZ* transcripts accumulate predominantly in the dorsal and ventral parts of the corolla. By contrast, *LvINA* has much lower levels of expression in floral organs, equivalent to that in the SAM. While a wild-type role in SAM maintenance is highly likely for both genes, the broad and high-level expression of *LvHIRZ* suggests additional roles for this gene in one or more processes related to floral development. Although it appears likely that *LvHIRZ* and *LvINA* are functionally redundant in the SAM, the results presented in this work suggest that in *L. vulgaris* these two highly similar gene paralogs may have differential activity outside of the SAM. Whilst it is clear that differential expression may be one such mechanism by which these similar genes carry out differential functions, such functional differences might also relate to differential interactors and/or targets (Krizek & Fletcher, 2005; Sablowski, 2007). At present such interactors/targets are unknown.

The broad expression patterns of *LvHIRZ* and *LvINA* are reminiscent of those determined for *AmHIRZ* and *AmINA* in the snapdragon mutants *Hirz-d153* and *Ina-d1*, respectively (Golz *et al.*, 2002). In these mutant lines, ectopic expression of the class 1 *KNOX* genes *AmHIRZ* and *AmINA* outside of the SAM in petals and leaves result in a range of pleiotropic phenotypes including altered leaf shape, excessive trichome formation and, most notably, an additional ectopic petal tube on the ventral part of the corolla, morphologically similar to the spurs of close relatives of snapdragon. These data suggest that the nectar spurs of closely related Antirrhineae are ectopic petal tubes formed as a result of elaboration of *KNOX* gene expression (Golz *et al.*, 2002). Similarly broad patterns of expression for the paralogs *LvHIRZ* and *LvINA* in developing floral organs strongly support the hypothesis formulated by Golz *et al.* (2002).

Given the uncertain but closely related phylogenetic nature of *L. vulgaris* and other spur-bearing genera in Antirrhineae, such as *Kickxia* (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004), the broad patterns of floral *KNOX* gene expression in *L. vulgaris* indicate that such an expression pattern is likely to be common amongst other spur-bearing Antirrhineae and may be integral to the evolution and development of nectar spurs in this group.

6.3.3 Possible causes of extra-meristematic expression of *LvHIRZ* and *LvINA*: the importance of cis-regulatory regions in morphological evolution

Changes to cis-regulatory regions of genes are a common method by which genes can escape ancestral patterns of gene expression and become expressed in novel locations thereby potentiating the evolution of new morphologies. Regulatory sequences are a richer source of variation than coding sequences, as they are less constrained by the need to maintain the triplet code. Therefore, rapid changes in cis-regulatory sequences might be more important than coding, or trans-regulatory changes, for morphological evolution in plants (Cong *et al.*, 2002; Wang *et al.*, 1999).

Recent advances in the understanding of *KNOX* gene regulatory mechanisms have identified several cis-regulatory targets that may be responsible for changes in the expression patterns of *KNOX* genes. The *K-box* is one such cis-regulatory target, identified in the 5′ UTR of a number of *Arabidopsis STM* orthologs in simple and compound-leafed species across monocots and eudicots. Responsible for the persistent repression of *STM* transcription in *Arabidopsis* leaves following initial downregulation in the leaf primordium (Uchida *et al.*, 2007), the *k-box* appears to be involved in mediating a cis-regulatory pathway involving known negative regulators of *KNOX* genes such as *ASYMMETRIC LEAVES1/2* (*AS1/2*). There is also considerable evidence for epigenetic regulation of *KNOX* genes, e.g. the chromatin modification factors encoded by *SERRATE*, *PICKLE* (Eshed *et al.*, 1999; Ogas *et al.*, 1999; Prigge & Wagner, 2001) and polycomb group proteins *CURLY LEAF*, *SWINGER* and *FERTILISATION INDEPENDENT ENDOSPERM*, all of which have methyltransferase activity that can repress *KNOX* activity by trimethylation of histone H3 at lysine 27 (H3K27me3) and subsequent chromatin re-modelling (Katz *et al.*, 2004; Schubert *et al.*, 2006; Xu & Shen, 2008). Mutations that disrupt the *K-box*, or similar elements, may permit expression of *KNOX* genes in areas from which they are normally excluded either by negative regulators such as *AS1/2* or through down-regulation by chromatin modification.

Changes in promoter sequences are a further cis-regulatory change that is commonly responsible for elaboration of gene expression patterns. A particularly good example of this comes from an investigation of the role of *KNOX* genes in the evolution and development of dissected leaves in *Cardamine hirsuta*, a close relative of *Arabidopsis* with dissected leaf morphology (Hay & Tsiantis, 2006). In *C. hirsuta* orthologs of *Arabidopsis STM* and *BP* are expressed outside the SAM in developing leaf primordia where they are associated with the development of dissected leaves. Promoter swapping experiments between *Arabidopsis* and *C. hirsuta* orthologs of *STM* and *BP* showed that the endoge-

nous gene expression pattern of the promoter sequence was different in *Arabidopsis* and *C. hirsuta STM* and *BP* orthologs. The *Arabidopsis* promoter sequences drive expression of *STM* and *BP* only in the SAM, whereas the *C. hirsuta* promoter sequences were able to drive *STM* and *BP* expression in the developing leaf primordia of *Arabidopsis*, from which *KNOX* gene expression is normally excluded (Hay & Tsiantis, 2006). By contrast, the activity of trans-acting negative regulators of *KNOX* genes, e.g. *AS1*, is conserved between *Arabidopsis* and *C. hirsuta*.

In the case of ectopic *AmHIRZ* and *AmINA* expression in the *Hirz-d153* and *Ina-d1* mutants of snapdragon, transposon insertions (Tam1 in the first intron of *AmHIRZ*, Tam3 in the 5′ UTR of *AmINA*) were found to be the causative agents of ectopic expression (Golz *et al.*, 2002). Snapdragon is famous for the activity of transposable elements, such activity is responsible for generating the plethora of mutants that has made *Antirrhinum* such a powerful model system (Schwarz-Sommer *et al.*, 2003). However, despite recent evidence that similar transposable elements may exist in *L. vulgaris* (Galego & Almeida, 2007), it remains unclear how significant a role transposon insertions have played in the elaboration of *LvHIRZ* and *LvINA* expression outside the SAM. Sequence analysis of the coding and untranslated sequences of the *LvHIRZ* and *LvINA* cDNAs shows no indication of cis-regulatory disruptions discussed above. In particular, there is no evidence of Tam1 and Tam3 transposon insertions responsible for ectopic *AmHIRZ* and *AmINA* expression in the *Hirz-d153* and *Ina-d1* snapdragon mutants (Golz *et al.*, 2002). However, this does not necessarily mean that transposons are absent.

If not apparent in the mRNA, the presence of transposable elements or other cis-regulatory changes to *LvHIRZ* and *LvINA* can be detected from gDNA sequence. Unfortunately, gDNA sequence data could not be obtained for *LvHIRZ* and *LvINA*, and in its absence it is impossible to determine whether transposon insertions have disrupted cis-regulation of *LvHIRZ* and *LvINA*. However, the possibility of large transposon sequences such as Tam1 and Tam3 in cis-regulatory regions of *LvHIRZ* and *LvINA* may explain the lack of success in isolating gDNA sequence for both of these genes. Transposon insertions are a likely cause of changes in cis-regulatory elements and may result in disruption of specific repressor binding sites or insulator elements that influence chromatin structure, as demonstrated in *Drosophila* (Cai & Shen, 2001).

In addition to their accepted roles as transcriptional partners for KNOX proteins, specific BEL proteins have also been demonstrated to negatively regulate *KNOX* gene activity. For example the *Bel1*-like homeodomain genes *SAWTOOTH1* (*SAW1*) and *SAWTOOTH2* (*SAW2*) repress *KNOX* gene expression in *Arabidopsis* leaves (Kumar *et al.*, 2007). The KNATM family of KNOX proteins, which lack a homeodomain, may also selectively interact with BEL proteins, competitively inhibiting other *KNOX* genes

via the formation of inactive KNATM-BEL heterodimers (Kimura *et al.*, 2008; Magnani & Hake, 2008). Alterations in protein:protein binding domains may be another possible mechanism by which cis-regulatory sequences could influence elaboration of expression domains. However, for *LvHIRZ* and *LvINA*, at least, this seems unlikely as the protein:protein binding MEINOX domains appear to be normal when compared to those of other class 1 *KNOX* relatives. Such interactions also provide a possible trans-regulatory pathway by altering levels of interacting partners, however, there is no evidence to suggest that changes in the expression levels of potential interacting partners are responsible for elaboration of *KNOX* expression in *L. vulgaris*. Indeed, this would be difficult to demonstrate before first identifying the downstream targets and interacting factors of LvHIRZ and LvINA proteins.

6.3.4 Floral expression of *LvHIRZ* and *LvINA*: Implications for the role of *KNOX* genes in nectar spur development

It is possible that the developing nectar spur could function in a manner analogous to the SAM or a developing root, with a growing point at the very tip that may require *KNOX* gene activity to maintain a population of undifferentiated cells to allow for sustained growth and elongation of the spur. According to this *meristematic model*, nectar spurs would cease to develop when all the meristematic cells at the tip become exhausted, perhaps in relation to cessation of *tip-localised KNOX* gene activity. Whilst it is possible to view the growing nectar spur in this way, this model appears to be inconsistent with ontogenetic analyses of nectar spur development (see Chapter 3). For example, in *L. vulgaris* (and *D. fuchsii*), detailed analyses of cell morphology throughout development of the nectar spur clearly show that there is no localised population of apparently meristematic cells at the tip of the nectar spur. In addition, Golz *et al.* (2002) were unable to detect expression of meristematic markers that normally overlap with the expression of *KNOX* genes in the SAM, in the ectopic petal tubes of *Hirz-d153* and *Ina-d1* snapdragon mutants. Similarly, Golz *et al.* (2002) were unable to demonstrate that ectopic *AmHIRZ* and *AmINA* expression is localised to any one part of the petal and whether foci of *KNOX* gene expression correspond spatially with the developing ectopic petal tubes of *Hirz-d153* and *Ina-d1* snapdragon mutants. Extra-meristematic patterns of *LvHIRZ* and *LvINA* expression in developing *L. vulgaris* flowers is also diffuse, strongly opposing the meristematic model of nectar spur development outlined above, although the absence of *in situ* hybridisation data may preclude this conclusion and reveal precise localisation within an organ not detectable by PCR-based gene expression methods.

The reversed organ polarity of ectopic petal tubes in snapdragon mutants prompted Golz *et al.* (2002) to propose a different model for the involvement of *KNOX* genes in

nectar spur development in Antirrhineae. This model implicated the presence of an *organiser*; a concept common in animal development in which a group of cells, *the organiser*, can determine the polarity of growth and fate of more distant cells (Lecuit & Cohen, 1997; MacWilliams, 1983). In animal systems organisers work by secreting morphogens, biologically active substances that govern the pattern of tissue development and, in particular, the positions of the various specialized cell types. Morphogens act over long distance to influence the fate and activity of more distant cell populations. In addition to expression in the SAM, wild-type *AmHIRZ* and *AmINA* are expressed in the floral meristem and receptacle, proximal to developing petals, where they are proposed to establish a proximal organiser at the base of the corolla that controls growth and fate of cells along the proximal-distal axis (Golz *et al.*, 2002). Ectopic *KNOX* gene expression in the corolla of *Hirz-d153* and *Ina-d1* snapdragon mutants may therefore induce a novel axis of growth by generating an additional ectopic *organiser* (Golz *et al.*, 2002). Such a phenomenon has also been implicated in the formation of ectopic structures, with reversed organ polarity, from the bract of barley as a result of *KNOX* gene misexpression (Müller *et al.*, 1995; Williams-Carrier *et al.*, 1997). Establishing additional ectopic organisers, in animals at least, requires the transplantation of groups of cells that can act as organisers or, as may be the case for the *Hirz-d153* and *Ina-d1* mutants, ectopic expression of genes sufficient for organiser function, i.e. *KNOX* genes.

Whilst the concept of a proximal-distal corolla tube organiser located at the base of the developing petal is consistent with the observed basipetal pattern of lateral organ differentiation, in which the cells at the tip of an organ exit the cell cycle before those at the base (Donnelly *et al.*, 1999; Nath *et al.*, 2003; Poethig & Sussex, 1985), there is no evidence of similar patterns of *KNOX* gene expression in the receptacle of other flowers with tubular corollas such as tobacco (Nishimura *et al.*, 1999). *KNOX* gene expression has been shown in the pedicel of *Arabidopsis* where it functions to promote pedicel and internode growth, e.g. *KNAT1/ATK1* (Douglas *et al.*, 2002; Venglat *et al.*, 2002), but does not appear to act as a proximal-distal organiser of petal development.

Despite the fact that broad expression patterns for *LvHIRZ* and *LvINA* in flowers suggests that *KNOX* genes are not involved in sculpting the morphology of the nectar spur itself, but rather in somehow promoting its development (this study), in the absence of *in situ* hybridisation data there is no conclusive evidence to support or refute the organiser model proposed by Golz *et al.* (2002). However, the absence of similar expression patterns in well characterised tubular flowers like tobacco, suggest it is unlikely that *KNOX* genes are organising the growth and fate of cells along the proximal-distal axis in the tubular corolla of *A. majus* (Golz *et al.*, 2002). This makes the ectopic organiser concept an unlikely explanation for the development of ectopic petal tubes in the *Hirz-d153* and *Ina-*

d1 mutants and therefore, the natural spurs observed in other Antirrhineae such as *L. vulgaris*. However, evidence for additional caveats of the model proposed by Golz *et al.* (2002), e.g. a requirement for additional spatially restricting factors, are firmly supported by the broad expression patterns of *KNOX* genes in *L. vulgaris*.

In the *Hirz-d153* and *Ina-d1* snapdragon mutants, ectopic *KNOX* gene expression is not localised to any one part of the corolla and does not appear to correspond with the development of the ectopic petal tube. Despite similarly broad expression patterns for *LvHIRZ*, only a single nectar spur develops on the ventral part of the corolla tube. To explain how such broad patterns of *KNOX* gene expression are related to the development of ectopic petal tubes on only one part of the corolla tube, Golz *et al.* (2002) implicated the involvement of additional spatially restricting factors that presumably limit the activity of the KNOX protein, reviewed by Sablowski (2007). Silencing of petal-determining B-function genes in *Aquilegia vulgaris* (Kramer *et al.*, 2007), naturally occurring five-spurred radially symmetric flowers of *L. vulgaris* mutants (Cubas *et al.*, 1999), and the presence of multiple ectopic petal tubes in *radialis*,*Hirz-d153* double mutants (Golz *et al.*, 2002), all strongly suggest that nectar spur development is associated with canonical ABC and floral symmetry breaking genes.

In addition to floral symmetry genes such as *RADIALIS* there are many other possibilities that may restrict the development of nectar spurs to certain parts of the corolla. The availability of interacting partners and downstream targets of KNOX proteins, such as BEL proteins, may also be responsible for spatially restricting KNOX protein activity in a broader field of *KNOX* gene expression. Such targets may themselves be restricted by canonical ABC and floral symmetry breaking genes. Isolating downstream targets/interacting factors of the LvHIRZ protein may provide further insights into nectar spur development.

KNOX protein movement should also be considered to explain the precise spatial positioning of the *L vulgaris* spur despite broad patterns of floral *KNOX* expression. The maize KNOX protein KNOTTED is capable of moving intercellularly via plasmodesmata, a property that may be integral to its role in meristem maintenance (Jackson *et al.*, 1994; Kim *et al.*, 2003; Lucas *et al.*, 1995). Patterns of KNOX protein movement may be integral to the function of the protein. Broad patterns of *LvHIRZ* and/or *LvINA* gene expression may be associated with localised accumulation of the respective proteins during floral nectar spur development. Studies of protein localisation using immunolocalisation techniques could provide further insights in this regard, although initial attempts (data not shown) did not provide reliable data.

6.3.5 *LvHIRZ* and *LvINA* may play a role in carpel development

Aside from the publication of the *Hirz-d153* and *Ina-d1* snapdragon mutants (Golz *et al.*, 2002), few authors have characterised patterns of floral *KNOX* expression from initiation of the floral meristem (FM) to the cessation of floral development. However, in a number of cases, most notably in *Arabidopsis*, investigations of floral *KNOX* gene expression patterns are becoming increasingly common (Endrizzi *et al.*, 1996; Pautot *et al.*, 2001; Scofield *et al.*, 2007). Contrary to popular belief, these studies have shown that rather than being exclusively involved in maintenance of the SAM, *KNOX* genes may also be essential to floral patterning and organ identity.

Such studies have highlighted a particular role in carpel initiation. Both *STM* (Scofield *et al.*, 2007, 2008) and *KNAT2* (Endrizzi *et al.*, 1996; Pautot *et al.*, 2001) have been implicated in such a role as a direct result of extensive transgenic and mutant analyses. Constitutive expression of both genes result in ectopic carpel formation, whilst inducible silencing experiments and loss-of-function mutants result in either the complete absence of carpel tissues at the centre of the flower, or inhibited carpel growth resulting in malformed ovules and carpel placental tissues (Endrizzi *et al.*, 1996; Pautot *et al.*, 2001; Scofield *et al.*, 2007). In addition to work conducted with *Arabidopsis*, the dominant *GN1-R* maize mutant also forms ectopic carpels in flowers (Foster *et al.*, 1999).

Scofield *et al.* (2007, 2008) proposed a model for the role of *STM* in carpel initiation, highlighting the fact that in *Arabidopsis*, the carpels are formed from a residual population of stem cells located at the centre of the floral meristem (FM) (reviewed in Blázquez *et al.* (2006)). *STM* functions to maintain this stem cell niche until carpel development is initiated. In weak *stm* mutants, the cell niche that gives rise to the central carpel whorl is established but not maintained and, as a result, is consumed by development of the anthers (Scofield *et al.*, 2007). Some degree of functional redundancy has been implicated between *STM* and *KNAT2* in this process (Endrizzi *et al.*, 1996; Pautot *et al.*, 2001; Scofield *et al.*, 2007). In addition to a somewhat classical role in maintaining the stem cell niche that gives rise to the carpel tissues, *STM* has also been shown to directly promote the development of carpels and the associated meristematic placental tissues of the ovary (Pautot *et al.*, 2001; Scofield *et al.*, 2007). The termination of stem cell maintenance in the FM and the subsequent development of reproductive floral organs normally requires the activity of the MADS-box gene *AGAMOUS* (AG), which is activated in the centre of the FM by the floral regulator *LEAFY* (*LFY*) and the indeterminacy-promoting factor *WUSCHEL* (*WUS*). *AG* inhibits *WUS* thereby terminating stem cell maintenance and permitting the reproductive tissues to develop (Busch *et al.*, 1999; Lenhard *et al.*, 2001; Lohmann *et al.*, 2001). *Arabidopsis STM* appears to be able to promote carpel initiation

independently of *LEAFY* (*LFY*) and *AGAMOUS* (*AG*) expression, which allows the stem cell population to differentiate into carpels, perhaps by providing subtle activation of *AG* or its downstream targets (Scofield *et al.*, 2007).

Interestingly, (Golz *et al.*, 2002) described wild-type expression of *AmHIRZ* and *AmINA* in the marginal regions between developing stamen and carpel tissues but failed to recognise a potential role for these genes in carpel development itself. In *L. vulgaris* *KNOX* gene expression is also consistently demonstrated in reproductive tissues, including the ovary. Expression of both *LvHIRZ* and *LvINA* is detectable in the reproductive tissues, which comprise tissue from both stamens and ovary (impossible to separate by conventional dissection techniques). In particular *LvHIRZ* is expressed at 4 to 5-fold greater levels in the reproductive tissues versus the SAM. In the absence of precise *in situ* localisation data for both *LvHIRZ* and *LvINA*, it is difficult to determine a precise role for these genes in carpel development.

6.3.6 Conclusions

The identification of floral expression patterns for *LvHIRZ* and *LvINA* supports the data from snapdragon mutants (Golz *et al.*, 2002), indicating a role for *KNOX* genes in the evolution and development of nectar spurs in Antirrhineae. However, the organiser model proposed by Golz *et al.* (2002) does not appear to provide a sufficient explanation for the role of *KNOX* genes in spur development. As such, the specific role that *KNOX* genes may play in nectar spur development remains uncertain. However, insights may be gleaned from recent developments in understanding compound leaf development (see Chapter 9). Whilst it appears that *KNOX*, ABC and floral zygomorphy genes are necessary for nectar spur development, it is highly likely that there are many additional unknown factors involved in nectar spur development that await identification.

Nectar spurs have evolved multiple times independently in a broad range of angiosperm taxa, many of which have quite different floral morphologies to the long-tubed flowers of Antirrhineae. Although nectar spurs are almost always petal derived, in some species, nectar spurs develop from other organs such as sepals. To assume that findings from such limited taxon sampling is representative of all plants with spurs is misleading, as nectar spurs in other angiosperms may have evolved by alternative means. However, observations of nectar spur ontogeny from divergent angiosperm taxa suggest that where spurs are present they develop in a very similar manner (Chapter 3). Such findings suggest that *KNOX* genes play a fundamental role in the development of floral nectar spurs, making it likely that floral *KNOX* gene expression may be shared among even distantly related angiosperm taxa with floral nectar spurs, e.g. orchids, see Chapter 8.

Chapter 7

Characterisation of *A. majus* and *L. vulgaris KNOX* gene function by constitutive expression in *Nicotiana tabacum* cv. Samsun

7.1 Introduction

Determining novel gene/protein function requires a multi-faceted approach. Whilst gene expression analyses allow hypotheses about protein function to be developed, functional characterisation requires biochemical analyses and careful observation of the phenotypes of mutants and transgenic plants. The fundamental role played by *KNOX* genes in plant development has attracted considerable attention to the gene family over the last 15 years. As such, a plethora of loss/gain-of-function mutants and transgenic plants have been described from most of the accepted model plants including *Arabidopsis*, maize, barley, rice, tomato and tobacco (reviewed by Hake *et al.* (2004); Reiser *et al.* (2000)). Such mutants and transgenic plants have been integral in demonstrating the function of KNOX proteins.

7.1.1 Typical *KNOX* mutant phenotypes

Typical *KNOX* gain-of-function (GoF) mutants include the maize *KNOTTED1* (*KN1*), *ROUGH SHEATH1* (*RS1*), *LIGULELESS3* (*LG3*) and *GNARLY1* (*GN1*) mutants, all of which are characterised by perturbations at the leaf blade-sheath boundary (Foster *et al.*, 1999; Muehlbauer *et al.*, 1999; Schneeberger *et al.*, 1995; Vollbrecht *et al.*, 1991). In maize, normal leaves have a proximal sheath and distal blade, separated by the ligule

and auricle. The dominant mutants *KN1, RS1, LG3* and *GN1* affect the organization of these four tissues, causing distal tissues to adopt proximal identities. For example, *KN1* leaf blades have ectopic sheath, auricle and ligule tissues generating *knots* in the leaf lamina. In contrast, *LG3* displaces the ligule into the blade (Fowler & Freeling, 1996; Freeling & Hake, 1985; Sinha & Hake, 1994). The strongly expressed *GN1-R* allele also affects the husk leaves and leads to ectopic carpels in flowers (Foster *et al.*, 1999).

Dominant GoF mutants from eudicots, such as the tomato *CURL* (*CU*) and *MOUSE-EAR* (*ME*) mutants, have leaves that show extensive ramification of the compound leaf, suppression of apical dominance, and retardation of growth (Parnis *et al.*, 1997). The GoF *HIRZINA* and *INVAGINATA* snapdragon mutants have small, rounded leaves that frequently develop ectopic trichomes and midribs with a buckled leaf lamina that results from excessive cell proliferation in the central and proximal regions of the leaf lamina. In addition to the leaf phenotype, *Hirz-d153* and *Ina-d1 Antirrhinum* mutants induce growth of additional ectopic petal tubes (Golz *et al.*, 2002) that have inspired much of the work presented in this thesis, see section 1.3.4.

Loss-of-function (LoF) mutants have proven to be equally informative in determining *KNOX* gene function. Plants carrying the *stm-1* allele in an *Arabidopsis* Landsberg erecta background produce cotyledons but no further components of the shoot system (Barton & Poethig, 1993; Long *et al.*, 1996). Similar recessive mutants have since been discovered, for example, maize LoF *kn1* mutants form only a limited shoot (Vollbrecht *et al.*, 2000). In each case such mutants provide firm support for the role of *KNOX* genes in meristem maintenance and complex leaf patterning. However, many of the *KNOX* genes identified from mutants such as these have been subject to transgenic analyses in both endogenous and heterologous hosts in order to better understand the function of these proteins.

7.1.2 Constitutive expression as a tool to infer protein function

Plant transformation provides a mechanism to manipulate the timing, level and tissue specificity of a particular gene's expression. Preliminary analyses of protein function utilise transgenic approaches in which a gene may be constitutively expressed (GoF) and/or specifically silenced (LoF) using techniques such as RNA interference (RNAi), antisense RNA or Viral Induced Gene Silencing (VIGS). Dramatically increasing or reducing the level of expression of a target gene, particularly genes such as transcription factors, often results in very obvious phenotypes. The function of endogenous and heterologous genes/proteins can be examined in this way using a suitable host, sensible experimental controls and sufficient experimental replication, such that the phenotype of transgenic plants may be confidently attributed to the gene/protein under test.

A large number of *KNOX* genes identified from both LoF and GoF mutants have

been assessed further using transgenic approaches to better understand the functions of the encoded proteins in plant development. Transgenic experiments such as these have focused predominantly on the model plants *Arabidopsis*, tobacco, tomato and rice, see below, but also a number of more unusual taxa such as Poplar (Groover *et al.*, 2006), potato (Rosin *et al.*, 2003) and dandelion (*Taraxacum officinale*; Müller *et al.* (2006)). Surprisingly, despite expression in a broad range of distantly related taxa, a number of common phenotypes, reminiscent of morphological abnormalities observed in *KNOX* misexpression mutants, can be observed in transgenic LoF and GoF lines.

Transgenic *Arabidopsis* plants constitutively expressing *KNOX* genes driven by the constitutive CaMV 35S promoter have highly lobed leaves and ectopic meristems (Chuck *et al.*, 1996; Jouannic *et al.*, 2007; Lincoln *et al.*, 1994; Matsuoka *et al.*, 1993; Sano *et al.*, 2005; Sundås-Larsson *et al.*, 1998). By contrast, in transgenic *Arabidopsis* plants in which *KNOX* activity has been silenced, e.g. *STM* RNAi *Arabidopsis* (Scofield *et al.*, 2007) and its close relative *Cardamine hirsuta STM* RNAi silenced plants (Hay & Tsiantis, 2006), the phenotype strongly resembles that of naturally occurring LoF mutants such as *Arabidopsis shoot meristemless*.

In tobacco, constitutive expression of several endogenous *KNOX* genes e.g. *NTH15*, plus heterologous genes such as maize *KN1* and *OSH1* from rice, results in leaf phenotypes such as rumpling, reduced leaf lamina, delayed senescence, and the formation of ectopic shoots on the leaves (Nishimura *et al.*, 2000; Ori *et al.*, 1999; Postma-Haarsma *et al.*, 1999; Sakamoto *et al.*, 1999, 2001; Sato *et al.*, 1996, 1998; Sinha *et al.*, 1993; Tamaoki *et al.*, 1997). Transgenic tomato plants constitutively expressing both endogenous or heterologous *KNOX* genes have a particularly dramatic phenotype. Like the tomato *CU* and *ME* mutants, transgenic tomato plants constitutively expressing *KNOX* have supercompound leaves (Hareven *et al.*, 1996; Janssen *et al.*, 1998a; Kimura *et al.*, 2008).

Constitutive *KNOX* gene expression phenotypes have also been well characterised in monocots. Rice plants constitutively expressing *KNOX* genes such as *OSH1*, *OSH3*, *OSH6*, *OSH15*, *OSH43* and *OSH71*, under the control of either the CaMV 35S or the rice *Act1* promoter (Nagasaki *et al.*, 2001b; Sato *et al.*, 2002; Sentoku *et al.*, 2000), were found to have severely malformed leaves with ectopic knots and displaced ligule formation resembling that observed in maize *KNOTTED1* mutants. In each case, constitutive expression of class 1 *KNOX* genes in monocots and eudicots is often associated with significant alterations in the levels of plant growth substances such as cytokinins and GA (see section 1.5.3.2).

Monocot and eudicot derived *KNOX* genes often induce similar phenotypic effects when constitutively expressed in the same heterologous host. However, the effects of constitutive *KNOX* gene expression in eudicots can be distinct from those in monocots. For

example, constitutive expression of maize *Kn1* in tomato generates plants with supercompounded leaves and loss of apical dominance (Hareven *et al.*, 1996). However, when maize *Kn1* is constitutively expressed in barley under a ubiquitin promoter, transgenic barley plants showed no abnormal leaf phenotype, although ectopic meristems were formed on the adaxial surface near the lemma/awn transition zone (Williams-Carrier *et al.*, 1997). These data indicate that monocots and eudicots differ in the competence of tissues to respond to, or modulate the function of, KNOX proteins (Williams, 1998).

Despite subtle phenotypic differences, the close resemblance of transgenic plants with a broad range of genetic backgrounds and evolutionary histories lends further support for the fundamental role of *KNOX* genes in maintaining cell indeterminacy in relation to plant growth substance such as cytokinins and GA (Hake *et al.*, 2004; Hay *et al.*, 2009; Reiser *et al.*, 2000).

7.1.3 Research aims and objectives

Golz *et al.* (2002) provided a thorough exploration of the *Antirrhinum Hirz-d153* and *Ina-d1* mutant phenotypes in relation to ectopic floral expression of the class 1 *KNOX* genes *HIRZINA* and *INVAGINATA*. Whilst phenotypic effects of ectopic class I *KNOX* gene expression in floral organs has been described for only a small number of LoF/GoF mutant and transgenic plants, e.g. Sinha *et al.* (1993), induction of a novel axis of petal growth appears unique to the *Hirz-d153* and *Ina-d1* mutations. This suggests that such a property may be the consequence of the genetic background within which ectopic *KNOX* expression had occurred, i.e. *Antirrhinum majus* and its close relatives, or a fundamental property of the AmHIRZ and AmINA proteins themselves. Transgenic analyses of *AmHIRZ* and *AmINA* to establish the functional properties of the encoded proteins were not conducted alongside initial expression analyses by Golz *et al.* (2002). The principal aim of this section of work is to determine whether Am/LvHIRZ and Am/LvINA are functionally equivalent to related KNOX proteins characterized in model species and to explore the sufficiency of these proteins to induce novel outgrowths on petals, resembling the ectopic petal tubes observed in the *Hirz-d153*/*Ina-d1* mutants and the nectar spurs of *L. vulgaris*.

The sufficiency of these proteins to induce petal outgrowth will be inferred by observing the phenotype of transgenic tobacco plants independently expressing the full-length coding sequence of the four *KNOX* genes driven by the constitutive CaMV 35S promoter (see section 2.7), which has been shown to drive expression in all tissue types in tobacco from an early developmental stage (Benfey *et al.*, 1989; Harpster *et al.*, 1988). In addition to constitutive expression, silencing of the endogenous *LvHIRZ* and *LvINA* transcripts will be attempted using Tobacco Rattle Virus-VIGS (TRV-VIGS, Ratcliff *et al.* (2001)),

employing a protocol optimised for *Aquilegia vulgaris* from the lab of Elena Kramer (Harvard University, USA, see section 2.8).

7.1.3.1 Tobacco: an appropriate heterologous host for transgenic analyses of *Am/LvHIRZ* and *Am/LvINA*

The context of gene expression is imperative to the function of the encoded protein. Therefore, when attempting to infer protein function by plant transformation the genetic background, i.e. the choice of host organism, is critical to the development of a successful transgenic experiment. An ideal transformation experiment to explore the function of a target gene/protein would involve transformation of the donor plant itself, i.e. gene/protein function would be explored endogenously. In this case such an experiment would involve genetic transformation of the developmental model plant *Antirrhinum majus* and its close relative, *L. vulgaris*. However, to date there are no reports in the literature of genetic transformation of *L. vulgaris* but, given the closely related nature of *L. vulgaris* and *A. majus* (Albach *et al.*, 2005; Ghebrehiwet *et al.*, 2000), one may expect that a transformation protocol for snapdragon could be employed in the analysis of the *HIRZ* and *INA* orthologs from both species. Despite numerous reports of stable genetic transformation of *A. majus* (Cui *et al.*, 2001, 2003, 2004; Senior *et al.*, 1995), transformation efficiency is very low and a reliable transformation protocol for *A. majus* has remained elusive (Hudson *et al.*, 2008). The absence of transformation protocols for *A. majus* and *L. vulgaris* necessitates the use of a heterologous host. Tobacco (*Nicotiana tabacum* cv. Samsun) and *Arabidopsis* have been used as heterologous hosts to investigate the function of many KNOX proteins, see section 7.1.2. Using either of these organisms as host for *Am/LvHIRZ* and *Am/LvINA* permits direct comparisons of the resultant phenotypes to those observed for other well characterised *KNOX* genes reported previously in the literature.

Both tobacco and *Arabidopsis* are amenable to genetic transformation and have rapid and reliable transformation protocols. Whilst transformation of *Arabidopsis* is quicker and easier than that of tobacco, taking only 6-8 weeks, flowers of *Arabidopsis* are small and simple. Such flowers are therefore unlikely to generate informative phenotypes related to floral nectar spur development. Also, *Arabidopsis* is a member of the rosid group of eudicots, which is phylogenetically relatovely distantly related to snapdragon and toadflax, both of which are members of the asterid group of eudicots (Stevens, 2001). In contrast to *Arabidopsis*, tobacco is also a member of the asterids (Stevens, 2001), therefore tobacco represents a closer phylogenetic background. Tobacco flowers are large, making them easy to manipulate and score phenotypes. Furthermore, tobacco flowers are characterised by a long, tubular corolla that is subtly zygomorphic. The presence of a

tubular corolla and floral zygomorphy means that tobacco is likely to possess a similar developmental-genetic background to the flowers of snapdragon and toadflax which are superficially similar, although considerably more complex in terms of structural modifications. From a practical perspective, tobacco is a hardy plant that is fast growing, self pollinating, and produces large quantities of seed that require no special treatment to germinate. Tobacco is amenable to stable genetic transformation using a well established leaf callus protocol and a modified laboratory strain of the plant pathogen *Agrobacterium tumefaciens*, strain GV3101, as a means of inserting the foreign DNA fragment (T-DNA) into the tobacco genome, see section 2.7 (Horsch *et al.*, 1985). Although this protocol requires a minimum of 14-18 weeks before flowering, transformation efficiency is relatively high, typically exceeding 50%.

7.2 Results

7.2.1 *L. vulgaris* TRV-VIGS

Four independent constructs were generated for the endogenous silencing of *LvHIRZ* and *LvINA* in *L. vulgaris* using TRV-VIGS (Figure 7.1); 1) full-length *LvHIRZ* cDNA, containing the highly conserved homeodomain to encourage global *KNOX* gene silencing; 2,3) the meinox encoding domain of *LvHIRZ* or *LvINA*, to promote specific silencing of either gene in the absence of the highly conserved homeodomain and; 4) the control gene *L. vulgaris phytoene desaturase* (*LvPDS*), which generates bleached white leaves when silenced. TRV-VIGS constructs were transferred to *L. vulgaris* by vacuum infiltration with *A. tumefaciens* strain GV3101 carrying the TRV-VIGS constructs, section 2.8.1. Unfortunately, although 100% of *L. vulgaris* plants survived the procedure, no phenotypes were observable, even amongst control *LvPDS* lines. Time limitations prevented further optimisation of TRV-VIGS for *L. vulgaris*.

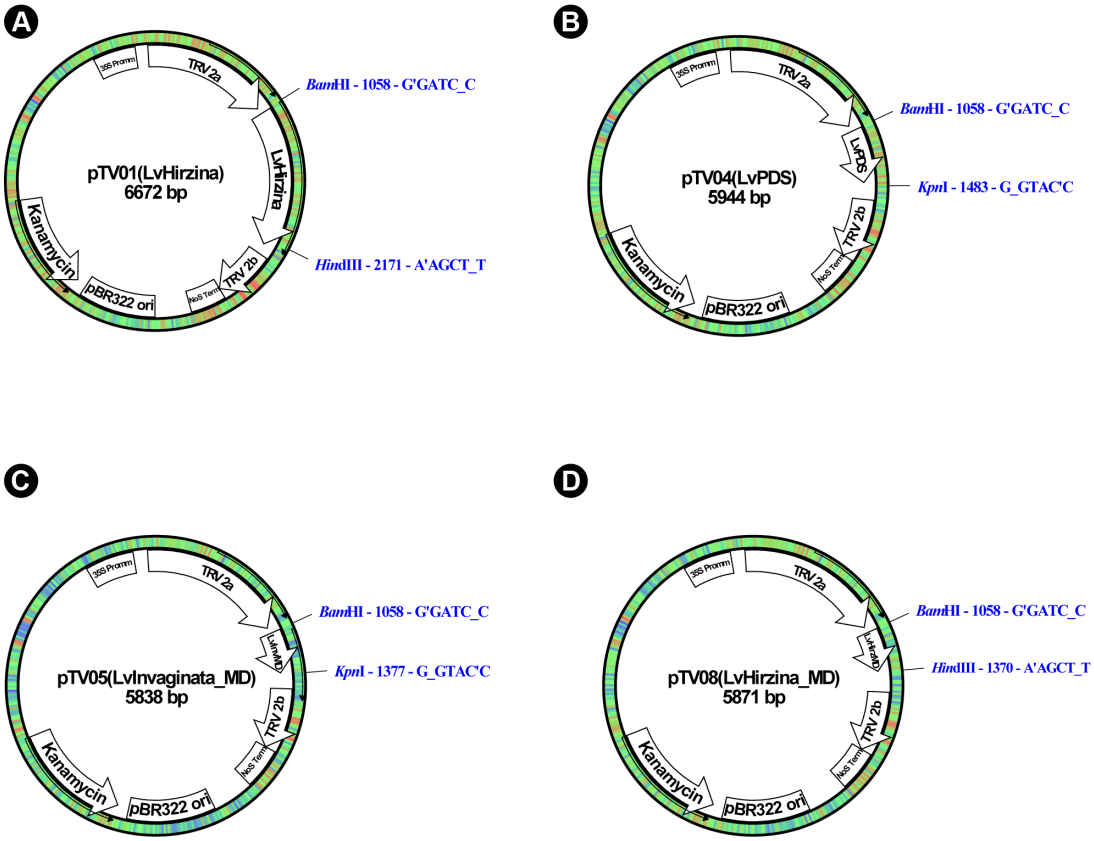


Figure 7.1: Vector maps of *L. vulgaris* TRV-VIGS constructs
Open reading frames are indicated by arrows on the plasmid backbone. Origins, promoters, terminators, operons and antibiotic resistance genes are indicated by block arrows. Primer annealing sites, restriction sites and multiple cloning sites are indicated by tags outside of the DNA backbone. A. pTV01, containing the full-length cDNA of *LvHIRZ*. B. pTV04, containing the positive control gene *L vulgaris phytoene desaturase* (*LvPDS*). C. pTV05, containing the meinox encoding domain of *LvINA*. D. pTV08, containing the meinox encoding domain of *LvHIRZ*.

7.2.2 Constitutive expression of *Am/LvHIRZ*, *Am/LvINA* in tobacco

Four independent constitutive expression constructs were generated according to the methods outlined in section 2.7.1 (Figure 7.2). In each case the full-length coding sequences of *AmHirz*, *AmINA*, *LvHIRZ* and *LvINA* was cloned, in sense orientation, between the CaMV 35S promoter and terminator to drive constitutive expression of each gene. Each construct was transferred to *Nicotiana tabacum* cv. Samsun via leaf disc transformation using *Agrobacterium tumefaciens* strain GV3101, using a modified protocol from Horsch *et al.* (1985) (see sections 2.7.3, 2.7.4). Presence of the T-DNA in transgenic plants transformed with each of the four constructs was confirmed by gDNA PCR and T-DNA expression confirmed by RT-PCR using full-length primers specific for each of the transgenes (Appendix B). Vegetative and floral phenotypes were recorded from multiple independent transgenic tobacco lines.

Only transgenic tobacco phenotypes for constitutive expression lines of *AmHIRZ*, *AmINA* and *LvHIRZ* are presented here, as transgenic plants constitutively expressing *LvINA* could not be regenerated. Whilst transgenic *LvINA* callus formed shoots, these failed to grow sufficiently and were much shorter than those derived for other constructs, as a result 35S::*LvINA* shoots failed to root when transferred to rooting media. Transgenic tobacco phenotypes were compared to wild-type tobacco plants and transgenic tobacco control plants transformed with an unaltered, *empty* 35S construct (Figure 7.3). Three independent empty 35S tobacco control lines were generated (35S control lines 1.1, 2.2 and 3.1). Presence of the T-DNA, consisting of only the 35S promoters and terminator sequences, was confirmed by gDNA PCR. In terms of vegetative and floral phenotype, empty 35S control lines were almost identical to wild-type tobacco plants, although the flower colour was slightly paler (Figure 7.2). Observed phenotypes for experimental transgenic tobacco lines were therefore likely to be the result of the transgene rather than an artefact of the transformation protocol.

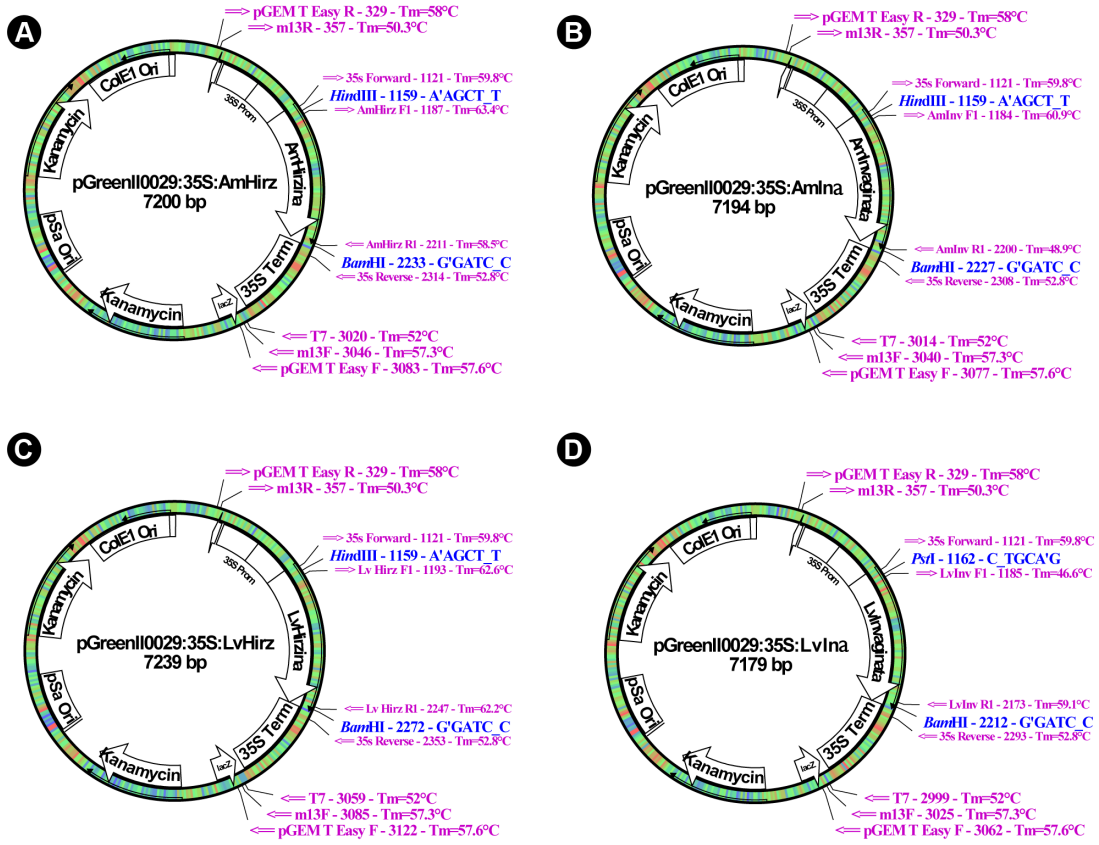


Figure 7.2: Vector maps of CaMV 35S constructs used in transgenic experiments
Open reading frames are indicated by arrows on the plasmid backbone. Origins, promoters, terminators, operons and antibiotic resistance genes are indicated by block arrows. Primer annealing sites, restriction sites and multiple cloning sites are indicated by tags outside of the DNA backbone. A. 35S::*AmHIRZ*. B. 35S::*AmINA*. C. 35S::*LvHIRZ*. D. 35S::*LvINA*.

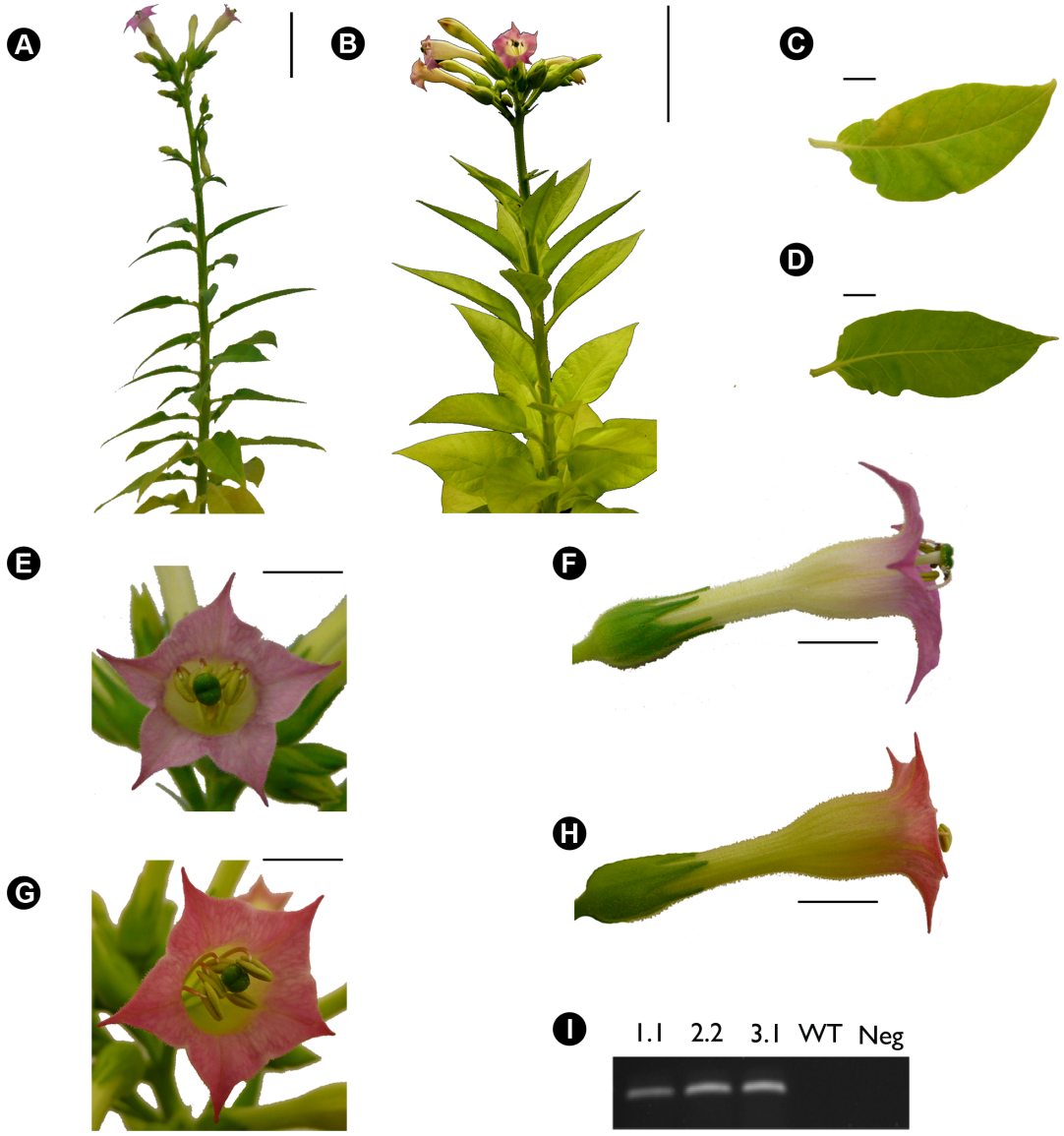


Figure 7.3: Wild-type versus transgenic empty 35S tobacco control plants
Empty 35S transgenic tobacco controls are identical to wild-type tobacco. A, B. Comparison of mature empty 35S transgenic (A) and wild-type (B) tobacco plants. Note the difference in scale. C, D. Mature leaves of empty 35S transgenic (C) and wild-type (D) tobacco plants from comparable internodes. E-H. Floral morphology of empty 35S (E [top], F [side]) and wild-type (G [top], H [side]) tobacco flowers at anthesis is identical, although flowers of transgenic control plants are a paler pink in colour. I. gDNA PCR of empty 35S transgenic control lines using the primers 35S F/R (148bp; Appendix B) confirms the presence of the T-DNA and its absence from wild-type (WT) plants. Scale bars = 5cm in A, B; 1cm in C-H.

7.2.2.1 Efficacy of the tobacco transformation protocol

The modified tobacco transformation protocol employed in this work was highly efficient. For each of the four *KNOX* gene constructs, and the empty 35S control, kanamycin resistant callus developed on tobacco leaf discs within 3-4 weeks of *A. tumefaciens* inoculation. In 4-6 weeks more than 90% of inoculated leaf discs produced callus. Regenerated shoots were visible from as early as five weeks but of the 90% of leaf discs that generated callus, only approximately 50-60% regenerated shoots. The incidence of shoot regeneration for the 35S::*LvINA* construct was particularly low, typically less than 10%, and none of these could be encouraged to root. Regenerated shoots from 35S::*AmHIRZ*, 35S::*AmINA* and 35S::*LvHIRZ* callus were transferred to rooting media by approximately 5-8 weeks and, by 8-10 weeks, had a sufficiently developed root system to support growth in compost in a controlled environment growth room. Genomic DNA was extracted from young tobacco shoots in rooting media and analysed by PCR to confirm T-DNA transfer. Tobacco *ubiquitin* was used as a positive control. In weak/intermediate phenotype plants a minimum of 14-18 weeks was required to complete the transformation protocol from *A. tumefaciens* inoculation to flowering. Plants with increasingly severe phenotypic abnormalities had significantly delayed flowering by up to 6 months (see sections 7.2.2.2, 7.2.2.3).

Using *A. tumefaciens* and the 35S constitutive expression vectors, the integration site of the transgene into the gDNA of the host plant is essentially random. As such there is always a possibility that the T-DNA may be inserted into heavily silenced areas of the genome. To confirm expression of the T-DNA total RNA was extracted from the leaves of transformants and expression of the T-DNA assessed by non-quantitative RT-PCR. Random integration also raises a second problem, the phenotype generated may be the result of the location of transgene integration rather than the sequence of the T-DNA itself. In order to counteract this problem multiple independent lines were generated for each of the constructs tested to be sure that any abnormal morphology observed in transgenic plants was due to the introduced gene, rather than to positional effects. Seven independent lines were identified constitutively expressing *AmHIRZ* (35S::*AmHIRZ* lines 6, 9, 17, 20, 23, 24, 26; Figures 7.4, 7.5), four independent lines, derived from separate callus, were identified constitutively expressing *AmINA* (35S::*AmINA* lines 3, 4, 5, 10; Figures 7.6) and a further four independent lines were identified constitutively expressing *LvHIRZ* (35S::*LvHIRZ* lines 1, 4, 5, 6; Figure 7.7). For each line multiple clones were derived from the same callus and sampled as additional representatives of each independent transgenic line, e.g. 3.1 and 3.2 would be clonal representatives of line 3.

7.2.2.2 Vegetative phenotypes of tobacco transformants

Tobacco transformants constitutively expressing *AmHIRZ*, *AmINA* and *LvHIRZ* differed significantly from wild-type tobacco and empty 35S transgenic control plants (Figures 7.3, 7.4, 7.6, 7.7). Both within and between transgenic lines the phenotype varied in severity but was consistent in a single plant. The morphology of the 15 independent primary transformant lines (seven for 35S::*AmHIRZ* and four each for 35S::*AmINA* and 35S::*LvHIRZ*), can be divided into three phenotypic categories ranging from weak to severe. Transgenic plants with a weak phenotype were rare, e.g. 35S::*AmHIRZ* 26.2 (Figures 7.4, 7.5), and had slightly wrinkled leaves with a curved mid-vein. In addition, the leaves were moderately shorter and wider than wild-type and control plants.

By far the most common phenotypic class are those plants with intermediate phenotypic perturbations. Four of the 35S::*AmHIRZ* lines (6, 9, 17, 20; Figure 7.4), representatives from all of the 35S::*AmINA* lines (3, 4, 5, 10; Figure 7.6), and some representatives from 35S::*LvHIRZ* lines 5 and 6 (Figure 7.7), can be classified as having intermediate phenotypic characters. Plants with intermediate phenotypes have elongated stems and thickened leaves of reduced size that are mildly to deeply lobed with disrupted symmetry across the less prominent, and shorter, mid-vein, e.g. 35S::*AmHIRZ* 6.3 (Figure 7.4). As phenotypic severity increases the length of the petiole decreases significantly such that the petiole of more severely affected transformants is indistinguishable from the leaf blade. Older leaves are wrinkled, similar to those found on plants with weak phenotypes (Figures 7.4, 7.6, 7.7). In some instances, e.g. 35S::*AmHIRZ* 6.3 (Figure 7.4) and 35S::*AmINA* 4.1 (Figure 7.6), additional shoots and meristems form on the adaxial leaf surfaces. In all lines vegetative growth was prolonged as a result of decreased leaf senescence and continued branching from axillary buds that are normally suppressed by strong apical dominance in wild-type tobacco.

Plants with the most severe phenotypes include 35S::*AmHIRZ* lines (23, 24; Figure 7.4) and representatives from all of the 35S::*LvHIRZ* lines (1, 4, 5, 6; Figure 7.7). Severe phenotypic perturbations included often barely recognisable, thickened leaves that were severely reduced in size. Such leaves lack a petiole and have disrupted leaf venation, lacking any obvious mid-vein. Apical dominance is almost entirely absent, axillary buds develop into vegetative stems, generating a bushy growth habit. Plants with severe phenotypes are significantly dwarfed as a result of reduced internode elongation, in the most extreme cases reaching no more than 5-10cm in height (e.g. 35S::*LvHIRZ* 5.1; Figures 7.4, 7.7) compared to wild-type tobacco, which can be up to 80-100cm tall (Figure 7.3).

7.2.2.3 Floral phenotypes of tobacco transformants

A grade of morphological abnormalities was also observable in flowers and inflorescences, which correlated with the aberrant leaf shapes (Figures 7.5, 7.6, 7.7). The number of floral organs, however, was never affected in any of the transformants. Flowers from plants with weak phenotypes were borne on inflorescences that were wild-type in terms of floral density and branching pattern. The flowers themselves were also wild-type in appearance with a five-lobed, although less vividly coloured, pink corolla. Flowering time of transformants with weak phenotype was relatively rapid, occurring within a similar time frame to wild-type tobacco plants, e.g. 35S::*AmHIRZ* 26.2 (Figure 7.5).

Flowers produced by transformants with intermediate phenotypes show a high degree of variability along a spectrum toward greater severity. All transformants with intermediate leaf phenotypes, e.g. 35S::*AmHIRZ* 6.3, 9.2, 17.2, 20 and 35S::*AmINA* lines (3, 4, 5, 10) plus representatives from 35S::*LvHIRZ* lines 5 and 6, produced flowers (Figures 7.5, 7.6, 7.7). However, flowering time was delayed in plants towards the more severe end of the spectrum. Often flowering is sustained for longer in transformants compared to wild-type as axillary branches, normally suppressed in wild-type tobacco, produce additional flowers. As the leaf phenotype increases in severity, flower colour becomes increasingly pale pink, and the corolla tube is disrupted, often narrow and wrinkled.

Perhaps the most significant result was that numerous distinct sac-like outgrowths of tissue form on the margins of the fused petals of the corolla tube, e.g. 35S::*AmHIRZ* 6.3 and 9.2 (Figure 7.5). Often only a single outgrowth forms between a pair of petals, most commonly on the ventral side of the corolla tube. Closer examination of these sac-like structures demonstrates that they are not merely folds of tissue but that they retain their shape throughout manipulation and longitudinal dissection, suggesting that they are discrete proximal-distal outgrowths of the corolla tube, often several millimetres in size. The corolla lobes at the proximal region of the tube itself may become more or less dissected, compare 35S::*AmHIRZ* 20 and 17.2, 35S::*AmINA* 4 and 35S::*LvHIRZ* representatives from lines 5 and 6 (Figures 7.5, 7.6, 7.7). The length of the corolla tube is reduced in flowers closer to the more severe end of the spectrum, such that the stigma is exserted far beyond the end of the corolla tube. In contrast, the anther filaments are significantly reduced in length, generating a pin-like flower with short stamens and a long pistil that disrupts self-pollination. However, if pollination is effective, either naturally or artificially, all plants bear numerous seeds.

Transformants with moderately-severe leaf phenotypes, e.g. 35S::*AmINA* 4 and representatives from 35S::*LvHIRZ* lines 1, 4 and 6, have more dramatic floral perturbations (Figures 7.6, 7.7). Flowering is delayed by up to six months in severe lines but is sustained for some time on axillary branches. Inflorescences become crowded by numerous small

flowers as a result of short internodes. Flowers are almost white and somewhat shorter in length than those of wild-type. The margins are severely dissected, and the tube is compressed around the central reproductive structures. The corolla tube is shorter and narrower than the flowers of intermediate phenotype transformants, with a style up to twice the length of the corolla tube, e.g. 35S::*AmINA* 4 and 35S::*LvHIRZ* 1 and 4 (Figures 7.6, 7.7). Flowers on plants with severe phenotypes have short, curled stamens with shruken anthers, e.g. 35S::*LvHIRZ* 1 and 4 (Figure 7.7). In many cases numerous small buds are formed and quickly aborted, the corolla tube and anthers of any remaining flowers failed to elongate (Figures 7.6, 7.7). As a result the flowers themselves were short-lived and almost completely infertile. In the most severely affected lines, e.g. 35S::*AmHIRZ* 23 and 24, plus 35S::*LvHIRZ* 5.1 (Figures 7.4, 7.7), transformants were maintained for a period of up to 12 months but failed to flower.

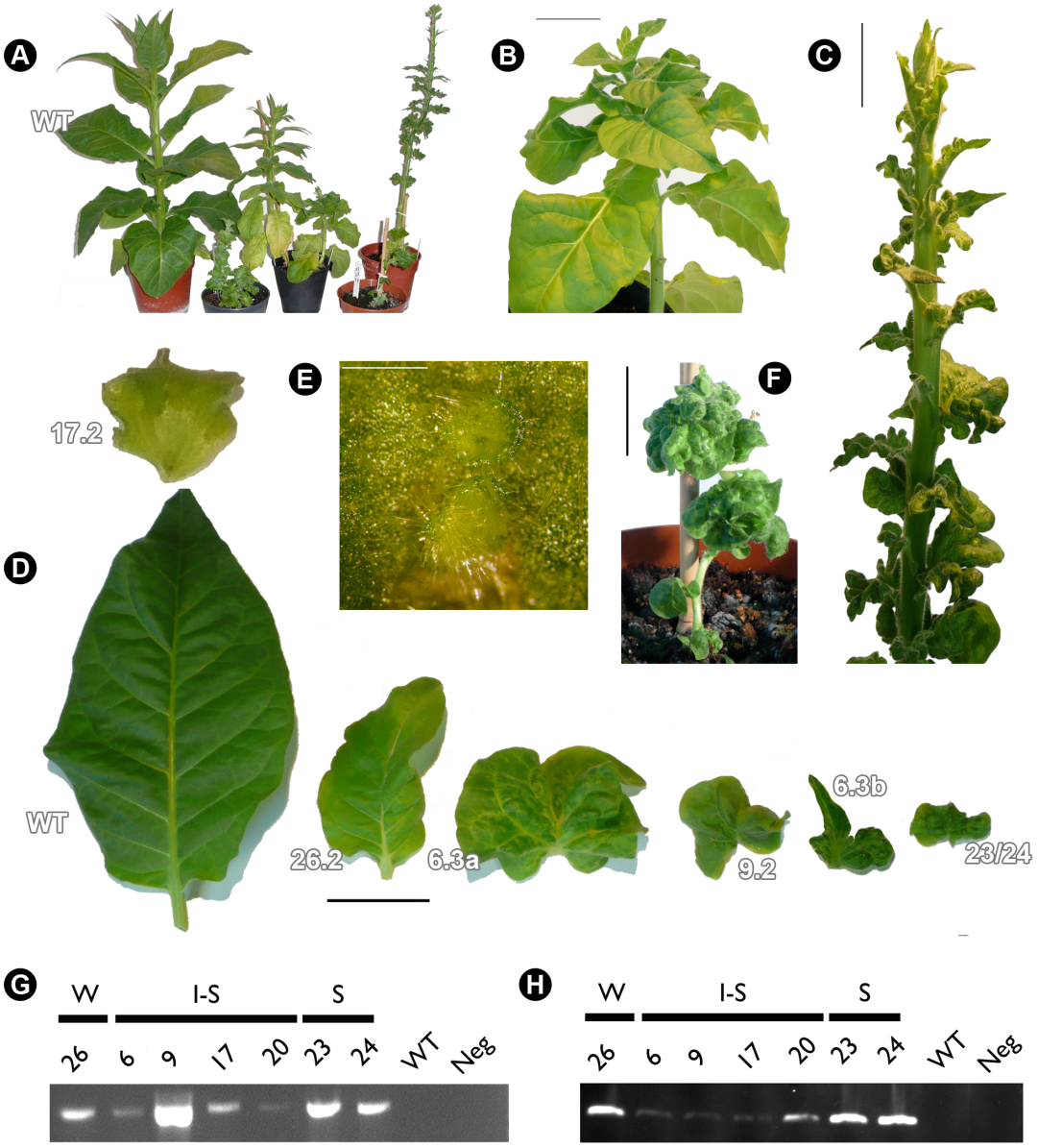


Figure 7.4: Vegetative phenotypes of 35S::*AmHIRZ* transformants
Phenotypes can be divided into three categories, weak (W), intermediate (I-S) and severe (S). A. 35S::*AmHIRZ* tobacco transformants exhibit a range of vegetative phenotypes. B. The mild vegetative phenotype, 26.2, has slightly wrinkled leaves. C, E. Plants with an intermediate phenotype, e.g. 6.3, have elongated stems and thickened leaves of reduced size that are mildly to deeply lobed with disrupted symmetry and ectopic shoots (E) on the adaxial leaf surface. D. The range of leaf phenotypes in 35S::*AmHIRZ* transformants, note the curved mid-vein in 26.2 and disrupted lateral leaf veins in lower leaves (6.3a) and leaf lobing in upper leaves (6.3b) of intermediate phenotype plants. Leaf shape/size is severely disrupted with increasing phenotypic severity (lines 23/24). F. The most severe lines, e.g. lines 23 and 24, show severe dwarfism. G, H. gDNA (G) and non-quantitative RT-PCR (H) of 35S::*AmHIRZ* transformants using gene specific primers (1,062bp; Appendix B) confirms the presence and expression of the T-DNA, which is absent from wild-type (WT) plants. Scale bars = 5cm in B, C, E, F; 1mm in D.

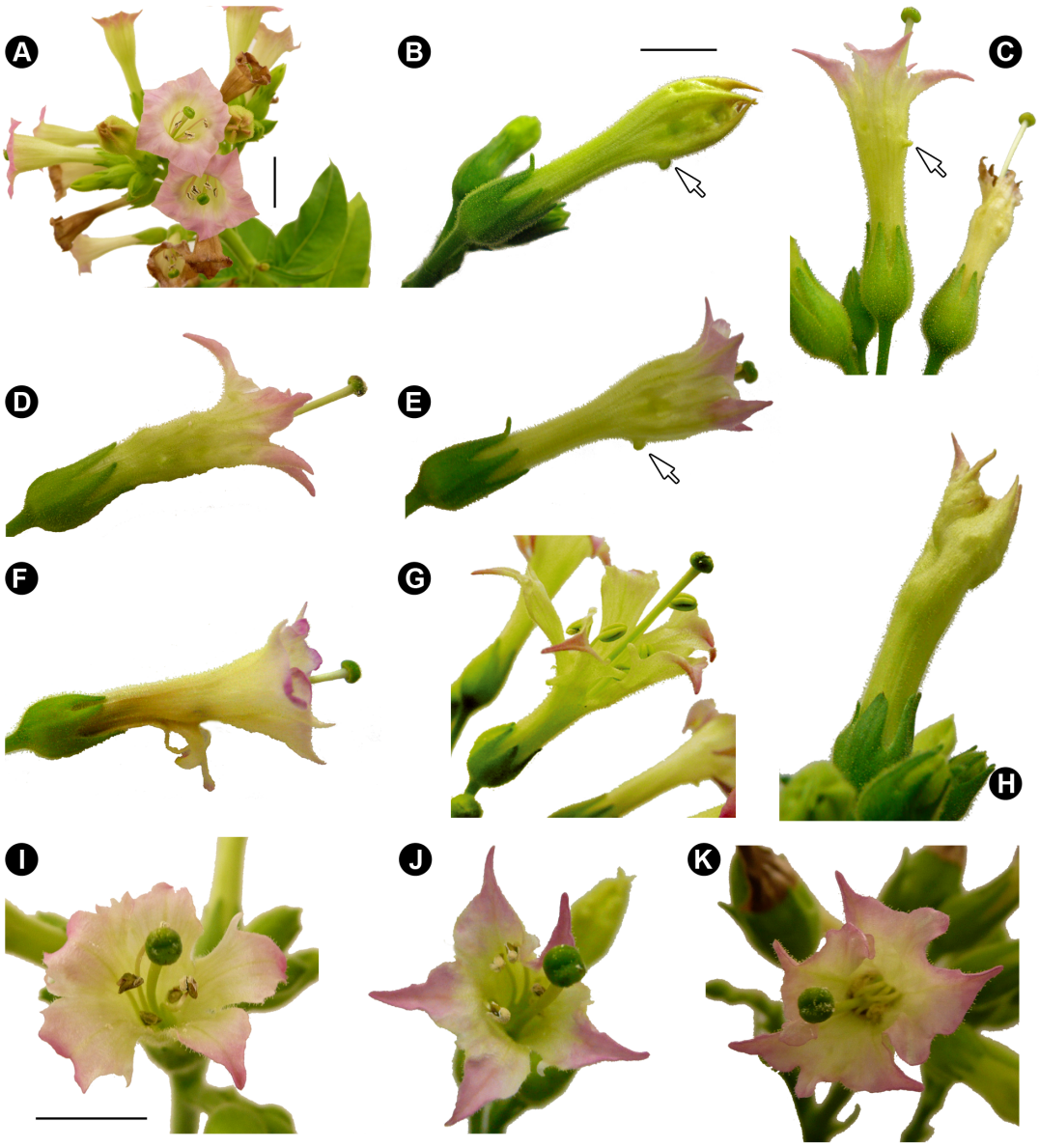


Figure 7.5: Floral phenotypes of 35S::*AmHIRZ* transformants
Severity of floral phenotypes of 35S::*AmHIRZ* tobacco transformants correspond to those in leaves, which varies within and between lines. As the severity of phenotype increases flowers are less darkly pigmented A. Flowers of weak phenotype lines (26.2) have normal flowers identical to wild-type and transgenic controls. B, C, E, K. Lines 6.3 (B, E) and 9.2 (C, K) have almost normal flowers with ectopic sac-like outgrowths of tissue that form on the margins of the fused petals of the corolla tube. Line 9.2 flowers show slightly more severe phenotypes (C, K) with a shorter, dissected corolla tube. D, F, I, J. Line 20 flowers have a short dissected corolla tube (I, J) exposing the stigma and style. In addition the corolla tube is heavily wrinkled with numerous perturbations (D) and additional tissue growth (F). G, H. Line 17.2 plants show intermediate leaf phenotypes and variable floral phenotypes including severely dissected (G) and wrinkled corolla tubes (H). Arrows indicate ectopic sac-like outgrowths. Scale bars = 1cm. Figures B-H were photographed to scale.

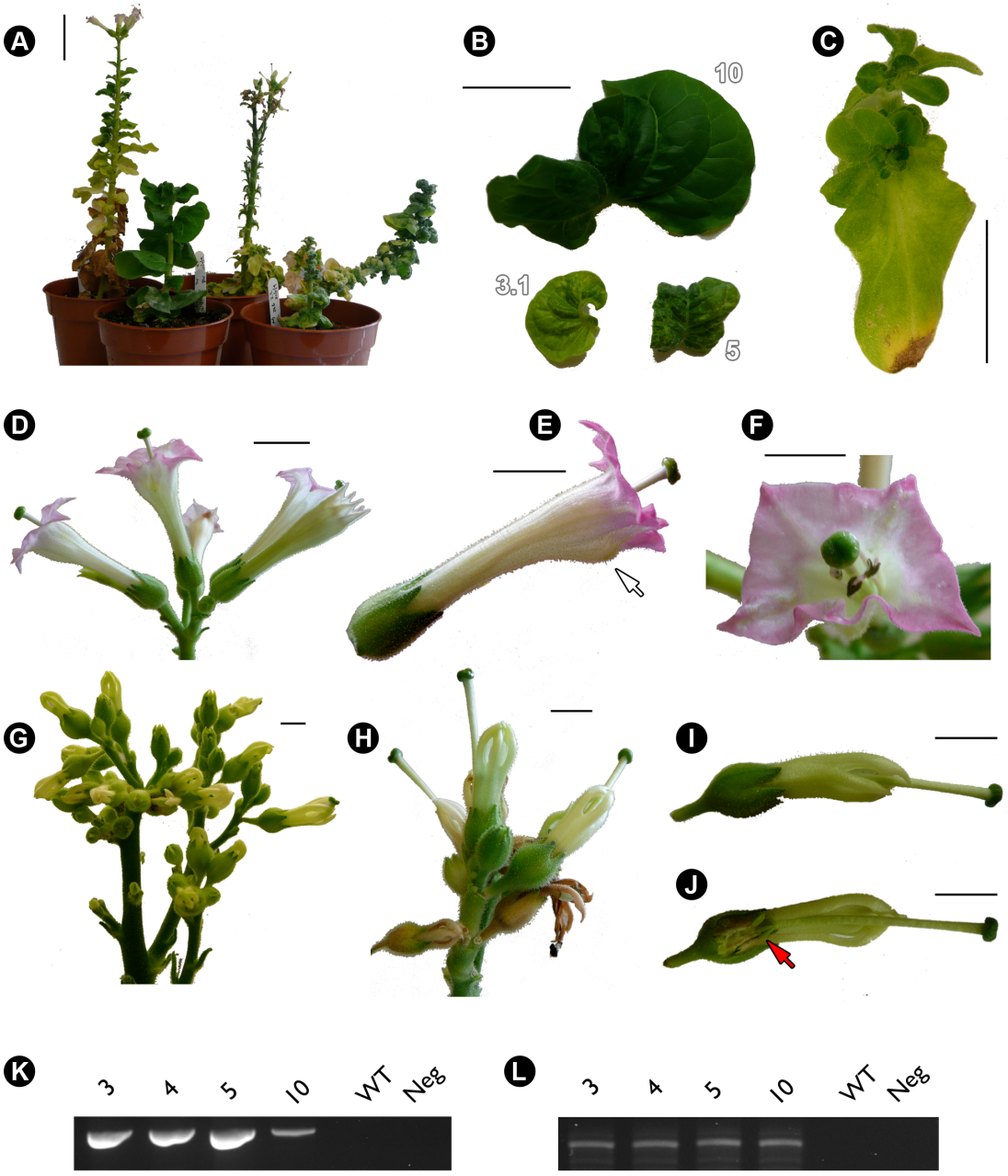


Figure 7.6: Vegetative and floral phenotypes of 35S::*AmINA* transformants
A. 35S::*AmINA* tobacco transformants exhibit a range of vegetative phenotypes. B, C. Leaves are wrinkled and reduced in size with a curved mid-vein (lines 3.1, 5, 10; B) and ectopic shoots on the adaxial leaf surface (line 4; C). D-J. In weak-intermediate lines, e.g. line 3.1, flowers are almost normal (D) but with a reduced corolla tube, ectopic bulges (white arrow; E) and an exserted style (E, F). Severe floral phenotypes include a crowded inflorescence, numerous short-lived buds, many of which are aborted (G, H). Flowers that continue development lack pigment, the petals fail to fuse and the corolla tube (I) and anthers (red arrow in J) fail to elongate. K, L. gDNA (K) and non-quantitative RT-PCR (L) of 35S::*AmINA* transformants using gene specific primers (1,056bp; Appendix B) confirms the presence and expression of the T-DNA, which is absent from wild-type (WT) plants. Scale bars = 5cm in A; 3cm in B, C; 1cm in D-J.

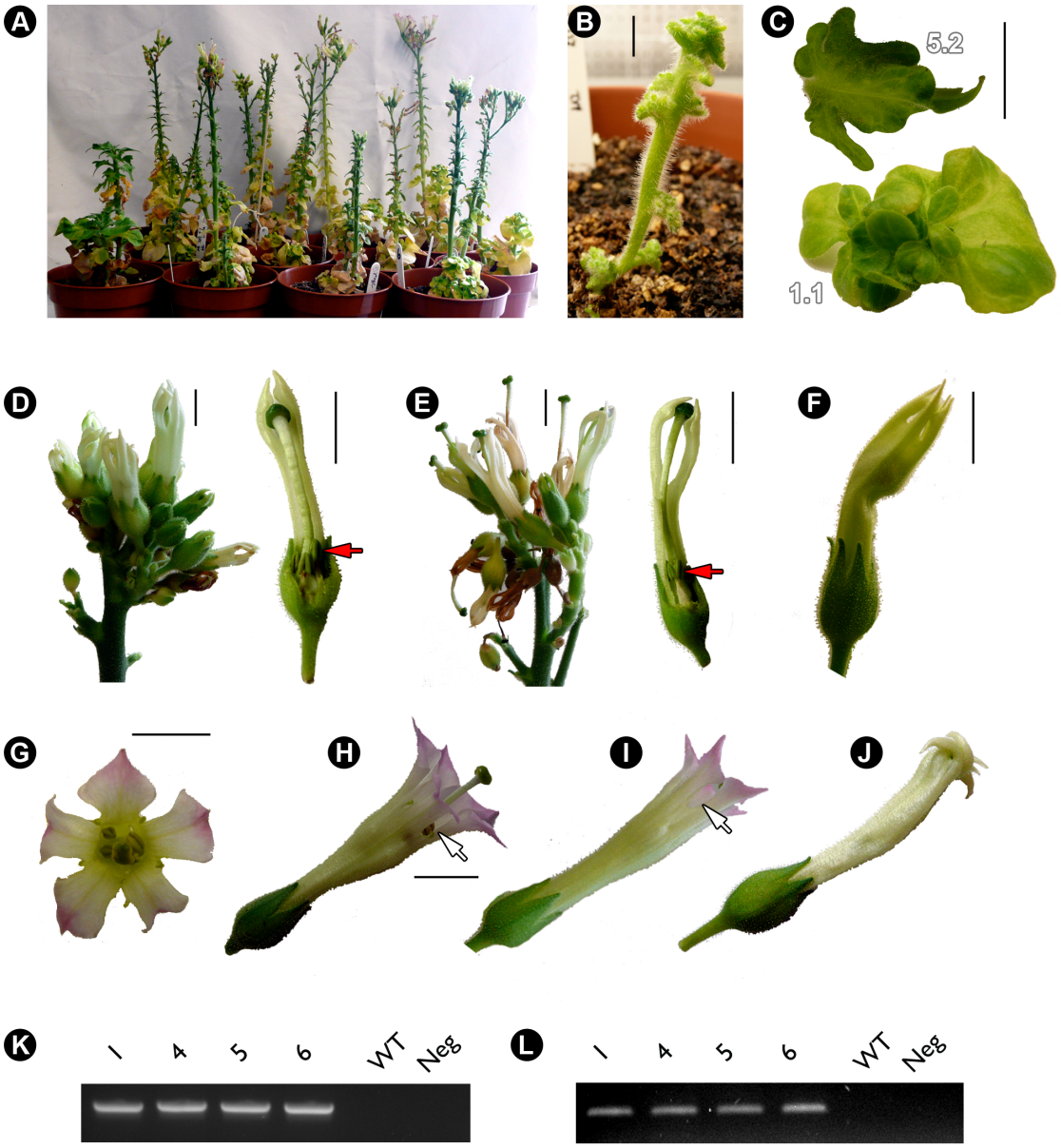


Figure 7.7: Vegetative and floral phenotypes of 35S::*LvHIRZ* transformants
A. 35S::*LvHIRZ* tobacco transformants exhibit a range of vegetative phenotypes. B. In severe lines, e.g. line 5.1, leaves fail to develop and dwarfing is severe. C. Leaves are wrinkled, reduced in size, have disrupted venation, lobing (line 5.2) and ectopic shoots on the adaxial leaf surface (line 1.1; C). D, E, J. In intermediate-severe lines, e.g. lines 1.1 (D), 4.3 (E) and 6.2 (J), inflorescences are crowded with numerous short-lived buds, many of which are aborted (D, E). Flowers that do develop are smaller and narrower than wild-type, lacking in pigment (D, E, J). Whilst the corolla tube does elongate (D, E), the anthers (red arrows in D, E) remain short. F-I. In more weakly affected plants, the corolla tube may be bent, e.g. line 5.2 (F) or dissected (whit arrows in H, I), e.g. lines 5.4 (G), 5.5 (H) and 6.3 (I), with reduced length (H). K, L. gDNA (K) and non-quantitative RT-PCR (L) of 35S::*LvHIRZ* transformants using gene specific primers (1,101bp; Appendix B) confirms the presence and expression of the T-DNA, which is absent from wild-type (WT) plants. Scale bars = 3cm in C; 1cm in B, D-J.

7.3 Discussion

Full-length cDNAs of the class 1 *KNOX* genes *AmHIRZ*, *AmINA* and *LvHIRZ* were introduced into tobacco and constitutively expressed under the CaMV 35S promoter. Introduction of these genes into transgenic tobacco clearly affects the morphology of tobacco shoots, leaves and flowers. Although unique phenotypes such as the ectopic sac-like structures in some 35S::*AmHIRZ* line have been identified, many of the vegetative and floral phenotypes presented in this chapter have been previously described for a broad range of class 1 *KNOX* genes that have been constitutively expressed in *Arabidopsis* and tobacco (Chuck *et al.*, 1996; Kano-Murakami *et al.*, 1993; Matsuoka *et al.*, 1993; Müller *et al.*, 1995; Nishimura *et al.*, 2000; Postma-Haarsma *et al.*, 1999; Sato *et al.*, 1996, 1998; Sinha *et al.*, 1993; Tamaoki *et al.*, 1997). The striking similarity of the *KNOX* constitutive expression lines described here to other published examples, suggests that *AmHIRZ*, *AmINA* and *LvHIRZ* encode typical KNOX proteins that function in a fundamentally similar manner, i.e. maintaining indeterminate cell fate and suppressing differentiation (Endrizzi *et al.*, 1996; Kerstetter *et al.*, 1997; Long *et al.*, 1996; Scofield & Murray, 2006; Vollbrecht *et al.*, 2000), as suggested by bioinformatic analyses that show normal domain structure for each gene (see Section 4.2.3).

7.3.1 Transgenic tobacco phenotypes can be divided into three categories of severity

The resultant aberrant morphology of transformed tobacco plants presented in this chapter can be divided into three phenotypic categories, 1) weak, 2) intermediate and 3) severe, although in practice these categories are arbitrary and phenotypic variation is better represented as a continuum of increasing severity. A number of authors have noted variability in the severity of constitutive *KNOX* expressing tobacco transformants (Kano-Murakami *et al.*, 1993; Matsuoka *et al.*, 1993; Nishimura *et al.*, 2000; Sato *et al.*, 1996, 1998; Tamaoki *et al.*, 1997). In many cases the severity of phenotype is correlated with the expression level of the transgene, e.g. high levels of *KN1* and *OSH1* expression are detectable in the most severely affected tobacco transformants (Kano-Murakami *et al.*, 1993; Sinha *et al.*, 1993).

Although a similar relationship has been identified for other class 1 *KNOX* genes, the relationship between expression level and severity of phenotype is not concrete. For example, constitutive expression of *OSH1* under the tobacco pathogenesis related protein (PR1a) promoter (Kano-Murakami *et al.*, 1993; Sato *et al.*, 1996) generates equally severe phenotypes despite low expression levels. Nishimura *et al.* (2000) suggested that phenotypic severity may also be a property of the *KNOX* gene introduced, demonstrating

that tobacco transformants constitutively expressing endogenous genes such as *NTH15* and *NTH20* tend to have a higher incidence of severe category phenotypes than those expressing *NTH1*, *NTH9* or *NTH22*. Domain exchange experiments suggest that severity of different genes may be related to the formation of an amphipathic structure in the C-terminal KNOX subdomain (Sakamoto *et al.*, 1999). Simple changes in timing of transgene expression may also determine the severity of the phenotype. However, the PR1a promoter drives constitutive expression in the SAM and early leaf primordia only, Despite brief exposure to low levels of *KNOX* gene expression, transgenic tobacco plants expressing *KNOX* genes under the PR1a promoter still develop severe leaf phenotypes.

In the absence of quantitative PCR data pertaining to transgene expression at different stages of development in the 35S::*KNOX* lines presented in this chapter, it is not possible to determine whether expression level and/or timing cause the observed phenotypic variation. However, considerable phenotypic variation is observable in all of the 35S lines presented in this work and in many of the previously published analyses irrespective of the identity of the gene under test. Therefore, although there does appear to be some relationship between phenotype and expression of particular groups of *KNOX* genes, it is more likely that variation in phenotypic severity is not merely the result of the particular gene expressed but is the combined result of all the above factors.

7.3.2 Vegetative aberrations in transgenic tobacco

Vegetative perturbations induced by constitutive expression of *AmHIRZ*, *AmINA* and *LvHIRZ* included small wrinkled/lobed leaves with a thickened leaf lamina and short mid-veins, dwarfism, loss of apical dominance and induction of meristematic identity in determinate organs, i.e. the formation of ectopic shoots on leaves, strongly supporting the fundamental function of KNOX proteins and their interrelationship with phytohormones such as auxin, cytokinin and GA, reviewed by Hay *et al.* (2004); Shani *et al.* (2006). Understanding vegetative phenotypes such as these have been the principal focus of previous investigations of KNOX function. As such, much is known about the morphogenetic processes that give rise to such developmental abnormalities. For example, the overall reduction in size of leaves and internodes has been shown to result from a reduction in cell division consistent with the role of *KNOX* genes in delaying cellular differentiation (Sinha *et al.*, 1993).

The formation of asymmetric leaves with a wrinkled, thickened lamina and curved midvein, has been shown to result from irregularities in leaf vein development which may, in part, reflect a localised decrease in cell division bought about by preferential activity of the CaMV 35S promoter in leaf veins (Jefferson *et al.*, 1987; Schneider *et al.*, 1990; Sinha *et al.*, 1993). In addition, constitutive expression of KNOX proteins disrupts the strict pattern of

anticlinal cell divisions that characterise early tobacco leaf development (Sato *et al.*, 1996). This results in disruption of the leaf mesophyll which becomes characterised by sporadic periclinal divisions in the developing leaf lamina, generating a thickened, irregular leaf blade, and severely disrupts lateral vascular differentiation as lateral veins differentiate from mesenchyma (McHale, 1993). Differences in the frequency of periclinal division in the left-right halves of the leaf disrupts mid-vein leaf symmetry, areas characterised by a high frequency of periclinal cell divisions result in reduced lateral leaf expansion generating mid-vein curvature.

Loss of apical dominance is also well understood. Morphological phenotypes observed in both naturally occurring mutants and transgenic plants with altered *KNOX* gene expression are very similar to plants that exhibit imbalances in plant growth substances such as cytokinins, GA and auxin. *KNOX* GoF phenotypes are also observed in transgenic plants that either overexpress a cytokinin biosynthetic gene or underproduce auxin, and therefore have elevated cytokinin to auxin ratios (Estruch *et al.*, 1991; Li *et al.*, 1992), strongly suggesting that the developmental pathways defined by plant growth regulators and *KNOX* genes are interrelated (Kerstetter *et al.*, 1997; Tsiantis & Langdale, 1998). Tsiantis *et al.* (1999a) examined the patterns of auxin transport in the maize *rough sheath 2* (*rs2*) mutant, which displays ectopic expression of *KNOX* genes due to loss of function of the *RS2* gene that encodes a MYB-like transcription factor that negatively regulates *KNOX* (Timmermans *et al.*, 1999; Tsiantis *et al.*, 1999b). The resulting phenotype included severe dwarfing, this was related to perturbations in auxin gradients both within the leaves and across the vegetative axis, resulting in reduced internode elongation.

Unfortunately, transformants could not be regenerated constitutively expressing *LvINA* due to an inability to promote root initiation. The interrelationship of *KNOX* and phytohormones such as auxins and cytokinins suggests that transgenic 35S::*LvINA* plants may be obtained by modifying hormone levels in shoot regeneration media as the ratios of auxin and cytokinin are critical for the production of shoots or roots in plant tissue culture (Skoog & Miller, 1957).

7.3.3 Floral aberrations in transgenic tobacco

A number of floral irregularities observed in 35S::*AmHIRZ*, 35S::*AmINA* and 35S::*LvHIRZ* lines, such as wrinkling of the corolla tube, increased corolla dissection and reduced corolla tube/anther length, have been reported previously in the literature (Kano-Murakami *et al.*, 1993; Sinha *et al.*, 1993). Interestingly reduced corolla tube length was also noted for the *Hirz-d153* and *Ina-d1* GoF mutants (Golz *et al.*, 2002). Few authors have discussed the significance of such floral phenotypes in relation to the potential function of *KNOX* genes in floral morphogenesis. Floral phenotypes such as reduced corolla tube and anther

filament length may be explained by a reduction in cell division, as observed in the much reduced leaves of transformants (Sinha *et al.*, 1993).

More difficult phenotypes to resolve are the increasingly dissected corolla tube of plants in the intermediate/severe phenotype category. Disruption of cotyledon separation is observed in the LoF *shoot meristemless* mutants of *Arabidopsis* (Chuck *et al.*, 1996), providing a sensible contrast that may implicate *KNOX* genes in developmental pathways related to organ separation. Such a theory has been demonstrated for other *KNOX* genes including *STM* and *KNAT6* in *Arabidopsis*, which are thought to act redundantly to define the boundaries between the SAM and the cotyledons and promote cotyledon separation in combination with the *CUP-SHAPED COTYLEDON CUC1-3* alleles (Aida *et al.*, 1997; Belles-Boix *et al.*, 2006). *KNOX* and *CUC* genes may play a similar role in maintaining petal organ boundaries in non-tubular flowers. Presumably disrupting *KNOX* gene expression in transgenic tobacco has resulted in re-activation of a petal boundary establishment system normally suppressed to form the corolla tube.

In addition to *common* constitutive *KNOX* expression floral phenotypes, the presence of multiple sac-like protrusions on the corolla tube among representatives of the 35S::*AmHIRZ* transformants, and, to a lesser extent the phenotypes of 35S::*AmINA* plants, is remarkable. The development of such structures most likely results from alterations in cell divisions and differentiation and has strong implications in relation to the potential function of *KNOX* genes in floral nectar spur development and evolution.

7.3.4 *KNOX* genes are sufficient to induce novel petal outgrowths

Prior to the reports of ectopic *KNOX* gene expression in relation to ectopic petal tube development in snapdragon (Golz *et al.*, 2002), few authors have paid close attention to floral phenotypes that result from misexpression of *KNOX*. The most intriguing floral phenotype presented in this thesis is the appearance of ectopic sac-like protrusions, several millimetres in size, on the corolla tube of 35S::*AmHIRZ* intermediate phenotype transformants. This phenotype was not observed in any of the other lines, although larger bulges were also prominent on flowers from 35S::*AmINA* lines.

The development of distinct sac-like structures on the corolla tube of many tobacco transformants constitutively expressing *AmHIRZ*, somewhat similar to the knots on the leaves of *KNOTTED1* mutants, has not previously been described in the literature. There are two possibilities to explain this. 1) The lack of attention that has been paid to *KNOX* genes and floral morphogenesis may have resulted in such a phenotype simply being overlooked in prior reports. If this is the case, such a phenotype may indeed be common for a broad range of *KNOX* genes that have been constitutively expressed in tobacco. In

order to answer this question, transgenic tobacco lines previously reported in the literature must be re-examined. Alternatively, 2) the phenotype is genuinely novel, in which case such a phenotype may be peculiar to the AmHIRZ protein.

However, there are few differences in the amino acid sequence of the *AmHIRZ* protein compared to other similar *KNOX* genes such as *LvHIRZ*, *Arabidopsis STM* or the less similar *KN1* gene from maize, to indicate that the phenotype of 35S::*AmHIRZ* trangenic tobacco plants results from specific aspects of the amino acid sequence of AmHIRZ. In addition, the apparent absence of such distinct sac-like structures on flowers of 35S::*AmINA* and 35S::*LvHIRZ* tobacco transformants in particular, suggests that the development of these structures is unlikely to be a specific property of the AmHIRZ protein. If *KNOX* genes such as *AmHIRZ* are involved in nectar spur development, and this seems likely, 35S::*LvHIRZ* transformants should also produce similar petal outgrowths when constitutively expressed in transgenic tobacco, as LvHIRZ is 79% similar to the AmHIRZ protein. It seems likely that the absence of such a phenotype in *AmINA* and *LvHIRZ* transgenic tobacco lines reflects differences in the context of *KNOX* gene expression, in which case it is probable that ectopic petal outgrowths may be a general property of the *KNOX* gene family rather than of the AmHIRZ protein itself. Given that *KNOX* constitutive expression phenotypes observed in vegetative and floral tissues are mostly explained by the induction of meristem-like identities in determinate organs (Postma-Haarsma *et al.*, 1999), it is possible that, given the right conditions, most class 1 *KNOX* genes are capable of inducing similar structures in transgenic tobacco.

Whether a novel or anonymous phenotype, the development of distinct sac-like structures on the corolla tube of 35S::*AmHIRZ* transgenic tobacco plants clearly indicates that, under certain biological conditions, *KNOX* genes are sufficient to induce outgrowth of petal tissue in a novel axis that could sensibly be regarded as a *pre-nectar spur*. This supports earlier claims made by Golz *et al.* (2002) in light of the *Hirz-d153* and *Ina-d1* snapdragon mutants (Golz *et al.*, 2002). Of course, the sac-like structures reported on transgenic tobacco in this chapter are highly dissimilar to the nectar spur of *L. vulgaris*, bearing closer resemblance to the snapdragon gibba in terms of size. However, there is a problem homologising the ectopic outgrowths of 35S::*AmHIRZ* transgenic tobacco and nectar spurs, the sac-like structures characterizing the flowers of transgenic 35S::*AmHIRZ* tobacco plants always form around the fused margins of the petals that constitute the corolla tube. By contrast petal spurs, like those of *L. vulgaris*, always form in the centre of the petal, as such it may be more appropriate to homologise the ectopic sac-like structures of 35S:*AmHIRZ* tobacco plants with the excessive proliferation of tissues at the leaf margins common in transgenic tobacco plants constitutively expressing other *KNOX* genes. Despite these issues, ectopic *KNOX* gene expression appears to be able to con-

fer a meristematic state upon petals. Such a property may have been important in the evolution of petal spurs in Antirrhineae.

7.3.5 Conclusions

Constitutive expression of *AmHIRZ*, *AmINA* and *LvHIRZ* in transgenic tobacco generates a range of vegetative and floral phenotypes that are typical of constitutive class 1 *KNOX* expression characterised for a broad range of *KNOX* genes that, in many ways, phenocopy natural GoF mutants. The result presented in this chapter strongly suggest that these genes have typical functional characteristics of other reported class 1 *KNOX* genes, i.e. regulating the switch from indeterminate to determinate cell fate.

In particular, constitutive *AmHIRZ* expression results in a previously undescribed floral phenotype. The production of ectopic sac-like structures on the corolla tube of tobacco transformants suggest that *KNOX* genes are sufficient to promote outgrowth of petal tissue, providing some support for the assertion that misexpression of class 1 *KNOX* genes may have been an important factor in the evolution and development of floral nectar spurs (Golz *et al.*, 2002). It is unknown whether this phenotype has been ignored in previous analyses of other transgenic tobacco lines reported in the literature and, therefore, whether all class 1 *KNOX* genes have the potential to induce novel petal outgrowths. As phenotypes resulting from constitutive *KNOX* gene expression generally result from induction of meristematic identity in determinate organs, it is highly likely that, under the correct circumstances, most class 1 *KNOX* genes are able to induce similar outgrowths of petal tissue.

However, the floral outgrowths observed in 35S::*AmHIRZ* plants are by no means bona-fide nectar spurs and are not homologous in terms of position. This observation indicates that additional factors are also important in the development of a true nectar spur. In order to better understand the sufficiency of *KNOX* genes to induce nectar spur development, transgenic experiments must be conducted in a more appropriate genetic background. For example, class 1 *KNOX* genes such as *LvHIRZ* must be constitutively expressed, and/or silenced, in *Linaria*. In order to do this reliable transformation protocols must be developed for *Linaria*. Developing TRV-VIGS to silence *LvHIRZ/INA* would be an obvious avenue of future research. Alternatively, further progress may be made by applying what has been learnt from snapdragon and *Linaria* to a more genetically tractable species with floral nectar spurs, such as *Aquilegia*, which is amenable to TRV-VIGS.

Chapter 8

KNOX gene function and expression in the nectar spur of Orchidinae (Orchidaceae)

8.1 Introduction

Phylogenetic (Bateman *et al.*, 2003), morphological and ontogenetic analyses (this study, Chapter 3) have demonstrated that changes in the length of the floral nectar spur may have been a principal factor in the generation of new species in the orchid subtribe Orchidinae (Bateman, 2005; Bateman & DiMichele, 2002; Bell *et al.*, 2009; Box *et al.*, 2008). Furthermore, comparing floral nectar spur ontogeny in ancestrally longer-spurred orchid species, e.g. *Dactylorhiza fuchsii*, to those with shorter spurs, e.g. *D. viridis*, demonstrates that differences in spur length between closely related species are driven principally by shifts in the timing of organ development (heterochrony; Chapter 3; Box *et al.* (2008)). In short-spurred species, such as *D. viridis*, nectar spur ontogeny proceeds at the same rate as that in the longer-spurred orchid *D. fuchsii* but is progenetically curtailed, resulting in a shortened spur representative of earlier stages in the ontogeny of the ancestor (Chapter 3). Bateman (2005); Bateman & DiMichele (2002) suggested that such length differences may have arisen by heterochronic shifts in the timing of expression of a neomorphic gene.

The identification of snapdragon *KNOX* misexpression mutants that cause ectopic petal tube formation, resembling the floral nectar spurs of closely related taxa such as *Linaria vulgaris*, has implicated *KNOX* genes in the evolution and development of floral nectar spurs (Golz *et al.*, 2002). Gene expression and transgenic analyses presented in this thesis support such a role (Chapters 6, 7), suggesting that *KNOX* genes are, to some degree, sufficient for nectar spur development in the wild-type flowers of *L. vulgaris*. Presently, the broader significance of *KNOX* genes in the evolution and development of

floral nectar spurs outside of the eudicot subtribe Antirrhineae (Plantaginaceae), to which snapdragon and *Linaria* belong (Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004), is unknown.

However, morphological analyses of nectar spur ontogeny in a number of different angiosperm taxa, e.g. *Aquilegia* (Tucker & Hodges, 2005), *L. vulgaris*, *D. fuchsii* and *D. viridis* (this study) plus members of the *Dactylorhiza* sister-genus *Gymnadenia* (Box *et al.*, 2008), demonstrate that nectar spur ontogeny is similar in phylogenetically disparate groups. This suggests that, although nectar spurs have evolved multiple times independently among angiosperms (Hodges, 1997; Hodges & Arnold, 1995), *KNOX* gene involvement in *L. vulgaris* nectar spur evolution and development may be more broadly applicable to other spur-bearing angiosperm taxa. As such, *KNOX* genes may well represent the neomorphic alleles that Bateman (2005); Bateman & DiMichele (2002) implicated in spur length variation among closely related members of the subtribe Orchidinae.

8.1.1 Research aims and objectives

The observation that differences in nectar spur length in the closely related orchids *D. fuchsii* and *D. viridis* are related to changes in the timing of nectar spur ontogeny, presents an intriguing opportunity to examine whether changes in the timing of *KNOX* gene expression are related to changes in nectar spur length in *D. fuchsii* and *D. viridis*. More fundamentally, such a question also provides an opportunity to test the broader role that *KNOX* genes may have played in the evolution and development of floral nectar spurs in angiosperms as orchids, compared to *L. vulgaris* (this study), represent a phylogenetically disparate angiosperm group.

To this end, four *KNOX* genes were isolated from medium-spurred *D. fuchsii* (*DfKKN1-4*) and two genes from short-spurred *D. viridis* (*DvKN1* and *DvKN4*, orthologous to *DfKN1* and *DfKN4*; see Chapter 4). Expression of each gene has been assessed by RT-PCR. The functions of genes expressed in the early developing labellum petal, bearing the floral nectar spur primordium, were inferred by observing the phenotype of transgenic tobacco plants expressing the full-length coding sequence of the target *KNOX* gene(s) driven by the constitutive CaMV 35S promoter (see section 2.7).

8.2 Results

8.2.1 Determining the pattern of *KNOX* gene expression in the orchids *Dactylorhiza fuchsii* and *D. viridis*

Gene-specific RT-PCR was used to assess the expression patterns of the novel class 1 *KNOX* genes *DfKN1-4* and *DvKN1, 4* in a variety of vegetative and floral tissues at the earliest stages of spur development. Expression in each case was determined relative to the housekeeping actin gene (*OrACT*) isolated from multiple species in the orchid subtribe Orchidinae. A minimum of three biological and technical replicates were carried out for all PCR based expression analyses. In order to avoid false positive results as a result of gDNA contamination, RT-PCR primers (Appendix B) were designed across a predicted intron sequence and RNA samples treated with DNase. Gene expression was sampled from mature flowers and flowers at the earliest stages of floral nectar spur development (Figure 8.1). Expression in the SAM was not sampled for either *D. fuchsii* or *D. viridis* due to limitations of plant material related to the determinate nature of the orchid SAM, which is lost after flowering, and the inaccessible location of the orchid SAM, which means sampling of the meristem likely results in death of the plant.

8.2.1.1 *D. fuchsii* RT-PCR

Transcript accumulation was assessed by RT-PCR at two levels; (1) vegetative vs floral expression during and after nectar spur initiation and (2) precise floral localisation by RT-PCR analysis of dissected floral organs from early developing floral buds during nectar spur initiation. Although not strictly quantitative, identical quantities of RNA were used to produce cDNA for each of the tissues in which expression was assessed by RT-PCR. Consequently, crude quantitative statements such as *high* and *low* expression are valid when comparing expression of the same gene in different tissues.

(1) *DfKN1, 2* and *DfKN4* transcripts are detectable in mature and developing floral buds, whilst only *DfKN2* and *DfKN4* transcripts can be detected in the leaves. *DfKN3* transcripts are undetectable by RT-PCR in either leaves or floral buds (Figure 8.1). At this level of analysis *DfKN1, 2* and *DfKN4* are all likely candidates for a role in nectar spur development in *D. fuchsii*. Each is expressed to a high level in developing floral buds with early initiating nectar spur primordia. Expression of these three genes continues throughout floral development and is readily detectable in later developing buds but at lower levels, suggesting possible roles in additional areas of floral development.

(2) Developing floral buds with early initiating nectar spur primordia were dissected into several parts to explore precise floral localisation of *DfKN1-4* transcripts. Floral

dissections divided the developing floral bud into the floral bract, the perianth parts (sepals and petals, i.e. tepals), the distinct labellum petal (which bears the nectar spur primordium), the gynostemium (fused androecial and non-ovary gynoecial elements) and the ovary itself containing numerous ovules (Figure 8.1). From this analysis it is clear that floral expression of *DfKN1* is predominantly due to transcript accumulation in the ovary. *DfKN2* transcripts are detectable in the bract, perianth, spur bearing labellum petal, gynostemium and ovary. Despite absence of *DfKN3* expression in analysis 1 (above), *DfKN3* transcripts are readily detectable in the bract, perianth and ovary, whilst *DfKN4* transcripts are barely detectable in the bract and ovary. *DfKN2* is the only class 1 KNOX gene from *D. fuchsii* to be expressed in the spur-bearing labellum petal. Based on this evidence *DfKN2* is the most likely candidate for a role in nectar spur development.

8.2.1.2 *D. viridis* RT-PCR

D. viridis is becoming increasingly rare in the UK (Bateman and Box, personal observation) and is now all but absent from many of the field sites commonly used for previous work on this orchid. For this reason much of the *D. viridis* material was sourced from Austria and grown in the UK. Unfortunately most of the Austrian plants cultivated in the UK for this work failed to flower annually generating an unexpected shortage of suitable material. As a result only a crude analysis of *DvKN1* and *DvKN4* transcript accumulation was possible. Expression of these genes was analysed by gene-specific RT-PCR in vegetative vs. floral tissue, during and after nectar spur initiation. Transcripts from both genes are absent from leaves but are detectable in early developing flowers. Both genes are also readily detectable in later developing floral buds (Figure 8.2). Based on such a crude analysis both genes may be considered candidates for a role in nectar spur development but equally may be involved in all manner of floral developmental processes. In the absence of more precisely defined expression patterns it is impossible to draw firm conclusions from these data.

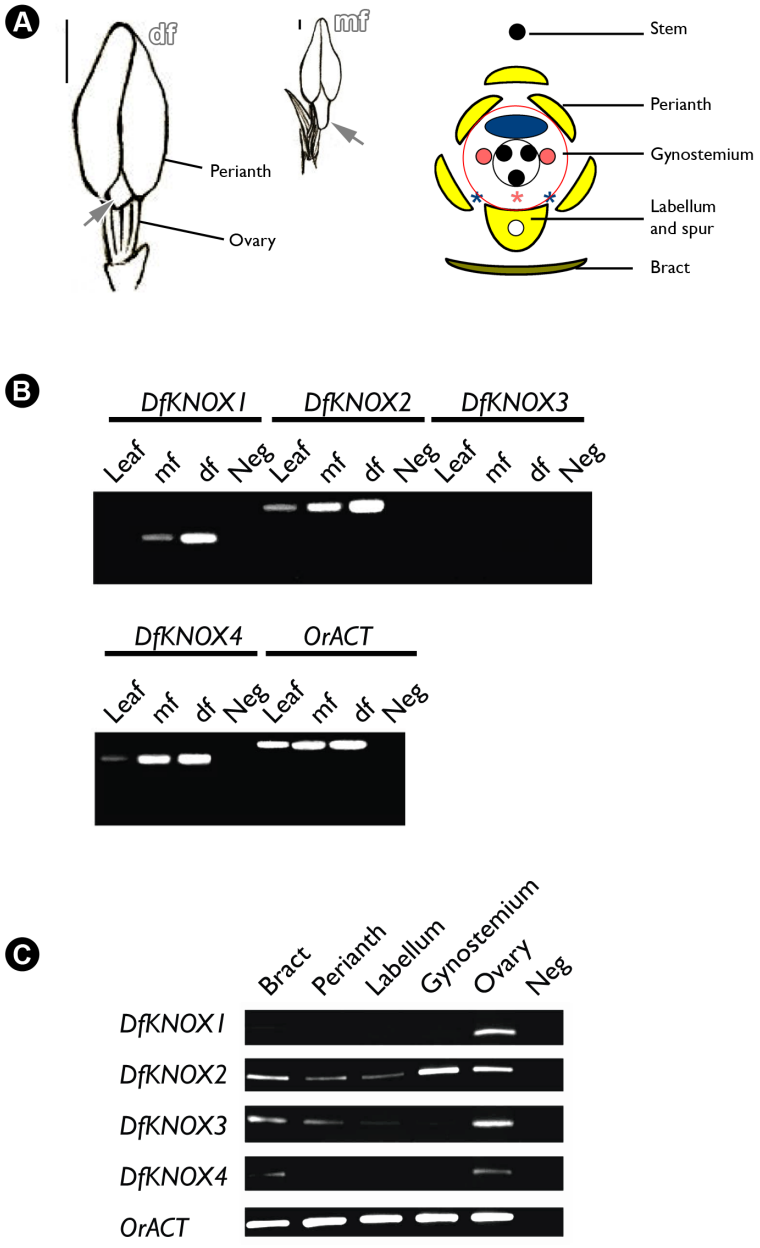


Figure 8.1: Gene-specific RT-PCR analysis of *DfKN1-DfKN4* transcript accumulation in various tissues of the developing floral bud.
Schematic representation of *D. fuchsii* floral dissection and the tissues used in subsequent gene-specific RT-PCR analysis of *DfKN1-DfKN4* relative to the housekeeping gene *OrACT*. A. Floral dissection plan of a *D. fuchsii* flower, arrows indicate developing nectar spurs. Developing (df) and mature (mf; inset) were sampled for RT-PCR. Developing flowers were subsequently dissected and the resultant RNA pools used in RT-PCR (bract, perianth, labellum/spur, gynostemium and ovary; C). B. Vegetative vs. floral gene-specific RT-PCR of *DfKNOX1* (211bp), *DfKNOX2* (359bp), *DfKNOX3* (338bp) and *DfKNOX4* (292bp) relative to *OrACT* (388bp). C. Floral tissue-specific RT-PCR of *DfKNOX1-DfKNOX4* relative to *OrACT*. Note expression of *DfKNOX2* in the developing labellum petal bearing the spur primordium. Scale bars = 50μm.

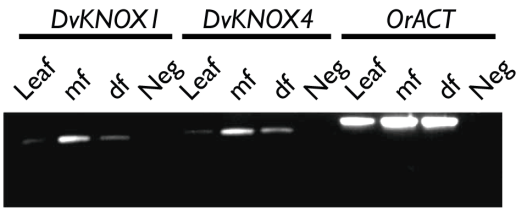


Figure 8.2: Gene-specific RT-PCR analysis of *DvKN1* and *DvKN4* transcript accumulation in vegetative and floral tissues.
Gene-specific RT-PCR of vegetative vs. floral tissues of *DvKN1* (297bp) and *DvKN4* (322bp) relative to *OrACT* (388bp). Expression of both transcripts is detectable in mature and developing flowers.

8.2.2 Constitutive expression of *DfKN2* in tobacco

RT-PCR (see section 8.2.1.2) has identified the *D. fuchsii KNOX* gene *DfKN2* as being expressed in the developing labellum bearing the floral nectar spur primordium (Figure 8.1). The full-length cDNA of *DfKN2* was obtained by RT-PCR and cloned, in sense orientation, into a constitutive expression vector between the CaMV 35S promoter and terminator according to the methods outlined in section 2.7.1 (Figure 8.3). The construct was transferred to *Nicotiana tabacum* cv. Samsun via leaf disc transformation using *Agrobacterium tumefaciens* strain GV3101, using a modified protocol from Horsch *et al.* (1985), (see sections 2.7.3, 2.7.4). Presence of the T-DNA in transgenic plants was confirmed by gDNA PCR and T-DNA expression by RT-PCR using full-length primers specific for the transgene (Appendix B). Vegetative and floral phenotypes were recorded from multiple independent transgenic tobacco lines. Transgenic tobacco phenotypes were compared to wild-type tobacco plants and transgenic tobacco control plants transformed with an unaltered, *empty* 35S construct (see section 7.2.2; Figures 7.3, 8.3).

8.2.2.1 Efficacy of the tobacco transformation protocol

The modified tobacco transformation protocol employed in this work was highly efficient generating comparable results to those described in section 7.2.2.1. Four independent lines were identified constitutively expressing *DfKN2* (35S::*DfKN2* lines 4, 5, 6, 8; Figure 8.3).

8.2.2.2 Vegetative phenotypes of tobacco transformants

Tobacco transformants constitutively expressing *DfKN2* did not differ significantly from wild-type and empty 35S transgenic control plants with respect to vegetative morphology. Each of the four 35S::*DfKN2* lines had normal leaves and grew to an equivalent height to wild-type and control plants (Figure 8.3). The absence of clear leaf phenotypes is highly unusual, especially with respect to the flowers (see section 8.2.2.3).

8.2.2.3 Floral phenotypes of tobacco transformants

In terms of floral morphology representatives from each of the four independent 35S::*DfKN2* lines 4, 5, 6 and 8 barely differed from the flowers of wild-type and empty 35S transgenic control plants (Figure 8.3). All transformants were classed as having weak/intermediate *KNOX* constitutive expression phenotypes. Flowering time of transformants was relatively rapid, occurring within a similar time frame to wild-type tobacco plants. Flowers from plants with weak phenotypes were borne on inflorescences that were wild-type in terms of floral density and branching pattern in all four independent lines. The flowers

themselves were almost wild-type in appearance with a five-lobed, pale pink corolla, e.g. line 5. In some cases, e.g. line 8, the corolla lobes were recurved.

Unusually, the severity of floral phenotypes varied on the same plant. In addition to weak phenotypes, floral traits characteristic of intermediate phenotypes were also observable in some flowers. Intermediate phenotype flowers predominantly affected the corolla lobes (Figure 8.3). The corolla lobes at the proximal region of the corolla tube were more dissected than wild-type, often with a reduced amount of lobe tissue resulting in flowers with a minimal corolla, e.g. line 6 (Figure 8.3C, arrow). Whilst the number of floral organs was not affected in any of the transformants, production of excessive petal tissue at the base of the corolla tube, fusion of neighbouring floral buds and/or reduction in pedicel length appeared to be common phenotypes, e.g. line 10. Where neighbouring buds appeared to be fused, floral organs failed to develop to maturity such that the corolla tube did not form. This extreme phenotype was not observable in every flower on the same plant. In all cases where flowers developed to anthesis, transgenic plants were able to self-fertilise producing copious quantities of seed.

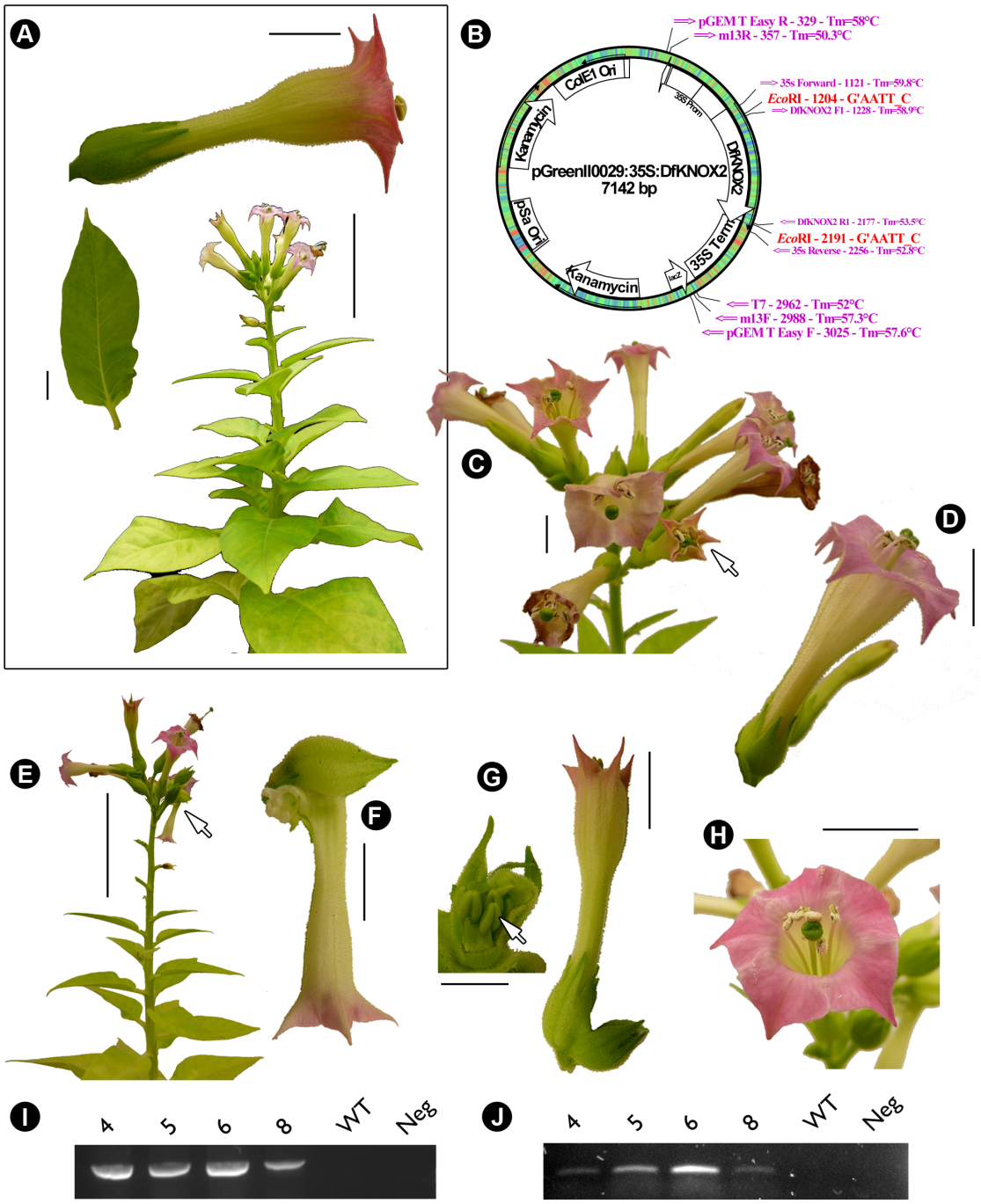


Figure 8.3: Vegetative and floral phenotypes of *35S::DfKN2* transformants, continued over the page...

Figure 8.3: Vegetative and floral phenotypes of *35S::DfKN2* transformants cont.
A. Wild-type tobacco flower, leaf and growth habit (boxed). B. Vector map of the 35S::*DfKN2* construct used in transgenic experiments. Open reading frames are indicated by arrows on the plasmid backbone. Origins, promoters, terminators, operons and antibiotic resistance genes are indicated by block arrows. Primer annealing sites, restriction sites and multiple cloning sites are indicated by tags outside of the DNA backbone. C, D. A range of floral deformities can be observed in the corolla lobe, e.g. reduced corolla lobe tissue (line 6; C, arrow) and recurved corolla lobes (line 8; D). E. Vegetative features of all 35S::*DfKN2* lines were identical to wild-type, e.g. line 10 (E). F, G. Floral aberrations were mild but commonplace, e.g. the production of additional petal tissue generates flowers in which the corolla tube elongates perpendicular to the ovary (arrow, E, enlarged in F). More severely affected flowers are fused to neighbouring buds. Dissection of the minor bud demonstrates the presence of a full complement of floral organs which fail to develop (line 10; arrow in G). H. The most weakly affected lines have flowers that appear like wild-type, e.g. line 5. I, J. gDNA (I) and RT-PCR (J) of 35S::*DfKN2* transformants using gene specific primers (981bp; Appendix B) confirms the presence and expression of the T-DNA, which is absent from wild-type (WT) plants. Scale bars = 5cm (whole plant), 1cm (leaf/flower) in A; 1cm in C-H.

8.3 Discussion

Expression patterns of the *KNOX* genes *DfKN1-4*, *DvKN1* and *DvKN4* from the orchids *D. fuchsii* and *D. viridis* are discussed in relation to *AmHIRZ* and *AmINA* from *Antirrhinum majus*, their *L. vulgaris* orthologs *LvHIRZ* and *LvINA*, and the only other published orchid *KNOX* gene, *DOH1* from *Dendrobium* grex Madame Thong-IN (Yu *et al.*, 2000). Particular emphasis is placed on the potential role of *KNOX* genes in the evolution and development of nectar spurs in Orchidinae.

Based on gene expression analyses, the orchid *KNOX* gene *DfKN2* was selected for constitutive expression analysis using transgenic tobacco. Vegetative and floral phenotypes are discussed in relation to other class 1 *KNOX* genes constitutively expressed in tobacco and the orchid *KNOX* gene *DOH1*, which has been constitutively expressed, and silenced, in transgenic *Dendrobium* grex Madame Thong-IN (Yu *et al.*, 2000).

8.3.1 Orchid *KNOX* genes are expressed in the SAM and flowers

In the case of both *D. fuchsii* and *D. viridis KNOX* genes, expression was not directly demonstrable in the SAM of these plants due to significant tissue limitations. However, the fundamental role demonstrated for class 1 *KNOX* genes in SAM maintenance (Endrizzi *et al.*, 1996; Kerstetter *et al.*, 1997; Long *et al.*, 1996; Vollbrecht *et al.*, 2000) and demonstrable expression of the orchid *KNOX* gene *DOH1* from *Dendrobium* grex Madame Thong-IN in the SAM (Yu *et al.*, 2000), makes it highly likely that one or more of the *D. fuchsii*/*D. viridis* genes are also expressed in the SAM. Further gene expression analyses are required to establish meristematic patterns of *D. fuchsii*/*D. viridis KNOX* gene expression.

RT-PCR demonstrates that transcript accumulation of the orchid *KNOX* genes *DfKN1-4* from *D. fuchsii* is clearly observable in mature and developing flowers. In developing tissues transcripts are detectable in leaf tissue (*DfKN2, 4*), the perianth (*DfKN2, 3*), the enlarged labellum petal, bearing the nectar spur primordium (*DfKN2*), and the reproductive tissues such as the gynostemium (*DfKN2*) and the ovary (*DfKN1-4*). Likewise, *DvKN1* and *DvKN4* transcripts are detectable in mature and developing *D. viridis* flowers. However, shortage of tissue for this species prevented precise localisation of transcript accumulation by RT-PCR analyses using organ specific pools of cDNA.

The broad patterns of extra-meristematic expression reported for *DOH1* in the vegetative apices, transitional buds, floral buds, provascular strands of leaf primordia, inflorescence meristems and floral primordia (Yu *et al.*, 2000) when combined with observations of extra-meristematic expression of *D. fuchsii*/*D. viridis KNOX* genes, further strength-

ens the idea that *KNOX* gene expression is not merely restricted to meristematic tissues such as the SAM, but that *KNOX* genes are also likely to play crucial roles in additional aspects of plant development, such as complex leaf morphogenesis (Hareven *et al.*, 1996; Hay & Tsiantis, 2006; Shani *et al.*, 2009), carpel development (Endrizzi *et al.*, 1996; Foster *et al.*, 1999; Pautot *et al.*, 2001; Scofield *et al.*, 2007) and nectar spur ontogenesis (Golz *et al.*, 2002).

8.3.2 Floral expression of *DfKN2* resembles that of *L. vulgaris HIRZ/INA* and the pattern observed for *AmHIRZ/AmINA* in snapdragon mutants: implications for a broader role of *KNOX* genes in nectar spur evolution and development

Despite a probable wild-type role in maintaining the SAM, transcripts of *DfKN1-4* have much broader patterns of expression than other well characterised class 1 *KNOX* genes from eudicots and monocots. In particular the *D. fuchsii KNOX* gene *DfKN2* shows particularly broad patterns of floral expression, strongly suggesting additional roles for this gene in one or more processes related to floral development.

The broad expression pattern of *DfKN2* resembles that of *AmHIRZ/INA* in snapdragon mutants (Golz *et al.*, 2002) and their *L. vulgaris* orthologs, *LvHIRZ/INA* (this study; 6). In the mutant snapdragon lines, ectopic expression of the class 1 *KNOX* genes *AmHIRZ* and *AmINA* outside of the SAM in petals and leaves result in a range of pleiotropic phenotypes including altered leaf shape, excessive trichome formation and, most notably, an additional ectopic petal tube on the ventral part of the corolla, morphologically similar to the spurs of close relatives of snapdragon. In *L vulgaris*, *KNOX* expression is largely concentrated in the petals, and has been related to the development of the floral nectar spur which is morphologically similar to the ectopic petal tubes induced in the *Hirz-d153* and *Ina-d1* snapdragon mutants.

Expression of *DfKN2* in the developing labellum petal, carrying the nectar spur primordium, indicates that *KNOX* genes such as *DfKN2* may be involved in the morphogenesis of the floral nectar spur in orchids as well as in eudicots such as *L. vulgaris*. Whilst nectar spurs have evolved multiple times independently in divergent taxa, the data presented in this chapter suggest that, at least in cases where nectar spur development is morphologically similar such as *D. fuchsii* and *L. vulgaris* (Chapter 3), it is feasible that monocot and eudicot taxa may utilise a similar developmental-genetic pathway regulated by *KNOX* genes.

However, whilst *DfKN2* is the only *D. fuchsii KNOX* gene to be expressed in the labellum, it is also expressed in all other parts of the flower. The class 1 *KNOX* gene *DOH1*

exhibits similarly broad patterns of floral expression. Expression of *DOH1* is detectable in the SAM and inflorescence meristem but also in leaf and developing floral primordia, particularly the petals and gynostemium (Yu *et al.*, 2000). Unfortunately, Yu *et al.* (2000) failed to discuss the relevance of floral *DOH1* expression. *Dendrobium* grex Madame Thong-In does not possess floral nectar spurs, strongly indicating that *KNOX* genes are important in additional aspects of floral organogenesis. However, *Dendrobium spp.* routinely possess spurs, albeit rather unsophisticated (conical; short and wide-mouthed) structures that form at the junction of the labellum and gynostemium, investigating *KNOX* expression in spurred *Dendrobium spp.* would be very interesting. Expression of *DOH1* in the gynostemium is of particular interest (see section 8.3.3). Whilst the presence of *DfKN2* expression in the same spatio-temporal pattern as nectar spur ontogeny does not imply causation, expression of *DfKN2* in additional floral organ parts does not preclude *DfKN2* from any possible involvement in nectar spur ontogeny.

Based solely on expression data one may expect that *DfKN2* is phylogenetically similar to *LvHIRZ/LvINA*. However, *DfKN2* is phylogenetically distinct from *L. vulgaris* and snapdragon *HIRZ* and *INA*, but similar to *OSH6* from rice. *Df/DvKN1* are in fact most closely related to *HIRZ* and *INA* from *L. vulgaris* and snapdragon, however *DfKN1* appears to be one of the least broadly expressed *D. fuchsii KNOX* genes. Likewise, in *Dendrobium*, *DOH1* is expressed in the petals sepals and gynostemium (Yu *et al.*, 2000), however the phylogenetically similar gene *DfKN4*, has a relatively restricted pattern of floral expression in the ovary.

The RT-PCR results presented in this chapter strongly support the assertion presented in Chapter 4 that gene function/expression patterns cannot be predicted from phylogeny. In addition, these results also suggest something remarkable. Although both orchid and Antirrhineae nectar spur development both seem to utilise a similar developmental-genetic pathway regulated by *KNOX* genes, the specific *KNOX* genes utilised need not be orthologous. Indeed, the basic function proposed for class 1 *KNOX* genes in regulating the switch from indeterminate to determinate development may explain the likely co-option of similar developmental-genetic pathways for nectar spur development over significant phylogenetic distances. In addition to nectar spur development, such properties may be fundamental to a variety of different developmental processes, e.g. carpel development.

8.3.3 Orchid *KNOX* genes may play a role in carpel development

Whilst few authors have paid close attention to the role of *KNOX* genes in floral development, some authors have highlighted a potential role in carpel initiation. Constitutive expression and/or silencing of *Arabidopsis STM* and *KNAT2* results in severely disrupted

carpel and ovule formation (Endrizzi *et al.*, 1996; Pautot *et al.*, 2001; Scofield *et al.*, 2007, 2008). Expression data presented in this thesis also indicate that *KNOX* genes are expressed in the carpels of developing *L. vulgaris* flowers (see Chapter 6). Likewise, the *Dendrobium KNOX* gene *DOH1* is expressed in the carpel-derived tissues of the gynostemium (Yu *et al.*, 2000). Interestingly, all four of the *KNOX* genes identified from *Dactylorhiza fuchsii* are expressed in the developing ovary (Figure 8.1), indicating that *KNOX* genes may play a broad role in the development of the angiosperm carpel.

As carpels are formed from a residual population of stem cells located at the centre of the floral meristem (FM) (reviewed in Blázquez *et al.* (2006)), it is perhaps, not surprising that *KNOX* gene expression may be detectable in developing ovaries. In fact, Scofield *et al.* (2007, 2008) have demonstrated that *Arabidopsis STM* functions to maintain this stem cell niche until carpel development is initiated. However, expression of *KNOX* in orchid carpel tissues is detectable long after floral transition and carpel initiation, suggesting that *KNOX* genes play a more direct role in carpel/ovary/locule development. Indeed, Pautot *et al.* (2001); Scofield *et al.* (2007) have shown that in addition to a somewhat classical role in maintaining the stem cell niche that gives rise to the carpel tissues, *Arabidopsis STM* can also directly promote the development of carpels and the associated meristematic placental tissues of the ovary, independently of *LEAFY* (*LFY*) and *AGAMOUS* (*AG*), which are normally required to terminate stem cell maintenance and permit development of reproductive tissues (Busch *et al.*, 1999; Lenhard *et al.*, 2001; Lohmann *et al.*, 2001).

In the absence of *in situ* localisation of the *D. fuchsii KNOX* genes, it is difficult to determine a precise role for these genes in carpel development. However, the expression of *KNOX* in the carpels of such a diverse range of monocots and eudicots strongly supports a general role for *KNOX* genes in the development of reproductive tissues. Precise gene expression/functional data for the *D. fuchsii* and *D. viridis KNOX* genes, amongst others from *Arabidopsis* and *Linaria*, is required to provide a convincing demonstration for a fundamental role of *KNOX* genes in the development of carpels.

8.3.4 Constitutive expression of *DfKN2* in transgenic tobacco

The only orchid *KNOX* gene that has been characterised by constitutive expression is *DOH1*, which was studied endogenously in *Dendrobium* grex Madame Thong-IN (Yu *et al.*, 2000). Constitutive expression of *DOH1* in orchid plants completely suppressed shoot organization and development. Transgenic orchid plants expressing antisense mRNA for *DOH1* generated multiple ectopic SAMs and exhibited early flowering phenotypes. Both the sense and anti-sense transformants showed defects in leaf development somewhat typical of those observed along other monocots. These findings strongly suggest that

KNOX genes play a key role in maintaining the basic plant architecture in much the same manner as those studied in other monocots and eudicots.

Given the phenotypes observed in *Dendrobium* grex Madame Thong-IN transformants (Yu *et al.*, 2000), it was highly likely that constitutive expression of *DfKN2* in transgenic tobacco would produce transformants with significantly altered vegetative traits. Transgenic tobacco plants with weak constitutive *KNOX* expression phenotypes typically alter vegetative traits, generating transformants with subtly wrinkled leaves but wild-type flowers (Kano-Murakami *et al.*, 1993; Matsuoka *et al.*, 1993; Nishimura *et al.*, 2000; Sato *et al.*, 1996; Sinha *et al.*, 1993; Tamaoki *et al.*, 1997). However, the phenotypes observed in 35S::*DfKN2* transgenic tobacco plants were unusual in that they predominantly affected floral rather than vegetative traits. Vegetatively, 35S::*DfKN2* plants were wild-type, but the flowers exhibited features that are more characteristic of 35S::*KNOX* tobacco plants with intermediate phenotypes. 35S::*DfKN2* plants have flowers with dissected corolla lobes, often with a reduced amount of lobe tissue resulting in flowers with a minimal corolla, produce excessive petal tissue at the base of the corolla tube and bear flowers that are sometimes fused to neighbouring floral buds.

The sac-like outgrowths that characterise 35S::*AmHIRZ* tobacco plants (Chapter 7), were not observed in 35S::*DfKN2* transgenic tobacco lines. It is possible that the absence of phenotype reflects the fact that DfKN2 is not phylogenetically close to AmHIRZ, this may not be surprising if the sac-like protrusions generated in 35S::*AmHIRZ* plants are a property of the AmHIRZ protein. However, phylogeny cannot be used as a tool to infer function in the *KNOX* gene family (see Chapter 4) and, furthermore, exploration of AmHIRZ function in Chapter 7 seemed to suggest that the sac-like protrusions of 35S::*AmHIRZ* transformants were unlikely to reflect a specific property of the AmHIRZ protein. Therefore, the absence of sac-like protrusions in the flowers of 35S::*DfKN2* transformants is more likely to reflect subtle differences in the timing and/or location of transgene expression in 35S::*DfKN2* transformants rather than aspects of the protein itself. However, the predominance of the floral phenotype in 35S::*DfKN2* plants may reflect specific properties of the DfKN2 protein.

The predominantly floral nature of 35S::*DfKN2* phenotypes can be explained in at least two different ways. 1) That *DfKN2* primarily interacts with floral targets in transgenic tobacco, reflecting a floral role for *DfKN2* in the orchid *D. fuchsii*. Alternatively, 2) specific aspects of the timing and/or pattern of *DfKN2* expression in transgenic tobacco plants, perhaps related to positional effects of transgene integration, may have generated an artifactual phenotype that does not indicate any specific role for *DfKN2* in floral development. Positional effects seem an unlikely explanation as the predominance of floral versus vegetative phenotypes was apparent in all four independent 35S::*DfKN2* lines.

More likely is that the unusual predominance of floral phenotypes is related to specific aspects of the timing and/or location of transgene expression. The CaMV 35S promoter has been shown to drive expression of transgenes in all tissues from an early developmental stage (Benfey *et al.*, 1989; Harpster *et al.*, 1988). Based on this fact, leaf tissue was used to screen for expression of the *DfKN2* transgene long before the transition to flowering had occurred, strongly indicating that the floral specific phenotype of 35S::*DfKN2* transformants is not related to floral-specific expression. It remains possible, however, that the *DfKN2* transgene is more highly expressed in floral tissues. Quantifying *DfKN2* expression in tobacco transformants is required to determine if this was the case.

In the absence of quantitative data pertaining to *DfKN2* expression, the unusual floral-specific phenotype of 35S::*DfKN2* transformants suggests that the DfKN2 protein predominantly interacts with floral factors in transgenic tobacco, reflecting a predominantly floral function for *DfKN2* in *D. fuchsii*. Such an explanation is consistent with the fact that *DfKN2* transcripts are readily detectable in a variety of floral tissues in *D. fuchsii* itself. These data may reflect a role for *DfKN2* in ovary and nectar spur development. In order to more fully understand the role of DfKN2 in floral development it is necessary to study the function of this protein in a more appropriate genetic background. Tobacco and orchids are phylogenetically distant. Transformation of *D. fuchsii* would be ideal, however few orchids are amenable to stable transformation, e.g. *Dendrobium* grex Madame Thong-In (Yu *et al.*, 2000), *Phalaenopsis* (Belarmino & Mii, 2000; Semiarti *et al.*, 2007) and *Oncidium* (Liau *et al.*, 2003), but are superficially highly variable in floral morphology and do not possess floral nectar spurs, although some species of *Dendorbium* do. The only other monocot species that develop floral nectar spurs are in the genus *Tricyrtis* (Rudall, pers comm.), which is reportedly amenable to stable genetic transformation using *A. tumefaciens* (Adachi *et al.*, 2004). Constitutive expression of *DfKN2* in *Tricyrtis* may be more informative with regard to nectar spur development. Investigating the function of DfKN1, 3 and 4 using similar techniques may also be fruitful, particularly in relation to their potential role in ovary development, as indicated by gene expression analyses.

8.3.5 How might differences in spur length between *D. fuchsii* and *D. viridis* be explained at the genetic level?

Bateman & DiMichele (2002) suggested that differences in nectar spur length of *D. fuchsii* and *D. viridis* are driven by mutation(s) of a neomorphic allele that altered the timing of its expression, resulting in the shifts in timing of nectar spur development that have been observed ontogenetically in this work (Chapter 3). The results presented here for *D. fuchsii* (and *L. vulgaris*, Chapters 6, 7), suggest that *KNOX* genes may well represent

the neomorphic alleles that Bateman (2005); Bateman & DiMichele (2002) hypothesized. Unfortunately, precise spatio-temporal patterns of expression are not available for *D. fuchsii* and *D. viridis*, making it difficult to draw any conclusions pertaining to the timing, duration and location of *KNOX* gene expression. Therefore it remains possible that changes in the timing/duration of *KNOX* expression may be a sufficient explanation for the morphological differences in spur length observed between *D. fuchsii* and *D. viridis*.

However, there is an additional theory that should also be explored. RT-PCR has demonstrated that *DfKN2* is the most likely candidate to play a role in nectar spur development in *D. fuchsii*. Assuming that the closely related nature of *D. fuchsii* and *D. viridis* would mean that the same *KNOX* gene was responsible for spur development in each species, the *D. viridis* ortholog of *DfKN2* could not be obtained despite extensive sampling. As such, it is also possible that the absence of this gene may reflect the difference in the spur lengths of *D. fuchsii* and *D. viridis*. Based on what is currently hypothesised regarding the role of *KNOX* genes in nectar spur development, i.e. that somehow *KNOX* genes act as proximal-distal organisers of petal growth (Golz *et al.*, 2002), or that *KNOX* expression delays petal differentiation (this study), the association between loss of *DvKN2* and nectar spur reduction is unlikely. It is more likely that the absence of *DvKN2* may be explained by failure to obtain the gene, rather than an evolutionary loss of developmental significance. In the event that *DvKN2* has been lost however, it is plausible that this would result in the absence of a nectar spur in accordance with the hypothesised role for *KNOX* proposed by Golz *et al.* (2002). If *KNOX* genes delay differentiation in petals (this study, Chapter 9), it is possible that loss of *KNOX* expression may result in a short spur by preventing elaboration of a spur development pathway that would otherwise permit the development of a longer spur. Confirmation of any such speculation would require significant further experimentation.

8.3.6 Conclusions

The orchid *KNOX* gene *DfKN2* is the most promising candidate for a role in orchid nectar spur development. *DfKN2* is expressed in the labellum petal, which also bears the nectar spur primordium. However, *DfKN2* transcripts can also be detected in other developing floral organs, most notably the ovary. In fact, expression of all four *D. fuchsii KNOX* genes is detectable in the ovary, strongly supporting results obtained from *Arabidopsis* (Scofield *et al.*, 2007) and *L. vulgaris* (Chapter 6, this study), that suggest a potential role for *KNOX* genes in carpel development.

Interestingly, both *DvKN1* and *DvKN4* are also expressed in mature and developing flowers, further supporting the idea that *KNOX* genes are important in floral development. Unfortunately, precise spatio-temporal patterns of expression are not presently available.

This makes it impossible to determine whether changes in nectar spur length between *D. fuchsii* and *D. viridis* are related to alterations in the timing of *KNOX* gene expression. In order to further clarify the role of *KNOX* genes in the evolution and development of orchid nectar spurs/carpel, precise spatio-temporal patterns of expression must be established for *D. fuchsii* and *D. viridis* using *in situ* hybridisation.

Constitutive expression of *DfKN2* in transgenic tobacco generates unusual plants with a mild/intermediate severity of floral phenotype, but with apparently normal vegetative traits. Typically, constitutive expression of *KNOX* genes in tobacco generates a vegetative phenotype before any detectable changes in floral morphology. The tendency for *DfKN2* to alter floral morphology in transgenic tobacco suggests that the DfKN2 protein may interact specifically with other floral factors in tobacco, indicating that DfKN2 may play a wild-type role in floral development in accordance with endogenous floral expression in *D. fuchsii*.

Although tobacco has been used to successfully understand the function of monocot *KNOX* genes such as *KN1*, it should be borne in mind that tobacco is by no means a close relative to *D. fuchsii*. To date, no orchid *KNOX* genes have been analysed using transgenic tobacco, making it difficult to interpret any precise functional data from 35S::*DfKN2* transgenic tobacco. In light of this fact it may be appropriate to use a more closely related host plant, e.g. *Tricyrtis*.

Chapter 9

General discussion

9.1 Towards a model of *KNOX* gene involvement in nectar spur development

Nectar spurs are a pervasive feature of angiosperms and have evolved independently in a large variety of flowering plant groups including both monocots (e.g. Orchidaceae and Liliaceae) and eudicots (e.g. Ranunculaceae, Plantaginaceae and Lentibulariaceae) (Endress & Matthews, 2006). The nectar spur is a structure of acknowledged function(s) that is known to influence speciation in a broad range of angiosperm taxa via simple differences in length, shape, orientation, colouration and even curvature. It is unusually evolutionarily labile (in terms of both acquisition and loss), commonly subjected to strong directional and/or disruptive selection, and is frequently discussed in the literature as a key innovation that has played a critical role in determining the high species-level diversity of spur-bearing taxa (Hodges, 1997; Hodges & Arnold, 1995; Hodges *et al.*, 2002). Further understanding the evolutionary significance of nectar spurs requires a fundamental knowledge of the developmental genetic mechanisms that determine nectar spur characteristics.

A number of key experiments have demonstrated that nectar spur development is influenced by canonical ABC genes and factors that determine floral zygomorphy, however it is only recently that clues have emerged that promise to identify additional genetic factors related to nectar spur development. Snapdragon is, ordinarily, spurless (Sutton, 1988). The discovery of snapdragon *KNOX* GoF mutants with ectopic petal tubes, resembling the nectar spurs of closely related taxa such as *Linaria vulgaris* (Golz *et al.*, 2002), has highlighted a potential role for *KNOX* genes in nectar spur development. Many authors have called for investigations of *KNOX* gene activity in the spur-bearing close relative, *L. vulgaris* (Damerval & Nadot, 2007; Galego & Almeida, 2007; Whitney & Glover, 2007).

In this thesis, the role of *KNOX* genes in spur development has been investigated in

L. vulgaris and in the orchids *D. fuchsii* and *D. viridis* in an evolutionary-developmental context. Exploring the role of *KNOX* genes in these taxa has significantly expanded upon the foundations provided by Golz *et al.* (2002), and has established several key insights into the development of floral nectar spurs in a taxon-specific setting. Whilst the specific results presented in this thesis have been discussed in detail in their respective chapters, the aim of this chapter is to explore the general principles relating to nectar spur development that have emerged from this work, and to present a model postulating a plausible mechanistic role for *KNOX* genes in nectar spur development.

9.1.1 Major insights relevant to nectar spur development

A number of key insights into nectar spur development have been gleaned from the work conducted in this thesis:

Nectar spur ontogeny is fundamentally similar over broad phylogenetic distances Despite the extensive differences in floral morphology and the considerable phylogenetic distance that separates the eudicot *L. vulgaris* from the orchids *D. fuchsii* and *D. viridis*, morphological assessment of the floral nectar spurs in these taxa has revealed surprising similarities. In all species, the floral nectar spur is a tubular structure located at the base of the ventral petal that starts development as an abaxial bulge initiated very late in petal ontogeny. Closer examination of the cells that comprise the nectar spur primordium revealed that early nectar spur initiation is characterised by a mound of small, isodiametric cells that are located at the base of the petal, indicating that early nectar spur ontogeny involves cell division in a novel axis. Later nectar spur ontogeny, i.e. spur elongation, seems to be predominantly characterised by longitudinal cell enlargement as cell dimensions increase dramatically along their longitudinal axis throughout the entire length of the nectar spur. These observations suggest that nectar spur length may be determined by different extents of longitudinal cell enlargement during later spur ontogeny, and/or differential cell division during early nectar spur initiation.

Whilst similar ontogeny does not reflect any evolutionary relationship between these taxa, which are separated by a considerable phylogenetic gulf, it does present the intriguing possibility that nectar spurs in these divergent taxa have independently evolved through a similar spur development pathway. Ontogenetic analyses of the petal-derived nectar spurs of *Aquilegia formosa* and *A. olympica* (Ranunculaceae) have highlighted similar results (Gottlieb, 1984; Tepfer, 1953; Tucker & Hodges, 2005), further indicating that, at least for simple petal-derived nectar spurs, the same spur developmental pathway has evolved independently on multiple occasions in angiosperms. However, simple petal-derived nectar spurs do not represent the true diversity of nectar spur morphologies, for

example the receptacle-derived nectar spurs of *Tropaeolum majus* (Tropaeolaceae) may have evolved by different means. The fundamental similarity of nectar spur ontogeny among members of Plantaginaceae, Ranunculaceae and Orchidaceae suggests that nectar spur evolution and development has fundamental requirements that may implicate members of the same gene family in each case, e.g. *KNOX* genes (Golz *et al.*, 2002), or members of other gene families.

Variation in nectar spur length is likely to involve heterochrony Comparative floral ontogeny in the orchids *D. fuchsii* and *D. viridis* has shown that changes in the timing of nectar spur development are responsible for the length-differences apparent between the medium-spurred flowers of *D. fuchsii* and the short-spurred flowers of *D. viridis*. In *D. viridis* development of the floral nectar spur is progenetically curtailed, producing a flower with a short spur at anthesis that resembles the spur at earlier stages of ontogeny in the longer-spurred ancestor, represented by the flowers of *D. fuchsii*. Similar results have been established for short-spurred members of the sister orchid genus *Gymnadenia* (Box *et al.*, 2008). Whilst the precise functional significance of changes in nectar spur length is debatable in the specific case presented here for *D. fuchsii* and *D. viridis*, changes in nectar spur length are largely thought to influence speciation by disrupting plant-pollinator interactions and promoting reproductive isolation (Hodges & Arnold, 1995).

In the absence of strong phylogenetic data to determine the evolutionary relationship between the spur of *L. vulgaris* and the gibba of *A. majus*, one can only speculate as to the importance of heterochrony in generating the morphological differences that characterise these two homologous organs. However, the principal difference between the spur and gibba in Antirrhineae is length. It is highly likely therefore that morphological differences between the gibba and spur result from subtle changes in the timing of organ development. This may also be true in generating the vast diversity of nectar spur lengths among other angiosperm taxa.

***KNOX* gene function cannot be predicted from phylogeny** Novel *KNOX* genes were identified from *L. vulgaris* (*LvHIRZ* and *LvINA*), *D. fuchsii* (*DfKN1-4*) and *D. viridis* (*DvKN1, 4*). The probable function of each of these genes was explored in a phylogenetic context. Whilst phylogenetic treatments established that *LvHIRZ* and *LvINA* are the orthologs of *HIRZ* and *INA* from *A. majus*, and that the genes *Df/DvKN1* belong to the *STM* group, *DfKN2* the *OSH6* group, *DfKN3* the *KNAT1* group and *Df/DvKN4* the *KNAT2* group, the most striking result was that phylogeny is not able to predict the function of *KNOX* genes. The same ortholog from two different species may be involved in the same process, commonly maintenance of the SAM, or may have evolved completely different functions following gene duplication, such as complex leaf de-

velopment. Often paralogous *KNOX* genes can share considerable functional redundancy (Byrne *et al.*, 2002), this can result in stochastic retention of ancestral gene function and co-option for a range of diverse developmental roles. Furthermore, the absence of any pattern relating *KNOX* gene function and phylogeny, strongly indicates that whatever developmental processes *KNOX* genes are involved in, whether maintenance of the SAM or complex leaf morphogenesis, these processes must require a fundamental, rather than specific, function of *KNOX* genes. This discovery provides valuable insight into the putative mechanism of *KNOX* gene action in nectar spur development and suggests that *KNOX* gene activity in the spur may be fundamentally similar to that in the SAM and compound leaf.

***KNOX* genes are expressed in the perianth of *L. vulgaris* and orchids** Remarkably, floral expression of class 1 *KNOX* genes is readily detectable in *L. vulgaris* and the orchids *D. fuchsii* and *D. viridis*. Most remarkable of all is that wild-type expression of *LvHIRZ* in the flowers of *L. vulgaris* is almost identical to that of *AmHIRZ* in the *Hirz-d153* snapdragon mutants (Golz *et al.*, 2002). Expression of *LvHIRZ* is predominantly focused in the ventral petal, which bears the floral nectar spur primordium. Likewise, in the orchid *D. fuchsii*, expression of the class 1 *KNOX* gene *DfKN2* is correlated with the development of the ventral labellum petal and the associated floral nectar spur primordium. Observation of class 1 *KNOX* expression in *L. vulgaris* and *D. fuchsii* clearly indicates that *KNOX* gene expression coincides spatially and temporally with nectar spur development, adding considerable weight to the theory that *KNOX* gene expression is a critical component of nectar spur development. However, *KNOX* genes in *L. vulgaris* and *D. fuchsii* are also expressed more broadly in other floral tissues, particularly the ovary. While these data provide some additional evidence for the hypothesis that *KNOX* genes are important in carpel development, it also indicates that *KNOX* gene expression is not the sole contributor to floral nectar spur development. Clearly, additional factors such as floral zygomorphy-determining factors and ABC genes, are critical to nectar spur development, as suggested by previous work (Cubas *et al.*, 1999; Golz *et al.*, 2002; Kramer *et al.*, 2007).

***KNOX* genes are able to induce ectopic petal outgrowths in transgenic tobacco** Constitutive expression of the class 1 *KNOX* genes isolated from *A. majus*, *L. vulgaris* and *D. fuchsii* generates a range of typical vegetative and floral phenotypes in transgenic tobacco (Kano-Murakami *et al.*, 1993; Matsuoka *et al.*, 1993; Müller *et al.*, 1995; Nishimura *et al.*, 2000; Sato *et al.*, 1996; Sinha *et al.*, 1993; Tamaoki *et al.*, 1997). However, constitutive expression of *AmHIRZ*, generated transgenic tobacco plants with a remarkable phenotype, ectopic sac-like outgrowths on the corolla tube. The appearance of

this phenotype is strong in 35S::*AmHIRZ* tobacco lines but less distinct when expressing other *KNOX* genes isolated from *L. vulgaris* and *D. fuchsii*. Extensive searches of the literature have not revealed any previous reports of this phenotype. However, constitutive gene expression does not provide a precise determination of protein function. As such, the uniqueness of the 35S::*AmHIRZ* phenotype is still uncertain. Similar, but less clear phenotypes, are observable in 35S::*AmINA* and 35S::*LvHIRZ* constitutive expression lines, strongly indicating that the ectopic sac-like structures in 35S::*AmHIRZ* lines represent a specific manifestation of a general property of *KNOX* genes. It is highly likely that many class 1 *KNOX* genes could generate the 35S::*AmHIRZ* phenotype. Such a theory requires further testing but implies that *KNOX* genes can induce novel outgrowths of petal tissue, as suggested by Golz *et al.* (2002).

9.1.2 A mechanistic role for *KNOX* genes in nectar spur development

Taken together, the data presented in this thesis clearly indicate that nectar spurs may have evolved by similar means in phylogenetically disparate taxa and that both eudicots and monocots may utilise *KNOX* genes to drive the initiation and elongation of the floral nectar spur. However, these data not only indicate a role for *KNOX* genes in nectar spur evolution and development, but also provide clues as to the mechanistic role of *KNOX* genes in nectar spur development. Previous work focused on investigating *KNOX* function in the SAM (Kerstetter *et al.*, 1997; Long *et al.*, 1996; Vollbrecht *et al.*, 2000) and complex leaf morphogenesis (Champagne *et al.*, 2007; Chen *et al.*, 1997; Hareven *et al.*, 1996; Hay & Tsiantis, 2006; Kimura *et al.*, 2008; Kumar *et al.*, 2007; Müller *et al.*, 2006; Shani *et al.*, 2009) suggests that most class 1 *KNOX* genes share a fundamental cellular function in maintaining indeterminacy and suppressing cellular differentiation. Such a function may also be important in nectar spur development.

Golz *et al.* (2002) noted expression of *AmHIRZ* and *AmINA* in the floral meristem and receptacle of wild-type snapdragon plants, proximal to developing petals, where they are proposed to establish a proximal organiser at the base of the corolla that controls growth and fate of cells along the proximal-distal axis. Ectopic expression of *KNOX* genes in the corolla of *Hirz-d153* and *Ina-d1* snapdragon mutants was suggested to induce a novel axis of growth by generating an additional ectopic *organiser*, resulting in the ectopic petal tubes that characterise these mutants (Golz *et al.*, 2002). In Chapter 6 a number of objections were presented that suggest that *KNOX* genes do not act as proximo-distal organisers. Most notable among these objections is the lack of evidence for similar patterns of *KNOX* gene expression in the receptacle of other flowers with tubular corollas, such as tobacco (Nishimura *et al.*, 1999). Furthermore, an organising role for *KNOX* genes

seems to conflict with the prevailing view that class 1 *KNOX* genes act as indeterminacy factors.

If *KNOX* genes are not acting as an organiser of corolla or nectar spur development then how else might they function? The role of *KNOX* genes in the development of complex and diverse leaf morphologies has received considerable attention in recent years (Bharathan *et al.*, 2002; Champagne *et al.*, 2007; Hareven *et al.*, 1996; Hay & Tsiantis, 2006; Jasinski *et al.*, 2007; Shani *et al.*, 2009). The insights gleaned from the role of *KNOX* genes in compound leaf development may be directly relevant to the potential function of *KNOX* genes in the development of floral nectar spurs. Ectopic expression of *KNOX* genes in leaves confers indeterminate features, resulting in ectopic meristems, or outgrowths capable of converting leaves from simple to lobed in transgenic *Arabidopsis* constitutively expressing *KNOX* (Chuck *et al.*, 1996; Sentoku *et al.*, 2000; Sinha *et al.*, 1993); from simple to dissected in the case of wild-type *C. hirsuta* (Hay & Tsiantis, 2006); or from compound to super-compound in transgenic plants and natural mutants of tomato (Hareven *et al.*, 1996; Janssen *et al.*, 1998a). These studies strongly suggest that persistence of *KNOX* gene expression in leaf primordia can lead to prolonged organogenic activity and the formation of complex leaf morphologies, perhaps by preventing the precocious exit of tissues from the cell cycle (Hay & Tsiantis, 2006).

Typically, *KNOX* GoF mutants, or constitutively expressing transgenic plants, have variable phenotypes, suggesting that the context of *KNOX* gene expression is of vital importance (reviewed in Hay *et al.* (2009); Tsiantis & Hay (2003)). In a recent publication Shani *et al.* (2009) provide an excellent example of this, showing that the phenotypic effects associated with expression of the tomato *KNOX* genes *TKN1* and *TKN2* are strongly context-dependent. *TKN* affected leaf shape only when expressed during primary maturation (PM) of the leaf. *TKN* expression during early initiation (EI) halts leaf development by extending the leaf initiation stage, generating a simple leaf. However, if expressed during PM, *TKN* expression significantly increases the organogenic potential of the leaf, resulting in dramatically increased leaf compounding. *KNOX* genes may act in a similar manner in promoting the development of floral nectar spurs. Rather than acting as an organiser *per se*, expression of *KNOX* genes at appropriate stages of petal development may increase the organogenic potential of the petal, maintaining the morphogenetic *competency* of the organ by delaying cellular differentiation, most likely by modulating levels of plant growth substances, such as GA and cytokinins (*KNOX* interactions with plant growth substances have been reviewed by Hay *et al.* (2004); Shani *et al.* (2006), see Chapter 1).

In the absence of an apparent *organiser* of petal development in the tubular flowers of tobacco, the *competency model* may provide a more satisfactory explanation for the broad

patterns of *KNOX* gene expression observed in *L. vulgaris* and snapdragon mutants. In particular, such a model may more adequately explain the phenotypic inconsistency of the *Hirz-d153* and *Ina-d1* snapdragon mutants, which do not always generate ectopic petal tubes on the corolla (Golz *et al.*, 2002). In the snapdragon mutants, *KNOX* gene expression may only result in the development of ectopic petal tubes if expressed within a critical window of petal maturation. The same may be said for the true spurs of *L. vulgaris*, *D. fuchsii* and *D. viridis* flowers. In accordance with this hypothesis *KNOX* gene expression may therefore be viewed as a *facilitator* of petal elaboration that acts at specific stages within the nectar spur development program to delay maturation and permit spur formation, rather than set the nectar spur route. If such a model is true, it indicates that *KNOX* genes are critical for floral nectar spur development, at least in Antirrhineae and orchids, but that additional unknown factors are also required and remain to be identified. This model may also implicate *KNOX* genes in a variety of additional organ elaborations that appear to be commonplace amongst angiosperms, and provide further support for the idea that plastic changes in the spatio-temporal progress of organ maturation is one of the key mechanisms employed by plants as a source of novel morphological diversity.

How does the *facilitator model* relate to the evolution of nectar spurs in *Antirrhineae* and orchids? Bateman & DiMichele (2002); Box *et al.* (2008) and Golz *et al.* (2002) implicated neomorphic mutations in spur evolution in orchids and Antirrhineae. Although *KNOX* genes are unlikely to be the sole regulators of nectar spurs, if *KNOX* genes function to *facilitate* spur development they may well represent the neomorphic alleles implicated in the evolution of petal tube diversity in Antirrhineae and orchids. Difference in nectar spur length in the orchids *D. fuchsii* and *D. viridis* may be explained by changes in the timing and/or duration of *KNOX* gene expression relative to particular stages of petal maturation. By altering the timing of *KNOX* gene expression the morphogenetic competency of the labellum petal, and therefore the potential for nectar spur development, may be significantly altered, resulting in a shorter nectar spur in *D. viridis* (see Chapter 8). Likewise, the apparent ease with which nectar spurs appear to have been lost or gained in Antirrhineae could be explained by alterations to the temporal expression patterns of *KNOX* genes. The contrasting pattern of *KNOX* gene expression in gibbous wild-type *A. majus* and spurred *L. vulgaris* may represent a neomorphic change that has been recapitulated in the *Hirz-d153* and *Ina-d1* snapdragon mutants. Neomorphic mutations such as these are not traditionally thought to account for evolutionary change, as the pleiotropic effects of altered gene expression are likely to be deleterious. However, the apparent functional redundancy observed among class 1 *KNOX* paralogs may buffer potentially deleterious effects of *KNOX* misexpression. Precise spatio-temporal patterns

of floral *KNOX* gene expression patterns in *L. vulgaris*, *D. fuchsii* and *D. viridis* are required to validate this hypothesis, while including additional orchid and Antirrhineae taxa would also add considerable support.

9.1.3 If *KNOX* genes are unlikely to be the sole regulators of nectar spur evolution and development, what else could be?

Alterations in the timing and or presence of *KNOX* gene expression may be a key genetic determinant of nectar spur development and evolution, at least in some angiosperm taxa. However, it is not surprising that the picture emerging of the developmental-genetic pathway contributing to nectar spur development seems to be a complex one. The emerging model for nectar spur evolution and development presented in this work (and the related model by Golz *et al.* (2002)) implicate a number of genes in the development of nectar spurs, including canonical ABC genes and floral symmetry and *KNOX* genes, not to mention the additional factors associated with regulating *KNOX* gene expression. However, the principal genetic determinants of nectar spurs *per se* remain elusive. If *KNOX* genes are not responsible for the outgrowth and sculpturing of the nectar spur itself, what other genes might be required? Other potential candidate genes include the TCP family of transcription factors and homologs of the gene *JAGGED* from *Arabidopsis*, both of which play significant roles in controlling the rates and patterning of cell division in the development of lateral organs (Crawford *et al.*, 2004; Dinneny *et al.*, 2004; Ohno *et al.*, 2004).

The *Antirrhinum* gene *CINCINNATA* (*CIN*) encodes a member of the TCP transcription factor family, which also contains well-characterised transcription factors such as CYC and DICH (Crawford *et al.*, 2004). *CIN* affects growth at the margins of petals, where it is specifically expressed and is believed to play a role in elaborating the marginal regions of *Antirrhinum* petals, generating the characteristic petal lobes of wild-type *A. majus* flowers. *CIN* achieves this by promoting growth in the petal lobes as *cin* mutants have smaller than normal petal lobes. Interestingly, other TCP transcription factors have also been shown to have regional specific effects on growth, such as CYC and DICH themselves (Crawford *et al.*, 2004; Galego & Almeida, 2002; Luo *et al.*, 1999). *CIN* promotes growth in the petal lobes by influencing the cell cycle, causing extra cell divisions (Crawford *et al.*, 2004). Such a function is consistent with what may be envisaged in the development of elaborate petal outgrowths such as nectar spurs. While it is true that *CIN* is not expressed in the corolla tube of spurless *A. majus* (Crawford *et al.*, 2004), there is currently no evidence to suggest whether it is expressed in the corolla tubes of

spur-bearing Antirrhineae, or the petals of other spur-bearing genera such as *Aquilegia* and orchids. Such a lack of evidence, however, results from the absence of experimentation. Exploring the expression patterns of *CIN* orthologs from *L. vulgaris* in particular may shed light on their possible role in nectar spur development.

In addition to TCP transcription factors, homologs of the *Arabidopsis* gene *JAGGED* (*JAG*) may be required for nectar spur development. The *JAG* gene encodes a zinc-finger domain protein involved in controlling the morphogenesis of lateral organs (Dinneny *et al.*, 2004; Ohno *et al.*, 2004). In *jag* mutants, petals (among other lateral organs) do not develop completely, with the strongest effects concentrated in the distal most regions (Dinneny *et al.*, 2004; Ohno *et al.*, 2004). *JAG* is thought to slow the cessation of cell division by reducing cell-cycle activity in the locality of *JAG* expression in the distal regions of lateral organs (Dinneny *et al.*, 2004), thereby maintaining meristematic competency of cells along a proximal-distal axis (Ohno *et al.*, 2004). This may allow differential development of structures in the proximal and distal regions of petals. In this sense *JAG* may function as a proximal-distal organiser that Golz *et al.* (2002), incorrectly, envisaged for *KNOX* genes. An additional factor that may play a complementary role to *JAG* in proximo-distal patterning is the AP2-like factor *AINTEGUMENTA* (Krizek, 1999; Mizukami & Fischer, 2000).

The involvement of any of these genes in the development of proximo-distal outgrowths of petals, such as floral nectar spurs, may be of significant interest. It is likely that if such genes are involved, they may function co-operatively with *KNOX* gene expression. Any such involvement for these factors remains to be demonstrated.

9.2 Nectar spurs, *KNOX* genes and the evolution of morphological novelty

It has long been understood that morphological evolution occurs through alterations in the pattern of ontogenesis (Gould, 1977). Understanding the genetic and molecular mechanisms governing such alterations and the evolution of morphological diversity is a major challenge in biology. An ever increasing body of comparative studies have revealed that morphological differences between taxa are caused by spatio-temporal differences in gene expression and the recruitment of existing developmental programs to carry out new developmental roles. These data are largely derived from investigations focused on understanding the evolution and development of the arthropod body plan (reviewed by Hughes & Kaufman (2002)), however in recent years, a number of excellent examples have also been shown in plants (reviewed in Rosin & Kramer (2009)). Taken together, these data strongly support the concept that evolutionary modifications of gene expression, i.e.

regulatory evolution, are the basis of morphological diversification.

Examining nectar spur development in both orchids and *Linaria* provides an excellent opportunity to address a number of fundamental evolutionary questions:

- Do similar molecular mechanisms underlie the repeated evolution of a single trait?
- Are existing developmental pathways redeployed in the evolution of novel biological structures?
- Can neomorphic mutations in key regulatory genes contribute to the origin of novel biological structures in plants, a phenomenon well characterised in animal models?

A number of plant examples demonstrating the importance of regulatory evolution may be cited, e.g. diversification of the MADS-box gene family and morphological diversity of flowers, or multiple independent incidences of TCP gene co-option in determining patterns of floral zygomorphy (reviewed in Rosin & Kramer (2009)). However, the *KNOX* gene family provides a particularly good example. Multiple independent taxa have recycled existing *KNOX*-dependent developmental pathways and modified them in order to generate biological novelty. Key to this phenomenon is the fundamental activity of class 1 *KNOX* genes themselves, which appears to lend itself to a variety of developmental processes. Despite an ancestral role in maintaining the SAM (Kerstetter *et al.*, 1997; Long *et al.*, 1996; Vollbrecht *et al.*, 2000), many species with independently derived compound leaves have co-opted the *KNOX*-dependent pathway for maintaining indeterminacy in the SAM to extend the morphogenetic potential of the developing leaf and generate compound leaf morphologies (Bharathan *et al.*, 2002; Hareven *et al.*, 1996; Hay & Tsiantis, 2006; Rosin *et al.*, 2003; Shani *et al.*, 2009).

The results presented in this thesis suggest that co-option of *KNOX*-dependent pathways may also have been pivotal in the evolution of the floral nectar spur in *L. vulgaris* and orchids. Furthermore, the broad phylogenetic distance separating the eudicot *L. vulgaris* from the monocots *D. fuchsii* and *D. viridis*, suggests that co-option of *KNOX* genes for nectar spur development may have been commonplace among angiosperms. Subsequently, the co-opted *KNOX* genes have been incorporated into existing petal development pathways, where it appears that activity at crucial stages of petal maturation is able to induce novel outgrowths of petal tissue that have subsequently been refined to generate the nectar spurs characteristic of *L. vulgaris* and *D. fuchsii*. It is highly likely that fine tuning of such a pathway through neomorphic mutations has generated the diversity in spur-length that appears to have been critical for the high species-level diversity of spur-bearing taxa.

9.2.1 Conclusions and future progress

KNOX genes appear to be a critical factor in the development of the petal derived nectar spurs of *L. vulgaris* and the orchids *D. fuchsii* and *D. viridis*, and may contribute to the evolution of diverse nectar spur morphologies via alterations in the spatio-temporal patterns of *KNOX* gene expression in relation to the particular stage of petal maturation. Furthermore, the broad phylogenetic gulf separating *L. vulgaris* and orchids indicates that many spur-bearing angiosperm taxa may have co-opted *KNOX* genes for a role in nectar spur development. Closer examination of the mechanistic role that *KNOX* genes may play in the spur development pathway indicates that, although *KNOX* genes appear to be able to induce outgrowths of petal tissue, it is likely that rather than setting the nectar spur fate, *KNOX* genes function to extend the morphogenetic potential of the petal from which the nectar spur develops, functioning in a fundamentally similar manner in the developing petals, compound leaves and the SAM. Taken together, the available data indicate that *KNOX* genes confer a meristematic state upon plant tissues in a variety of morphogenetic contexts, making the gene family a potentially versatile tool to mediate evolutionary transformations.

Although this work has identified a key component of angiosperm nectar spur development, it has also demonstrated that nectar spur development appears to require a number of additional genetic factors, such as ABC and TCP genes, which regulate floral organ fate and zygomorphy. Additional components related to the specific cell fates and micromorphological sculpturing of the spur itself remain to be identified. The interaction between *KNOX* genes, their regulators and downstream targets has become increasingly well understood in recent years but remains largely unexplored (reviewed by Hay *et al.* (2009)). Understanding the interactions between *KNOX* genes and their downstream targets is likely to provide further clues to the development of nectar spurs in angiosperms. Expanding investigations of *KNOX* gene activity to additional spur-bearing taxa, establishing precise spatio-temporal patterns of *KNOX* expression, and developing techniques to transgenically alter *KNOX* gene expression patterns in *Linaria* and orchids, e.g. silencing of *LvHIRZ*, will provide further clarification of the role of *KNOX* genes in spur development. Further understanding is also likely to emerge from the burgeoning field of deep sequencing and comparative transcriptomics.

The development of new genetic models with more diverse floral traits, including nectar spurs, is also likely to provide valuable insights, e.g. development of the spur-bearing eudicot *Aquilegia* as a genetic model (Kramer, 2009). Development of new genetic models should provide an enhanced toolkit with which to explore the development of nectar spurs in the future. In addition to *Aquilegia*, however, spur bearing genera in the Antirrhineae, e.g. *Linaria*, are ideal candidates for model development, benefiting from the existence

of a close relative that is already a well-established model, *A. majus*. The existence of naturally occurring mutants of *L. vulgaris* specifically affecting spur development (Cubas *et al.*, 1999), could provide particularly powerful insights when combined with modern deep sequencing techniques. Analysing the transcriptome of wild-type vs ectopic spur-bearing petals of peloric *Linaria* mutants could help unravel the entire genetic pathway regulating nectar spur development in Antirrhineae. Although not ideally suited for model development, recent advances in sequencing technologies means that similar experiments may also be practicable in orchids such as *D. fuchsii* and *D. viridis*, plus a variety of additional angiosperm taxa more representative of the full diversity of nectar spur morphologies. *Biology has never been so well equipped to unravel the mysteries of evolution.*

Bibliography

Abouheif, E. (1997). Developmental genetics and homology: a hierarchical approach. *Trends in Ecology & Evolution*, **12**, 405–408.

Abouheif, E., Akam, M. & Dickinson, W. (1997). Homology and developmental genes. *Trends in Genetics*, **13**, 432–433.

Adachi, Y., Mori, S. & Nakano, M. (2004). *Agrobacterium*-mediated production of transgenic plants in *Tricyrtis hirta* (Liliaceae). *ISHS Acta Horticulturae: IX International Symposium on Flower Bulbs*, **673**.

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H. & Tasaka, M. (1997). Genes lnvolved in Organ Separation in *Arabidopsis*: An Analysis of the *cup-shaped cotyledon* mutant. *The Plant Cell*, **9**, 841–857.

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

Albach, D., Meudt, H. & Oxelman, B. (2005). Piecing together the "new" Plantaginaceae. *American Journal of Botany*, **92**, 297–315.

Alberch, P., Gould, S., Oster, G. & Wake, D. (1979). Size and shape in ontogeny and phylogeny. *Paleobiology*, **5**, 296–317.

Almeida, J., Rocheta, M. & Galego, L. (1997). Genetic control of flower shape in *Antirrhinum majus*. *Development*, **124**, 1387–1392.

Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.

Arnold, R. (1982). Pollination, predation and seed set in *Linaria vulgaris* (Scrophulariaceae). *American Midland Naturalist*, **107**, 360–369.

Averof, M. & Akam, M. (1995). Hox genes and the diversification of insect and crustacean body plans. *Nature*, **376**, 420–423.

Avivi, Y., Lev-Yadun, S., Morozova, N., Libs, L., Williams, L., Zhao, J., Varghese, G. & Grafi, G. (2000). *Clausa*, a Tomato Mutant with a Wide Range of Phenotypic Perturbations, Displays a Cell Type-Dependent Expression of the Homeobox Gene *LeT6/TKn2*. *Plant Physiology*, **124**, 541–551.

Awasthi, D., Kumar, V. & Murty, Y. (1984). Flower development in *Antirrhinum majus* L. (Scrophulariaceae) with a comment upon corolla tube formation. *Botanical Magazine Tokyo*, **97**, 13–22.

Barton, M. & Poethig, R. (1993). Formation of the Shoot Apical Meristem in *Arabidopsis thaliana* - An Analysis of Development in the Wild-Type and in the shoot meristemless mutant. *Development*, **119**, 823–831.

Bateman, R.M. (2001). Evolution and classification of European orchids: insights from molecular and morphological characters. *Journal Europäischer Orchideen*, **33**, 33–119.

Bateman, R.M. (2005). Circumscribing and interpreting closely related orchid species: *Platanthera, Dactylorhiza* and the crucial role of mutation. *Journal of the Hardy Orchid Society*, **2**, 104–111.

Bateman, R.M. & DiMichele, W.A. (1994). Saltational evolution of form in vascular plants: a neoGoldschmidtian synthesis. In D.S. Ingram & A. Hudson, eds., *Shape and Form in Plants and Fungi*, 63–102, Linnean Society Symposium Series, Academic Press, London.

Bateman, R.M. & DiMichele, W.A. (2002). Generating and filtering major phenotypic novelties: neoGoldschmidtian saltation revisited. In Q.C.B. Bronk, R.M. Bateman & J.A. Hawkins, eds., *Developmental Genetics and Plant Evolution*, 109–159, Taylor Francis, London.

Bateman, R.M. & Rudall, P.J. (2006). The good, the bad, and the ugly: Using naturally occurring terata to distinguish the possible from the impossible in Orchod Floral Evolution. *Aliso*, **22**, 481–496.

Bateman, R.M. & Sexton, R. (2008). Is spur length of *Platanthera* species in the British Isles adaptively optimized or an evolutionary red herring? *Watsonia*, **27**, 1–21.

Bateman, R.M., Pridgeon, A. & Chase, M. (1997). Phylogenetics of subtribe Orchidinae (Orchidoideae, Orchidaceae) based on nuclear *ITS* sequences. 2. Infrageneric

relationships and reclassification to achieve monophyly of *Orchis sensu stricto*. *Lindleyana*, **12**, 113–141.

Bateman, R.M., Hollingsworth, P.M., PRESTON, J., Luo, Y.B., Pridgeon, A.M. & Chase, M.W. (2003). Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). *Botanical Journal of the Linnean Society*, **142**, 1–40.

Bateman, R.M., Rudall, P.J. & James, K. (2006). Phylogenetic context, generic affinities and evolutionary origin of the enigmatic Balkan orchid *Gymnadenia frivaldii* Hampe ex Griseb. *Taxon*, **55**, 107–118.

Bauer, P., Lubkowitz, M., Tyers, R., Nemoto, K., Meeley, R.B., Goff, S.A. & Freeling, M. (2004). Regulation and a conserved intron sequence of *liguleless3/4 Knox* class-I homeobox genes in grasses. *Planta*, **219**, 359–368.

Belarmino, M. & Mii, M. (2000). *Agrobacterium*-mediated genetic transformation of a *Phalaenopsis* orchid. *Plant Cell Reports*, **19**, 435–442.

Bell, A., Roberts, D., Hawkins, J., Rudall, P., Box, M. & Bateman, R. (2009). Comparative micromorphology of nectariferous and nectarless labellar spurs in selected clades of subtribe Orchidinae (Orchidaceae). *Botanical Journal of the Linnean Society*, **160**, 369–387.

Bellaoui, M., Pidkowich, M., Samach, A. & Kushalappa, K. (2001). The Arabidopsis BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. *Plant Cell*, **13**, 2455–2470.

Belles-Boix, E., Hamant, O., Witiak, S. & Morin, H. (2006). *KNAT6*: An Arabidopsis Homeobox Gene Involved in Meristem Activity and Organ Separation. *The Plant Cell*, **18**, 1900–1907.

Bendahmane, A., Querci, M. & Kanyuka, K. (2000). *Agrobacterium* transient expression system as a tool for the isolation of disease resistance. *Plant Journal*, **21**, 73–81.

Benfey, P.N., Ren, L. & Chua, N. (1989). The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. *EMBO Journal*, **8**, 2195–2202.

Bennetzen, J. (2002). Mechanisms and rates of genome expansion and contraction in flowering plants. *Genetica*, **115**, 29–36.

Bertolino, E., Reimund, B. & Wildt-Perinic, D. (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., Kellogg, E. & Sinha, N. (1997). Did homeodomain proteins duplicate before the origin of angiosperms, fungi, and metazoa? *Proceedings of the National Academy of Sciences*, **94**, 13749–13753.

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

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

Billeter, M., Qian, Y., Otting, G. & Müller, M. (1993). Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *Journal of Molecular Biology*, **234**, 1084–1097.

Blamey, M. & Grey-Wilson, C. (1989). Flora of Britain and Northern Europe. *Hodder & Stoughton, UK*, 544.

Blázquez, M., Ferrandiz, C., Madueño, F. & Parcy, F. (2006). How floral meristems are built. *Plant Molecular Biology*, **60**, 855–870.

Bolduc, N. & Hake, S. (2009). The Maize Transcription Factor KNOTTED1 Directly Regulates the Gibberellin Catabolism Gene *ga2ox1*. *The Plant Cell*, 1–12.

Box, M., Bateman, R., Glover, B. & Rudall, P. (2008). Floral ontogenetic evidence of repeated speciation via paedomorphosis in subtribe Orchidinae (Orchidaceae). *Botanical Journal of the Linnean Society*, **157**, 429–454.

Bradley, D., Carpenter, R., Sommer, H., Hartley, N. & Coen, E. (1993). Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the plena locus of *Antirrhinum*. *Cell*, **72**, 85–95.

Bremer, B., Bremer, K., Chase, M., Fay, M., Reveal, J., Soltis, D., Soltis, P., Stevens, P., Anderberg, A., Moore, M., Olmstead, R., Rudall, P., Sytsma, K., Tank, D., Wurdack, K., Xiang, J. & Zmarzty, S. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, **161**, 105–121.

Burger, W. (1981). Why are there so many kinds of flowering plants? *Bioscience*, **31**, 572–581.

Bürglin, T. (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.

Bürglin, T. (1998). The PBC domain contains a MEINOX domain: Coevolution of Hox and TALE homeobox genes. *Development Genes and Evolution*, **208**, 113–116.

Burns-Balogh, P. & Funk, V. (1986). A phylogenetic analysis of the Orchidaceae. *Smithsonian Contributions to Botany*, **61**, 1–79.

Busch, M., Bomblies, K. & Weigel, D. (1999). Activation of a floral homeotic gene in *Arabidopsis*. *Science*, **285**, 585–587.

Byrne, M., Barley, R., Curtis, M., Arroyo, J. & Dunham, M. (2000). *Asymmetric leaves1* mediates leaf patterning and stem cell function in *Arabidopsis*. *Nature*, **408**, 967–971.

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

Cai, H. & Shen, P. (2001). Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science*, **291**, 493–495.

Champagne, C. & Ashton, N. (2001). Ancestry of *KNOX* Genes Revealed by Bryophyte (*Physcomitrella patens*) Homologs. *New Phytologist*, **150**, 23–36.

Champagne, C. & Sinha, N. (2004). Compound leaves: equal to the sum of their parts? *Development*, **131**, 4401–4412.

Champagne, C., Goliber, T. & Wojciechowski, M. (2007). Compound Leaf Development and Evolution in the Legumes. *The Plant Cell*, **19**, 3369–3378.

Chan, R., Gago, G., Palena, C. & Gonzalez, D. (1998). Homeoboxes in plant development. *BBA-Gene Structure and Expression*.

Chase, M. (2004). Monocot relationships: an overview. *American Journal of Botany*, **91**, 1645–1655.

Chen, H., Rosin, F., Prat, S. & Hannapel, D. (2003). Interacting Transcription Factors from the Three-Amino Acid Loop Extension Superclass Regulate Tuber Formation. *Plant Physiology*, **132**, 1391–1404.

Chen, J., Janssen, B., Williams, A. & Sinha, N. (1997). A Gene Fusion at a Homeobox Locus: Alterations in Leaf Shape and lmplications for Morphological Evolution. *The Plant Cell*, **9**, 1289–1304.

Chuck, G., Lincoln, C. & Hake, S. (1996). *KNAT1* induces lobed leaves with ectopic meristems when overexpressed in *Arabidopsis*. *The Plant Cell*, **8**, 1277–1289.

Coen, E. & Meyerowitz, E. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature*, **353**, 31–37.

Coen, E., Romero, J., Doyle, S., Elliott, R., Murphy, G. & Carpenter, R. (1990). *Floricaula*: a homeotic gene required for flower development in *Antirrhinum majus*. *Cell*, **63**, 1311–1322.

Cole, M., Nolte, C. & 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.

Cong, B., Liu, J. & Tanksley, S. (2002). Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. *Proceedings of the National Academy of Sciences*, **99**, 13606–13611.

Coyne, J. & Lande, R. (1985). The genetic basis of species differences in plants. *Am Nat*, **126**, 141–145.

Cozzolino, S. & Widmer, A. (2005). Orchid diversity: an evolutionary consequence of deception? *Trends in Ecology & Evolution*, **20**, 487–494.

Crawford, B., Nath, U., Carpenter, R. & Coen, E. (2004). *CINCINNATA* controls both cell differentiation and growth in petal lobes and leaves of *Antirrhinum*. *Plant Physiology*, **135**, 244–253.

Crepet, W. (2008). The Fossil Record of Angiosperms: Requiem or Renaissance? *Annals of the Missouri Botanical Garden*, **95**, 3–33.

Crepet, W. & Niklas, K. (2009). Darwin's second'abominable mystery': Why are there so many angiosperm species? *American Journal of Botany*, **96**, 366–381.

Cubas, P. (2004). Floral zygomorphy, the recurring evolution of a successful trait. *BioEssays*, **26**, 1175–1184.

Cubas, P., Vincent, C. & Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*, **401**, 157–161.

Cui, M., Takayanagi, K., Kamada, H., Nishimura, S. & Handa, T. (2001). Efficient shoot regeneration from hairy roots of *Antirrhinum majus* L. transformed by the rol type MAT vector system. *Plant Cell Reports*, **20**, 55–59.

Cui, M., Handa, T. & Ezura, H. (2003). An improved protocol for *Agrobacterium*-mediated transformation of *Antirrhinum majus* L. *Molecular Genetics and Genomics*, **270**, 296–302.

Cui, M., Ezura, H., Nishimura, S., Kamada, H. & Handa, T. (2004). A rapid *Agrobacterium*-mediated transformation of *Antirrhinum majus* L. by using direct shoot regeneration from hypocotyl explants. *Plant Science*, **166**, 873–879.

Damerval, C. & Nadot, S. (2007). Evolution of Perianth and Stamen Characteristics with Respect to Floral Symmetry in Ranunculales. *Annals of Botany*, **100**, 631–640.

Darwin, C. (1862). *On the various contrivances by which Orchids are Fertilized by Insects*. John Murray, London, 1st edn.

Darwin, F. & Seward, A. (1903). *More letters of Charles Darwin: a record of his work in a series of hitherto unpublished letters*, vol. 2. Appleton, New York.

Davies, B., Motte, P., Keck, E., Saedler, H., Sommer, H. & Schwarz-Sommer, Z. (1999). PLENA and FARINELLI: redundancy and regulatory interactions between two Antirrhinum MADS-box factors controlling flower development. *EMBO Journal*, **18**, 4023–4034.

Dean, G., Casson, S. & Lindsey, K. (2004). *KNAT6* gene of *Arabidopsis* is expressed in roots and is required for correct lateral root formation. *Plant Molecular Biology*, **54**, 71–84.

Delforge, P. (2006). *Orchids of Europe, North Africa and the Middle East*. A C Black, London.

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

Dilcher, D. (2000). Toward a new synthesis: major evolutionary trends in the angiosperm fossil record. *Proceedings of the National Academy of Sciences*, **97**, 7030–7036.

Dinneny, J., Yadegari, R., Fischer, R., Yanofsky, M. & Weigel, D. (2004). The role of *JAGGED* in shaping lateral organs. *Development*, **131**, 1101–1110.

Dobzhansky, T. (1937). *Genetics and the origin of species*. Columbia University Press, New York., 1st edn.

Doebley, J. & Lukens, L. (1998). Transcriptional regulators and the evolution of plant form. *The Plant Cell*, **10**, 1075–1082.

Donnelly, P., Bonetta, D., Tsukaya, H., Dengler, R. & Dengler, N. (1999). Cell cycling and cell enlargement in developing leaves of *Arabidopsis*. *Developmental Biology*, **215**, 407–419.

Donoghue, M. & Doyle, J. (2000). Seed plant phylogeny: Demise of the anthophyte hypothesis? *Current Biology*, **10**, R106–R109.

Donoghue, M.J. & Mathews, S. (1998). Duplicate genes and the root of angiosperms, with an example using phytochrome sequences. *Molecular Phylogenetics and Evolution*, **9**, 489–500.

Douglas, S., Chuck, G. & Dengler, R. (2002). KNAT1 and ERECTA regulate inflorescence architecture in *Arabidopsis*. *The Plant Cell*, **14**, 547–558.

Dressler, R. (1993). *Phylogeny and classification of the orchid family*. Dioscorides, Portland.

Elisens, W. & Freeman, C. (1988). Floral nectar sugar composition and pollinator type among New World genera in tribe Antirrhineae (Scrophulariaceae). *American Journal of Botany*.

Endress, P. & Matthews, M. (2006). Elaborate petals and staminodes in eudicots: diversity, function, and evolution. *Organisms Diversity & Evolution*, **6**, 257–293.

Endrizzi, K., Moussian, B., Haecker, A., Levin, T. & 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*. *The Plant Journal*, **10**, 967–979.

Eshed, Y., Baum, S. & Bowman, J. (1999). Distinct mechanisms promote polarity establishment in carpels of *Arabidopsis*. *Cell*, **99**, 199–209.

Estruch, Prinsen, E., van Onckelen, H., Schell, J. & Spena, A. (1991). Viviparous leaves produced by somatic activation of an inactive cytokinin-synthesizing gene. *Science*, **254**, 1364–1367.

Farris, J. (1973). A Probability Model for Inferring Evolutionary Trees. *Systematic Zoology*, **22**, 297–309.

Felsenstein, J. (1973). Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Zoology*, **22**, 240–249.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, **27**, 401–410.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.

Fink, W. (1982). The Conceptual Relationship Between Ontogeny and Phylogeny. *Paleobiology*, **8**, 254–264.

Fitch, W. (1970). Distinguishing homologous from analogous proteins. *Systematic Biology*, **19**, 99–113.

Fitch, W. (1971). Toward defining the course of evolution: minimum change for a specified tree topology. *Systematic Zoology*, **10**, 406–416.

Foster, T., Yamaguchi, J., Wong, B.C., Veit, B. & Hake, S. (1999). *Gnarley1* Is a Dominant Mutation in the *knox4* Homeobox Gene Affecting Cell Shape and Identity. *The Plant Cell*, **11**, 1239–1252.

Fowler, J.E. & Freeling, M. (1996). Genetic analysis of mutations that alter cell fates in maize leaves: Dominant *Liguleless* mutations. *Developmental Genetics*, **18**, 198–222.

Freeling, M. & Hake, S. (1985). Developmental genetics of mutants that specify knotted leaves in maize. *Genetics*, **111**, 617–634.

Freudenstein, J. & Rasmussen, F. (1999). What does morphology tell us about orchid relationships? - A cladistic analysis. *American Journal of Botany*, **86**, 225–248.

Friis, E., Pedersen, K. & Crane, P. (1999). Early angiosperm diversification: the diversity of pollen associated with angiosperm reproductive structures in early cretaceous floras from Portugal. *Annals of the Missouri Botanical Garden*, **86**, 259–296.

Frohman, M., Dush, M. & Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences*, **85**, 8998–9002.

Frugis, G., Giannino, D., Mele, G. & Nicolodi, C. (2001). Overexpression of *KNAT1* in Lettuce Shifts Leaf Determinate Growth to a Shoot-Like Indeterminate Growth Associated with an Accumulation of Isopentenyl-Type Cytokinins. *Plant Physiology*, **126**, 1370–1380.

Galego, L. & Almeida, J. (2002). Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. *Genes and Development*, **16**, 880–891.

Galego, L. & Almeida, J. (2007). Molecular genetic basis of flower colour variegation in *Linaria*. *Genetical Research*, **89**, 129–134.

Garcês, H.M.P., Champagne, C.E.M., Townsley, B.T., Park, S., Malhó, R., Pedroso, M.C., Harada, J.J. & Sinha, N.R. (2007). Evolution of asexual reproduction in leaves of the genus *Kalanchoë*. *Proceedings of the National Academy of Sciences*, **104**, 15578–83.

Garstang, W. (1922). The theory of recapitulation: a critical restatement of the biogenetic law. *Zoological Journal of the Linnean Society*, **35**, 81–101.

Gascuel, O. & Steel, M. (2006). Neighbor-joining revealed. *Molecular Biology and Evolution*, **23**, 1997–2000.

Gehring, W., Qian, Y., Billeter, M., Furukubo-Tokunaga, K., Schier, F., Resendez-Perez, D., Affolter, M., Otting, G. & Wüthrich, K. (1994). Homeodomain-DNA recognition. *Cell*, **78**, 211–223.

Gehring, W.J. (1998). *Master control genes in the development and evolution: the homeobox story.*. Yale University Press.

Ghebrehiwet, M., Bremer, B. & Thulin, M. (2000). Phylogeny of the tribe Antirrhineae (Scrophulariaceae) based on morphological and *ndhF* sequence data. *Plant Systematics and Evolution*, **220**, 223–239.

Golz, J.F., Keck, E.J. & Hudson, A. (2002). Spontaneous mutations in *KNOX* genes give rise to a novel floral structure in *Antirrhinum*. *Current Biology*, **12**, 515–22.

Gordon, S., Chickarmane, V., Ohno, C. & Meyerowitz, E. (2009). Multiple feedback loops through cytokinin signaling control stem cell number within the *Arabidopsis* shoot meristem. *Proceedings of the National Academy of Sciences*, **106**, 16529–16534.

Gottlieb, L. (1984). Genetics and morphological evolution in plants. *American Naturalist*, **123**, 681–709.

Gould, S. (1977). *Ontogeny and Phylogeny*. Harvard University Press, Cambridge, MA.

Gould, S. (1992). Roots: Ontogeny and phylogeny-revisited and reunited. *BioEssays*, **14**, 275–279.

Gourlay, C., Hofer, J. & Ellis, T. (2000). Pea Compound Leaf Architecture Is Regulated by Interactions among the Genes *UNIFOLIATA*, *COCHLEATA*, *AFILA*, and *TENDRILLESS*. *The Plant Cell*, **12**, 1279–1294.

Grant, V. (1949). Pollination systems as isolating mechanisms in angiosperms. *Evolution*, **3**, 82–97.

Grant, V. (1963). *The origin of adaptations*. Columbia University Press, New York.

Greenboim-Wainberg, Y., Maymon, I. & Borochov, R. (2005). Cross Talk between Gibberellin and Cytokinin: The Arabidopsis GA Response Inhibitor SPINDLY Plays a Positive Role in Cytokinin Signaling. *The Plant Cell*, **17**, 92–102.

Groover, A.T., Mansfield, S.D., Difazio, S.P., Dupper, G., Fontana, J.R., Millar, R. & Wang, Y. (2006). The *Populus* homeobox gene *ARBORKNOX1* reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. *Plant Molecular Biology*, **61**, 917–932.

Guillet-Claude, C., Isabel, N. & Pelgas, B. (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.

Gutowski, J. (1990). Pollination of the orchid *Dactylorhiza fuchsii* by longhorn beetles in primeval forests of Northeastern Poland. *Biological Conservation*, **51**, 287–275.

Hake, S., Smith, H., Holtan, H., Magnani, E. & Mele, G. (2004). The role of*KNOX* genes in plant development. *Annual Review of Cell and Developmental Biology*, **20**, 125–151.

Hareven, D., Gutfinger, T., Parnis, A. & Eshed, Y. (1996). The making of a compound leaf: Genetic Manipulation of Leaf Architecture in Tomato. *Cell*, **84**, 735–744.

Harpster, M.H., Townsend, J.A., Jones, D.G., Bedbrook, J. & Dunsmuir, P. (1988). Relative strengths of the 35S califlower mosaic virus, 1', 2', and nopaline synthase promoters in transformed tobacco sugarbeet and oilseed rape callus tissue. *Molecular and General Genetics*, **212**, 182–190.

Harrison, J. & Langdale, J. (2006). A step by step guide to phylogeny reconstruction. *The Plant Journal*, **45**, 561–572.

Harrison, J., Corley, S., Moylan, E. & Alexander, D. (2005a). Independent recruitment of a conserved developmental mechanism during leaf evolution. *Nature*, **434**, 509–514.

Harrison, J., Möller, M., Langdale, J. & Cronk, Q. (2005b). The Role of *KNOX* Genes in the Evolution of Morphological Novelty in *Streptocarpus*. *Plant Cell*, **17**, 430–443.

Hay, A. & Tsiantis, M. (2006). The genetic basis for differences in leaf form between *Arabidopsis thaliana* and its wild relative *Cardamine hirsuta*. *Nature Genetics*, **38**, 942–947.

Hay, A., Kaur, H., Phillips, A., Hedden, P., Hake, S. & Tsiantis, M. (2002). The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. *Current Biology*, **12**, 1557–1565.

Hay, A., Craft, J. & Tsiantis, M. (2004). Plant hormones and homeoboxes: bridging the gap? *BioEssays*, **26**, 395–404.

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

Hellens, R., Mullineaux, P. & Klee, H. (2000). Technical focus: a guide to *Agrobacterium* binary Ti vectors. *Trends in Plant Science*, **5**, 446–451.

Hendy, M. & Penny, C. (1989). A framework for the quantitative study of evolutionary trees. *Systematic Biology*, **38**, 297–309.

Hernandez, C., Mathis, A., Brown, D. & Bol, J. (1995). Sequence of RNA 2 of a nematode-transmissible isolate of tobacco rattle virus. *Journal of General Virology*, **76**, 2847–2851.

Hertzberg, M., Aspeborg, H. & Schrader, J. (2001). A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences*, **98**, 14732–14737.

Hibara, K., Takada, S. & Tasaka, M. (2003). *CUC1* gene activates the expression of SAM-related genes to induce adventitious shoot formation. *Plant Journal*, **36**, 687–696.

Hileman, L. & Baum, D. (2003). Why Do Paralogs Persist? Molecular Evolution of *CYCLOIDEA* and Related Floral Symmetry Genes in Antirrhineae (Veronicaceae). *Molecular Biology and Evolution*, **20**, 591–600.

Hirayama, Y., Yamada, T., Oya, Y., Ito, M., Kato, M. & Imaichi, R. (2007). Expression patterns of class I *KNOX* and *YABBY* genes in *Ruscus aculeatus* (Asparagaceae) with implications for phylloclade homology. *Development Genes and Evolution*, **217**, 363–372.

Hodges, S. (1995). The Influence of Nectar Production on Hawkmoth Behavior, Self Pollination, and Seed Production in *Mirabilis multiflora* (Nyctaginaceae). *American Journal of Botany*, **82**, 197–204.

Hodges, S. (1997). Floral Nectar Spurs and Diversification. *International Journal of Plant Sciences*, **158**, S81–S88.

Hodges, S. & Arnold, M. (1994). Floral and ecological isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *Proceedings of the National Academy of Sciences*, **91**, 2493–2496.

Hodges, S. & Arnold, M. (1995). Spurring Plant Diversification: Are Floral Nectar Spurs a Key Innovation? *Proceedings of the Royal Society of London B: Biological Sciences*, **262**, 343–348.

Hodges, S., Whittall, J., Fulton, M. & Yang, J. (2002). Genetics of floral traits influencing reproductive isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *American Naturalist*, **159**, S51–S60.

Hodges, S., Fulton, M., Yang, J. & Whittall, J. (2004). Verne Grant and evolutionary studies of *Aquilegia*. *New Phytologist*, **161**, 113–120.

Holder, M. & Lewis, P. (2003). Phylogeny estimation: traditional and Bayesian approaches. *Nature Reviews Genetics*, **4**, 275–284.

Horsch, R., Fry, J., Hoffmann, N., Eichholtz, D., Rogers, S.G. & Fraley, R.T. (1985). A simple and general method for transferring genes into plants. *Science*, **227**, 1229–1231.

Hudson, A., Critchley, J. & Erasmus, Y. (2008). The Genus *Antirrhinum* (Snapdragon): A Flowering Plant Model for Evolution and Development. *Cold Spring Harbor Protocols*, 105–118.

Huelsenbeck, J. & Ronquist, F. (2001). MrBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.

Hughes, C. & Kaufman, T. (2002). Hox genes and the evolution of the arthropod body plan. *Evolution and Development*, **4**, 459–499.

Huijser, P., Klein, J., Lönnig, W., Meijer, H., Saedler, H. & Sommer, H. (1992). *Bracteomania*, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene squamosa in *Antirrhinum majus*. *EMBO Journal*, **11**, 1239–1249.

Huson, D.H., Richter, D.C., Rausch, C., Dezulian, T., Franz, M. & Rupp, R. (2007). Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics*, **8**, 460–465.

Ito, Y., Eiguchi, M. & Kurata, N. (1999). Expression of novel homeobox genes in early embryogenesis in rice. *BBA-Gene Structure and Expression*, **1444**, 445–450.

Iwakawa, H., Ueno, Y., Semiarti, E. & Onouchi, H. (2002). The *ASYMMETRIC LEAVES2* gene of *Arabidopsis thaliana*, required for formation of a symmetric flat leaf lamina, encodes a member of a novel family of proteins characterized by cysteine repeats and a leucine zipper. *Plant and Cell Physiology*, **43**, 467–478.

Jackson, D., Veit, B. & Hake, S. (1994). Expression of maize *KNOTTED1* related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. *Development*, **120**, 405–413.

Janssen, B., Lund, L. & Sinha, N. (1998a). Overexpression of a Homeobox Gene, *LeT6*, Reveals Indeterminate Features in the Tomato Compound Leaf. *Plant Physiology*, **117**, 771–786.

Janssen, B., Williams, A., Chen, J. & Mathern, J. (1998b). Isolation and characterization of two *knotted*-like homeobox genes from tomato. *Plant Molecular Biology*, **36**, 417–425.

Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., Phillips, A., Hedden, P. & Tsiantis, M. (2005). KNOX Action in *Arabidopsis* Is Mediated by Coordinate Regulation of Cytokinin and Gibberellin Activities. *Current Biology*, **15**, 1560–1565.

Jasinski, S., Kaur, H., Tattersall, A. & Tsiantis, M. (2007). Negative regulation of *KNOX* expression in tomato leaves. *Planta*, **226**, 1255–1263.

Jefferson, Kavanagh, T. & Bevan, M. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal*, **6**, 3901–3907.

Johnson, S. & Steiner, K. (1997). Long-Tongued Fly Pollination and Evolution of Floral Spur Length in the *Disa draconis* Complex (Orchidaceae). *Evolution*, **51**, 45–53.

Johnson, S., Linder, H. & Steiner, K. (1998). Phylogeny and radiation of pollination systems in *Disa* (Orchidaceae). *American Journal of Botany*, **85**, 402–411.

Joshi, R., Passner, J.M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M.A., Jacob, V., Aggarwal, A.K., Honig, B. & Mann, R.S. (1997). Functional Specificity of a Hox Protein Mediated by the Recognition of Minor Groove Structure. *Cell*, **131**, 530–543.

Jouannic, S., Collin, M., Vidal, B., Verdeil, J.L. & Tregear, J.W. (2007). A class I *KNOX* gene from the palm species *Elaeis guineensis* (Arecaceae) is associated with meristem function and a distinct mode of leaf dissection. *New Phytologist*, **174**, 551–568.

Kampny, C. & Harris, E. (1998). Heterochrony: The basis of floral shape evolution. *American Journal of Botany*, **Supplement 4**.

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

Katz, A., Oliva, M., Mosquna, A. & Hakim, O. (2004). FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene during sporophyte development. *Plant Journal*, **37**, 707–719.

Kenrick, P. & Crane, P. (1997). *The origin and early diversification of land plants: a cladistic study*. Smithsonian Institute Press, Washington DC.

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

Kerstetter, R., Laudencia-Chingcuanco, D., Smith, L.G. & Hake, S. (1997). Loss-of-function mutations in the maize homeobox gene, *knotted1*, are defective in shoot meristem maintenence. *Development*, **124**, 3045–3054.

Kim, J., Yuan, Z. & Jackson, D. (2003). Developmental regulation and significance of KNOX protein trafficking in *Arabidopsis*. *Development*, **130**, 4351–4362.

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

Kleisner, K. (2007). The Formation of the Theory of Homology in Biological Sciences. *Acta Biotheoretica*, **55**, 317–340.

Knoepfler, P., Calvo, K. & Chen, H. (1997). Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1. *Proceedings of the National Academy of Sciences*, **94**, 14553–14558.

Kramer, E. & Irish, V. (1999). Evolution of genetic mechanisms controlling petal development. *Nature*, **399**, 144–148.

Kramer, E., Jaramillo, A. & Stilio, V.D. (2004). Patterns of Gene Duplication and Functional Evolution During the Diversification of the *AGAMOUS* Subfamily of MADS Box Genes in Angiosperms. *Genetics*, **166**, 1011–1023.

Kramer, E.M. (2009). *Aquilegia*: a new model for plant development, ecology, and evolution. *Annual Review of Plant Biology*, **60**, 261–77.

Kramer, E.M., Holappa, L., Gould, B., Jaramillo, M.A., Setnikov, D. & Santiago, P.M. (2007). Elaboration of B Gene Function to Include the Identity of Novel Floral Organs in the Lower Eudicot *Aquilegia*. *The Plant Cell*, **19**, 750–766.

Krizek, B. (1999). Ectopic expression of *AINTEGUMENTA* in Arabidopsis plants results in increased growth of floral organs. *Developmental Genetics*, **25**, 224–236.

Krizek, B. & Fletcher, J. (2005). Molecular mechanisms of flower development: an armchair guide. *Nature Reviews Genetics*, **6**, 688–698.

Krusell, K., Rasmussen, I. & Gausing, K. (1997). DNA binding sites recognised in vitro by a knotted class 1 homeodomain protein encoded by the hooded gene, k, in barley (*Hordeum vulgare*). *FEBS Letters*, **408**, 25–29.

Kumar, R., Kushalappa, K. & Godt, D. (2007). The *Arabidopsis* BEL1-LIKE HOMEODOMAIN Proteins SAW1 and SAW2 Act Redundantly to Regulate *KNOX* Expression Spatially in Leaf Margins. *The Plant Cell*, **19**, 2719–2735.

Kumaran, M.K. (2002). *YABBY* Polarity Genes Mediate the Repression of *KNOX* Homeobox Genes in *Arabidopsis*. *The Plant Cell*, **14**, 2761–2770.

Kurzweil, H. (1998). Floral ontogeny of orchids: a review. *Beiträge zur Biologie der Pflanzen*, **71**, 45–100.

Larkin, M., Blackshields, G., Brown, N., Chenna, R., Mcgettigan, P., Mcwilliam, H., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Thompson, J., Gibson, T. & Higgins, D. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

Lawrence, P. & Morata, G. (1994). Homeobox genes: their function in *Drosophila* segmentation and pattern formation. *Cell*, **78**, 181–189.

Lawton-Rauh, A. (2003). Evolutionary dynamics of duplicated genes in plants. *Molecular Phylogenetics and Evolution*, **29**, 396–409.

Lecuit, T. & Cohen, S. (1997). Proximal–distal axis formation in the Drosophila leg. *Nature*, **388**, 139–145.

Lenhard, M., Bohnert, A., Jürgens, G. & Laux, T. (2001). Termination of stem cell maintenance in *Arabidopsis* floral meristems by interactions between WUSCHEL and AGAMOUS. *Cell*, **105**, 805–814.

Lenhard, M., Jurgens, G. & Laux, T. (2002). The *WUSCHEL* and *SHOOTMERISTEMLESS* genes fulfil complementary roles in *Arabidopsis* shoot meristem regulation. *Development*, **129**, 3195–3206.

Li, P. & Johnston, M. (2000). Heterochrony in plant evolutionary studies through the twentieth century. *The Botanical Review*, **66**, 57–88.

Li, W. (1983). Evolution of duplicated genes. In M. Nei & R. Koehn, eds., *Evolution of Genes and Proteins*, 14–37, Sinauer Associates.

Li, Y.I., Hagen, G. & Guilfoyle, T. (1992). Altered morphology in transgenic tobacco plants that overproduce cytokinins in specific tissues and organs. *Developmental Biology*, **153**, 386–395.

Liau, C.H., You, S.J., Prasad, V., Hsiao, H.H., Lu, J.C., Yang, N.S. & Chan, M.T. (2003). *Agrobacterium tumefaciens*-mediated transformation of an *Oncidium* orchid. *Plant Cell Reports*, **21**, 993–998.

Lin, W., Shuai, B. & Springer, P. (2003). The *Arabidopsis* LATERAL ORGAN BOUNDARIES –Domain Gene *ASYMMETRIC LEAVES2* Functions in the Repression of *KNOX* Gene Expression and in Adaxial-Abaxial Patterning. *The Plant Cell*, **15**, 2241–2252.

Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K. & 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. *The Plant Cell*, **6**, 1859–1876.

Little, K., Dieringer, G. & Romano, M. (2005). Pollination ecology, genetic diversity and selection on nectar spur length in *Platanthera lacera* (Orchidaceae). *Plant Species Biology*, **20**, 183–190.

Lohmann, J., Hong, R., Hobe, M., Busch, M., Parcy, F., Simon, R. & Weigel, D. (2001). A molecular link between stem cell regulation and floral patterning in *Arabidopsis*. *Cell*, **105**, 793–803.

Long, J. & Barton, M. (2000). Initiation of Axillary and Floral Meristems in *Arabidopsis*. *Developmental Biology*, **218**, 341–353.

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

Lu, R., Martin-Hernandez, A., Peart, J., Malcuit, I. & Baulcombe, D. (2003). Virus-induced gene silencing in plants. *Methods*, **30**, 296–303.

Lucas, W., Bouché-Pillon, S. & Jackson, D. (1995). Selective Trafficking of KNOTTED1 Homeodomain Protein and Its mRNA Through Plasmodesmata. *Science*, **270**, 1980–1983.

Luo, D., Carpenter, R., Copsey, L., Vincent, C. & Clark, J. (1999). Control of organ asymmetry in flowers of *Antirrhinum*. *Cell*, **99**, 367–376.

Macior, L. (1967). Pollen-foraging behavior of *Bombus* in relation to pollination of nototribic flowers. *American Journal of Botany*, **54**, 359–364.

MacWilliams, H. (1983). Hydra Transplantation Phenomena and the Mechanism of Hydra Head Regeneration. II. Properties of the Head Activation. *Developmental Biology*, **96**, 239–257.

Magnani, E. & Hake, S. (2008). KNOX Lost the OX: The *Arabidopsis KNATM* Gene Defines a Novel Class of KNOX Transcriptional Regulators Missing the Homeodomain. *The Plant Cell*, **20**, 875–887.

Mallory, A., Dugas, D., Bartel, D. & Bartel, B. (2004). MicroRNA regulation of NAC-domain targets is required for proper formation and separation of ancient embryonic, vegetative and floral organs. *Current Biology*, **14**, 1035–1046.

Manak, J. & Scott, M. (1994). A class act: conservation of homeodomain protein functions. *Development*, **Supplement**, 61–77.

Matsuoka, M., Ichikawa, H., Saito, A. & Tada, Y. (1993). Expression of a rice homeobox gene causes altered morphology of transgenic plants. *The Plant Cell*, **5**, 1039–1048.

Mattanovich, D., Rüker, F. & Machado, A. (1989). Efficient transformation of *Agrobacterium* spp. by electroporation. *Nucleic Acids Research*, **17**, 6747.

Mayr, E. (1942). *Systematics and the origin of species*. Harvard University Press.

McGinnis, W., Levine, M., Hafen, E., Kuroiwa, A. & Gehring, W. (1984). A conserved DNA sequence in homoeotic genes of the *Drosophila Antennapedia* and *bithorax* complexes. *Nature*, **308**, 428–433.

McHale, N. (1993). *LAM-1* and *FAT* genes control development of the leaf blade in *Nicotiana sylvestris*. *The Plant Cell*, **5**, 1029–1038.

Meisel, L. & Lam, E. (1996). The conserved ELK-homeodomain of *KNOTTED-1* contains two regions that signal nuclear localization. *Plant Molecular Biology*, **30**, 1–14.

Mele, G. (2003). The *knotted1*-like homeobox gene *BREVIPEDICELLUS* regulates cell differentiation by modulating metabolic pathways. *Genes and Development*, **17**, 2088–2093.

Meyer, A. (1996). The evolution of body plans: HOM/Hox cluster evolution, model systems and the importance of phylogeny. In P. Harvey, A.L. Brown, J. Smith & S. Nee, eds., *New uses for new Phylogenies*, 322–340, Oxford Univeristy Press.

Mizukami, Y. & Fischer, R. (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proceedings of the National Academy of Sciences*, **97**, 942–947.

Mondragón-Palomino, M. & Theissen, G. (2008). MADS about the evolution of orchid flowers. *Trends in Plant Science*, **13**, 51–59.

Moore, G., Devos, K., Wang, Z. & Gale, M. (1995). Grasses, line up and form a circle. *Current Biology*, **5**, 737–739.

Muehlbauer, G., Fowler, J., Girard, L. & Tyers, R. (1999). Ectopic expression of the maize homeobox gene *Liguleless3* alters cell fates in the leaf. *Plant Physiology*, **119**, 651–662.

Mukherjee, K. & Bürglin, T. (2007). Comprehensive analysis of animal TALE homeobox genes: new conserved motifs and cases of accelerated evolution. *Journal of Molecular Evolution*, **65**, 137–153.

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

Müller, J., Wang, Y., Franzen, R., Santi, L., Salamini, F. & 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. *The Plant Journal*, **27**, 13–23.

Müller, K., Romano, N., Gerstner, O., Garcia-Maroto, F., Pozzi, C., Salamini, F. & Rohde, W. (1995). The barley *Hooded* mutation caused by a duplication in a homeobox gene intron. *Nature*, **374**, 727–730.

Müller, K., He, X., Fischer, R. & Prüfer, D. (2006). Constitutive *knox1* gene expression in dandelion (*Taraxacum officinale*, Web.) changes leaf morphology from simple to compound. *Planta*, **224**, 1023–1027.

Mushegian, A. & Koonin, E. (1996). Sequence analysis of eukaryotic developmental proteins: ancient and novel domains. *Genetics*, **144**, 817–828.

Nagasaki, H., Sakamoto, T., Sato, Y. & Matsuoka, M. (2001a). Functional analysis of the conserved domains of a rice KNOX homeodomain protein, OSH15. *The Plant Cell*, **13**, 2085–2098.

Nagasaki, H., Sakamoto, T., Sato, Y. & Matsuoka, M. (2001b). Functional analysis of the conserved domains of a rice KNOX homeodomain protein, OSH15. *The Plant Cell*, **13**, 2085–98.

Nath, U., Crawford, B., Carpenter, R. & Coen, E. (2003). Genetic control of surface curvature. *Science*, **299**, 1404–1407.

Neal, P., Dafni, A. & Giurfa, M. (1998). Floral symmetry and its role in plant-pollinator systems: terminology, distribution, and hypotheses. *Annual Review of Ecology and Systematics*, **29**, 345–373.

Nepi, M., Pacini, E., Nencini, C. & Collavoli, E. (2003). Variability of nectar production and composition in *Linaria vulgaris* (L.) Mill. (Scrophulariaceae). *Plant Systematics and Evolution*, **238**, 109–118.

Newman, D. & Thomson, J. (2005). Interactions among nectar robbing, floral herbivory, and ant protection in *Linaria vulgaris*. *Oikos*, **110**, 497–506.

Nilsson, L. (1988). The evolution of flowers with deep corolla tubes. *Nature*, **334**, 147–149.

Nilsson, L.A. (1983). Processes of isolation and introgressive interplay between *Platanthera bifolia* (L.) Rich. and *P. chlorantha* (Custer) Reichb. (Orchidaceae). *Botanical Journal of the Linnean Society*, **87**, 325–350.

Nishimura, A., Tamaoki, M., Sato, Y. & Matsuoka, M. (1999). The expression of tobacco *knotted1*-type class 1 homeobox genes correspond to regions predicted by the cytohistological zonation model. *The Plant Journal*, **18**, 337–347.

Nishimura, A., Tamaoki, M., Sakamoto, T. & Matsuoka, M. (2000). Overexpression of tobacco *knotted1*-type class1 homeobox genes alters various leaf morphology. *Plant Cell Physiology*, **41**, 583–590.

Nusslein-Volhard, C. (1994). Of flies and fishes. *Science*, **266**, 572–574.

Nylander, J., Ronquist, F., Huelsenbeck, J. & Nieves-Aldrey, J. (2004). Bayesian Phylogenetic Analysis of Combined Data. *Systematic Biology*, **53**, 47–67.

Ogas, J., Kaufmann, S., Henderson, J. & Somerville, C. (1999). PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, **96**, 13839–13844.

Ohno, C., Reddy, G., Heisler, M. & Meyerowitz, E. (2004). The Arabidopsis *JAGGED* gene encodes a zinc finger protein that promotes leaf tissue development. *Development*, **131**, 111–1122.

Ohno, S. (1970). *Evolution by gene duplication*. Springer, Berlin.

Ohyanagi, H., Tanaka, T., Sakai, H., Shigemoto, Y., Yamaguchi, K., Habara, T., Fujii, Y., Antonio, B.A., Nagamura, Y., Imanishi, T., Ikeo, K., Itoh, T., Gojobori, T. & Sasaki, T. (2006). The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. japonica genome information. *Nucleic Acids Research*, **34**, D741–D744.

Ollerton, J., Alarcon, R., Waser, N. & Price, M. (2009). A global test of the pollination syndrome hypothesis. *Annals of Botany*, **103**, 1471–1480.

Olmstead, R. & Reeves, P. (1995). Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbcL* and *ndhF* sequences. *Annals of the Missouri Botanical Garden*, **82**, 176–193.

Olmstead, R., dePamphilis, C. & Wolfe, A. (2001). Disintegration of the Scrophulariaceae. *American Journal of Botany*, **88**, 348–361.

Ori, N., Juarez, M., Jackson, D. & Yamaguchi, J. (1999). Leaf Senescence Is Delayed in Tobacco Plants Expressing the Maize Homeobox Gene *knotted1* under the Control of a Senescence-Activated Promoter. *The Plant Cell*, **11**, 1073–1080.

Ori, N., Eshed, Y., Chuck, G., Brown, J. & Hake, S. (2000). Mechanisms that control *Knox* gene expression in the *Arabidopsis* shoot. *Development*, **127**, 5523–5532.

Otting, G., Qian, Y., Billeter, M. & Müller, M. (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.

Oxelman, B., Backlund, M. & Bremer, B. (1999). Relationships of the Buddlejaceae s. 1. investigated using parsimony jackknife and Branch Support Analysis of Chloroplast *ndhF* and *rbcL* Sequence Data. *Systematic Botany*, **24**, 164–182.

Oxelman, B., Kornhall, P., Olmstead, R. & Bremer, B. (2005). Further disintegration of Scrophulariaceae. *Taxon*, **54**, 411–425.

Oyama, R. & Baum, D. (2004). Phylogenetic relationships of North American *Antirrhinum* (Veronicaceae). *American Journal of Botany*, **91**, 918–925.

Parnis, A., Cohen, O., Gutfinger, T. & Hareven, D. (1997). The Dominant Developmental Mutants of Tomato, *Mouse-ear* and *Curl*, Are Associated with Distinct Modes of Abnormal Transcriptional Regulation of a *Knotted* Gene. *The Plant Cell*, **9**, 2143–2158.

Patterson, C. (1982). Homology in Classical and Molecular Biology. *Molecular Biology and Evolution*, **5**, 603–625.

Pautot, V., Dockx, J., Hamant, O. & Kronenberger, J. (2001). *KNAT2*: evidence for a link between *knotted*-like genes and carpel development. *The Plant Cell*, **13**, 1719–1734.

Pfaffl, M. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, **29**, e45.

Pfaffl, M. (2004). Quantification strategies in real-time PCR. In S. Bustin, ed., *A-Z of Quantitative PCR*, 87–112, International University Line (IUL) La Jolla, CA, USA.

Piazza, P., Jasinski, S. & Tsiantis, M. (2005). Evolution of Leaf Developmental Mechanisms. *New Phytologist*, **167**, 693–710.

Poethig, R. & Sussex, I. (1985). The developmental morphology and growth dynamics of the tobacco leaf. *Planta*, **165**, 158–169.

Postma-Haarsma, A., Verwoert, I. & Stronk, O. (1999). Characterization of the *KNOX* class homeobox genes *Oskn2* and *Oskn3* identified in a collection of cDNA libraries covering early stages of rice embryogenesis. *Plant Molecular Biology*, **39**, 257–271.

Prazmo, W. (1965). Cytogenetic studies on the genus Aquilegia III. *Acta Societatis Botanicorum Poloniae*, **34**, 403–437.

Prigge, M. & Wagner, D. (2001). The *Arabidopsis SERRATE* gene encodes a zinc-finger protein required for normal shoot development. *The Plant Cell*, **13**, 1263–1279.

Qian, Y., Billeter, M., Otting, G. & Müller, M. (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell*, **61**, 573–580.

Qiu, Y., Lee, J. & Bernasconi-Quadroni, F. (1999). The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature*, **402**, 404–407.

Ratcliff, F., Martin-Hernandez, A. & Baulcombe, D. (2001). Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant Journal*, **25**, 237–245.

Rechsteiner, M. & Rogers, S. (1996). PEST sequences and regulation by proteolysis. *Trends in Biochemical Sciences*, **21**, 267–271.

Reeves, P. & Olmstead, R. (1998). Evolution of novel morphological and reproductive traits in a clade containing *Antirrhinum majus* (Scrophulariaceae). *American Journal of Botany*, **85**, 1047–1056.

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

Reveal, J., Judd, W. & Olmstead, R. (1999). Proposal to conserve the name Antirrhinaceae against Plantaginaceae (Magnoliophyta). *Taxon*, **48**, 182.

Rolland-Lagan, A.G., Bangham, J.A. & Coen, E. (2003). Growth dynamics underlying petal shape and asymmetry. *Nature*, **422**, 161–163.

Ronquist, F., Huelsenbeck, J. & van der Mark, P. (2005). *MrBayes 3.1 Manual*. http://mrbayes.csit.fsu.edu/manual.php.

Rose, T., Schultz, E., Henikoff, J. & Pietrokovski, S. (1998). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Research*, **26**, 1628–1635.

Rosin, F., Hart, J., Horner, H., Davies, P. & Hannapel, D.J. (2003). Overexpression of a *knotted*-like homeobox gene of potato alters vegetative development by decreasing Gibberellin accumulation. *Plant Physiology*, **132**, 106–117.

Rosin, F.M. & Kramer, E.M. (2009). Old dogs, new tricks: Regulatory evolution in conserved genetic modules leads to novel morphologies in plants. *Developmental Biology*, **332**, 25–35.

Ruberti, I., Sessa, G., Lucchetti, S. & Morelli, G. (1991). A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. *EMBO Journal*, **10**, 1787–1791.

Rudall, P.J. & Bateman, R.M. (2002). Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. *Biological Reviews*, **77**, 403–441.

Rudall, P.J., Manning, J. & Goldblatt, P. (2003). Evolution of floral nectaries in Iridaceae. *Annals of the Missouri Botanical Garden*, **90**, 613–631.

Rutjens, B., Bao, D., van Eck-Stouten, E., Brand, M., Smeekens, S. & Proveniers, M. (2009). Shoot apical meristem function in *Arabidopsis* requires the combined activities of three BEL1-like homeodomain proteins. *The Plant Journal*, **58**, 641–654.

Ryoo, H., Marty, T., Casares, F. & Affolter, M. (1999). Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. *Development*, **126**, 5137–5148.

Sablowski, R. (2007). The dynamic plant stem cell niches. *Current Opinion in Plant Biology*, **10**, 639–644.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.

Sakamoto, T., Nishimura, A., Tamaoki, M., Kuba, M., Tanaka, H., Iwahori, S. & Matsuoka, M. (1999). The conserved KNOX domain mediates specificity of tobacco KNOTTED1-type homeodomain proteins. *The Plant Cell*, **11**, 1419–1431.

Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahorl, S. & 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.

Sambrook, J. & Russel, D. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 3rd edn.

Saner, M., Clements, D., Hall, M., Doohan, D. & Crompton, C. (1995). The biology of Canadian weeds. 105. *Linaria vulgaris* Mill. *Canadian Journal of Plant Science*, **75**, 525–538.

Sano, R., Juarez, C., Hass, B., Sakakibara, K. & Ito, 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.

Sargent, R. (2004). Floral symmetry affects speciation rates in angiosperms. *Proceedings: Biological Sciences*, **271**, 603–608.

Sato, T., Tamaoki, M., Murakami, T., Yamamoto, N., KanoMurakami, Y. & Matsuoka, M. (1996). Abnormal cell divisions in leaf primordia caused by the expression of the rice homeobox gene *OSH1* lead to altered morphology of leaves in transgenic tobacco. *Molecular and General Genetics*, **251**, 13–22.

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

Sato, Y., Aoki, Y. & Matsuoka, M. (2002). A Loss-of-Function Mutation in the Rice *KNOX* Type Homeobox Gene, *OSH3*. *Plant and Cell Physiology*, **43**, 44–51.

Sawa, S., Watanabe, K., Goto, K., Kanaya, E., Morito, E. & Okada, K. (1999). *FILAMENTOUS FLOWER*, a meristem and organ identity gene of *Arabidopsis*, encodes a protein with a zinc finger and HMG-related domains. *Genes and Development*, **13**, 1079–1088.

Scanlon, M., Henderson, D. & Bernstein, B. (2002). *SEMAPHORE1* functions during the regulation of ancestrally duplicated *Knox* genes and polar auxin transport in maize. *Development*, **129**, 2663–2673.

Schiestl, F. (2005). On the success of a swindle: pollination by deception in orchids. *Naturwissenschaften*, **92**, 255–264.

Schlüter, P. & Schiestl, F. (2008). Molecular mechanisms of floral mimicry in orchids. *Trends in Plant Science*, **13**, 228–235.

Schneeberger, R., Becraft, P., Hake, S. & Freeling, M. (1995). Ectopic expression of the *Knox* homeo box gene *rough sheath1* alters cell fate in the maize leaf. *Genes and Development*, **9**, 2292–2304.

Schneider, M., Ow, D. & Howell, S. (1990). The in vivo pattern of firefly luciferase expression in transgenic plants. *Plant Molecular Biology*, **14**, 935–947.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T. & Goodrich, J. (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO Journal*, **25**, 4638–4649.

Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P., Hansen, R., Tetens, F., Lonnig, W., Saedler, H. & Sommer, H. (1992). Characterization of the *Antirrhinum* floral homeotic MADS-box gene *Deficiens*: evidence for DNA binding and autoregulation of its persistent expression throughout flower development. *EMBO Journal*, **11**, 251–263.

Schwarz-Sommer, Z., Davies, B. & Hudson, A. (2003). An everlasting pioneer: the story of *Antirrhinum* research. *Nature Reviews Genetics*, **4**, 655–664.

Scofield, S. & Murray, J. (2006). *KNOX* Gene Function in Plant Stem Cell Niches. *Plant Molecular Biology*, **60**, 929–946.

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

Scofield, S., Dewitte, W. & Murray, J. (2008). A model for *Arabidopsis* class-1 *KNOX* gene function. *Plant Signaling & Behavior*, **3**, 257–259.

Scopece, G., Musacchio, A., Widmer, A. & Cozzolino, S. (2007). Patterns of reproductive isolation in Mediterranean deceptive orchids. *Evolution*, **61**, 2623–2642.

Scott, M. & Weiner, A. (1984). Structural relationships among genes that control development: sequence homology between the *Antennapedia*, *Ultrabithorax*, and *Fushi tarazu* loci of Drosophila. *Proceedings of the National Academy of Sciences*, **81**, 4115–4119.

Scott, M., Tamkun, J. & III, G.H. (1989). The structure and function of the homeodomain. *Biochemica et Biophysica Acta*, **989**, 25–48.

Semiarti, E., Indrianto, A., Purwantoro, A., Isminingsih, S., Suseno, N., Ishikawa, T., Yoshioka, Y., Machida, Y. & Machida, C. (2007). *Agrobacterium*-mediated transformation of the wild orchid species *Phalaenopsis amabilis*. *Plant Biotechnology*, **24**, 265–272.

Senior, I., Holford, P., Cooley, R. & Newbury, H. (1995). Transformation of *Antirrhinum majus* using *Agrobacterium rhizogenes*. *Journal of Experimental Botany*, **46**, 1233–1239.

Sentoku, N., Sato, Y., Kurata, N., Ito, Y., Kitano, H.H. & Matsuoka, M. (1999). Regional Expression of the Rice *KN1*-Type Homeobox Gene Family during Embryo, Shoot, and Flower Development. *The Plant Cell*, **11**, 1651–1663.

Sentoku, N., Sato, Y. & Matsuoka, M. (2000). Overexpression of rice *OSH* genes induces ectopic shoots on leaf sheaths of transgenic rice plants. *Developmental Biology*, **220**, 358–364.

Serikawa, K. & Mandoli, D. (1999). *Aaknox1*, a *kn1*-like homeobox gene in *Acetabularia acetabulum*, undergoes developmentally regulated subcellular localization. *Plant Molecular Biology*, **41**, 785–793.

Serikawa, K., Martinez-Laborda, A., Kim, H. & Zambryski, P. (1997). Localization of expression of *KNAT3*, a class 2 *knotted1*-like gene. *Plant Journal*, **11**, 853.

Shani, E., Yanai, O. & Ori, N. (2006). The role of hormones in shoot apical meristem function. *Current Opinion in Plant Biology*, **9**, 484–489.

Shani, E., Burko, Y., Ben-Yaakov, L., Berger, Y., Amsellem, Z., Goldshmidt, A., Sharon, E. & Ori, N. (2009). Stage-Specific Regulation of Solanum lycopersicum Leaf Maturation by Class 1 *KNOTTED1*-LIKE HOMEOBOX Proteins. *The Plant Cell*, **21**.

Siegfried, K., Eshed, Y., Baum, S., Otsuga, D., Drews, G.N. & Bowman, J.L. (1999). Members of the *YABBY* gene family specify abaxial cell fate in *Arabidopsis*. *Development*, **126**, 4117–4128.

Singh, D.P., Filardo, F.F., Storey, R., Jermakow, A.M., Yamaguchi, S. & Swain, S.M. (2010). Overexpression of a gibberellin inactivation gene alters seed development, *KNOX* gene expression, and plant development in *Arabidopsis*. *Physiologia Plantarum*, **138**, 74–90.

Singh, V. & Jain, D. (1979). Floral organogenesis in *Antirrhinum majus* (Scrophulariaceae). *Proceedings of the Indian Academy of Science. Sect. B.*, **88**, 183–188.

Sinha, N. & Hake, S. (1994). The *Knotted* leaf blade is a mosaic of blade, sheath, and auricle identities. *Developmental Genetics*, **15**, 401–414.

Sinha, N.R., Williams, R.E. & Hake, S. (1993). Overexpression of the maize homeo box gene, *KNOTTED-1*, causes a switch from determinate to indeterminate cell fates. *Genes and Development*, **7**, 787–795.

Skoog, F. & Miller, C. (1957). Chemical regularion of growth and organ formation in plant fissue cultured in vitro. *Symposia of the Society for Experimental Biology*, **11**, 118–131.

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

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

Smith, L., Greene, B., Veit, B. & Hake, S. (1992). A dominant mutation in the maize homeobox gene, *Knotted-1*, causes its ectopic expression in leaf cells with altered fates. *Development*, **116**, 21–30.

Soltis, D., Soltis, P. & Albert, V. (2002). Missing links: the genetic architecture of flower and floral diversification. *Trends in Plant Science*, **7**, 22–31.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, **98**, 503–517.

Stebbins, G. (1970). Adaptive radiation of reproductive characteristics in Angiosperms, I: pollination mechanisms. *Annual Review of Ecology and Systematics*, **1**, 307–326.

Stevens, P. (2001). Angiosperm phylogeny website. version 9, june 2008 [and more or less continuously updated since].

Stewart, W. & Rothwell, G. (1993). Paleobotany and the Evolution of Plants.

Stout, J., Allen, J. & Goulson, D. (2000). Nectar robbing, forager efficiency and seed set: bumblebees foraging on the self incompatible plant *Linaria vulgaris* (Scrophulariaceae). *Acta Oecologica*, **21**, 277–283.

Stpiczynska, M. (2001). Osmophores of the fragrant orchid *Gymnadenia conopsea* L. (Orchidaceae). *Acta Societatis Botanicorum Poloniae*, **70**, 91–96.

Stpiczynska, M. (2003). Floral longevity and nectar secretion of *Platanthera chlorantha* (Custer) Rchb.(Orchidaceae). *Annals of Botany*, **92**, 191–197.

Stpiczynska, M. & Matusiewicz, J. (2001). Anatomy and ultrastructure of spur nectary of *Gymnadenia conopsea* (L.) orchidaceae. *Acta Societatis Botanicorum Poloniae*, **70**, 267–272.

Stuurman, J., Jaggi, F. & Kuhlemeier, C. (2002). Shoot meristem maintenance is controlled by a GRAS-gene mediated signal from differentiating cells. *Genes and Development*, **16**, 2213–2218.

Sundås-Larsson, A., Svenson, M., Liao, H. & Engström, 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*, **95**, 15118–15122.

Sutton, D. (1988). *A revision of the tribe Antirrhineae*. British Museum (Natural History) Oxford University Press, London Oxford.

Svensson, M. & Engstrom, P. (2002). Closely related MADS-box genes in club moss (Lycopodium) show broad expression patterns and are structurally similar to, but phylogenetically distinct from, typical seed plant MADS-box genes. *New Phytologist*, **154**, 439–450.

Szlachetko, D. & Rutkowski, P. (2000). Gynostemia orchidalium I. Apostasiaceae, Cypripediaceae, Orchidaceae (Thelymitroideae, Orchidoideae, Tropidioideae, Spiranthoideae, Nettioideae, Vanilloideae). *Acta Botanica Fennica*, **169**, 1–380.

Takhtajan, A. (1969). *Flowering plants: origin and dispersal*. Smithsonian, Washington, DC.

Takhtajan, A. (1991). *Evolutionary Trends in Flowering Plants*. Columbia University Press, New York.

Takhtajan, A. (2009). *Flowering Plants*. Springer Verlag, New York.

Takumi, S., Kosugi, T., Murai, K., Mori, N. & Nakamura, C. (2000). Molecular cloning of three homoeologous cDNAs encoding orthologs of the maize *KNOTTED1* homeobox protein from young spikes of hexaploid wheat. *Gene*, **249**, 171–181.

Tamaoki, M., Kusaba, S., Kano-Murakami, Y. & Matsuoka, M. (1997). Ectopic expression of a tobacco homeobox gene, *NTH15*, dramatically alters leaf morphology and hormone levels in transgenic tobacco. *Plant Cell Physiology*, **38**, 917–927.

Tanaka, M., Kato, N., Nakayama, H., Nakatani, M. & Takahata, Y. (2008). Expression of class I *knotted1*-like homeobox genes in the storage roots of sweetpotato (*Ipomoea batatas*). *Journal of Plant Physiology*, **165**, 1726–1735.

Tepfer, S. (1953). Floral anatomy and ontogeny in *Aquilegia formosa* var. *truncata* and *Ranunculus repens*. *University of California Publications in Botany*, **25**, 513–648.

Theissen, G. & Saedler, H. (1999). The golden decade of molecular floral development (1990-1999): a cheerful obituary. *Developmental Genetics*, **25**, 181–193.

Theissen, G., Kim, J. & Saedler, H. (1996). Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. *Journal of Molecular Evolution*, **43**, 484–516.

Theissen, G., Becker, A., Winter, K., Munster, T. & Kirchner, C. (2002). How the land plants learned their floral ABCs: the role of MADS-box genes in the evolutionary origin of flowers. In Q.C.B. Bronk, R.M. Bateman & J.A. Hawkins, eds., *Developmental Genetics and Plant Evolution*, 173–205, Taylor Francis, London.

Thomas, S., Rieu, I. & Steber, C. (2005). Gibberellin metabolism and signaling. *Vitamins and Hormones*, **72**, 289–338.

Timmermans, M., Hudson, A. & Becraft, P. (1999). *ROUGH SHEATH2*: A Myb Protein That Represses *knox* Homeobox Genes in Maize Lateral Organ Primordia. *Science*, **284**, 151–153.

Tioni, M., Gonzalez, D. & Chan, R. (2003). *Knotted1*-like genes are strongly expressed in differentiated cell types in sunflower. *Journal of Experimental Botany*, **54**, 681–690.

Treisman, J., Gonczy, P., Vashishtha, M., Harris, E. & Desplan, C. (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell*, **59**, 553–562.

Truernit, E., Siemering, K.R., Hodge, S., Grbic, V. & Haseloff, J. (2007). A Map of *KNAT* Gene Expression in the *Arabidopsis* Root. *Plant Molecular Biology*, **2**, 10–12.

Tsiantis, M. & Hay, A. (2003). Comparative plant development: the time of the leaf? *Nature Reviews Genetics*, **4**, 169–80.

Tsiantis, M. & Langdale, J. (1998). The formation of leaves. *Current Opinion in Plant Biology*, **1**, 43–48.

Tsiantis, M., Brown, M., Skibinski, G. & Langdale, J. (1999a). Disruption of auxin transport is associated with aberrant leaf development in maize. *Plant Physiology*, **121**, 1163–1168.

Tsiantis, M., Schneeberger, R. & Golz, J. (1999b). The maize *rough sheath2* gene and leaf development programs in monocot and dicot plants. *Science*, **284**, 154–156.

Tucker, S. & Hodges, S. (2005). Floral ontogeny of *Aquilegia*, *Semiaquilegia*, and *Enemion* (Ranunculaceae). *International Journal of Plant Science*, **166**, 557–574.

Uchida, N., Townsley, B., Chung, K. & Sinha, N. (2007). Regulation of *SHOOT MERISTEMLESS* genes via an upstream-conserved noncoding sequence coordinates leaf development. *Proceedings of the National Academy of Sciences*, **104**, 15953–15958.

van der Cingel, N. (1995). An Atlas of orchid pollination: European orchids. Rotterdam, Balkema.

Vargas, P., Rosselló, J., Oyama, R. & Güemes, J. (2004). Molecular evidence for naturalness of genera in the tribe Antirrhineae (Scrophulariaceae) and three independent evolutionary lineages from the New World and the Old. *Plant Systematics and Evolution*, **249**, 151–172.

Veit, B. (2006). Stem cell signalling networks in plants. *Plant Molecular Biology*, **60**, 793–810.

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

Vincent, C. & Coen, E. (2004). A temporal and morphological framework for flower development in *Antirrhinum majus*. *Canadian Journal of Botany*, **82**, 681–690.

Viola, I. & Gonzalez, D. (2006). Interaction of the BELL-like protein ATH1 with DNA: role of homeodomain residue 54 in specifying the different binding properties of BELL and KNOX proteins. *Biological Chemistry*, **387**, 31–40.

Vogel, S. (1998). Remarkable nectaries: structure, ecology, organophyletic perspectives III. *Flora*, **193**, 113–131.

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

Vollbrecht, E., Reiser, L. & Hake, S. (2000). Shoot meristem size is dependent on inbred background and presence of the maize homeobox gene, *knotted1*. *Development*, **127**, 3161–3172.

Waites, R., Selvadurai, H. & Oliver, I. (1998). The *PHANTASTICA* Gene Encodes a MYB Transcription Factor Involved in Growth and Dorsoventrality of Lateral Organs in *Antirrhinum*. *Cell*, **93**, 779–789.

Wallace, A.R. (1867). Creation by law. *Quarterly Journal of Science*, **S140**, 471–486.

Wang, R., Stec, A., Hey, J., Lukens, L. & Doebley, J. (1999). The limits of selection during maize domestication. *Nature*, **398**, 236–239.

Watillon, B., Kettmann, R., Boxus, P. & Burny, A. (1997). *Knotted1*-like homeobox genes are expressed during apple tree (*Malus domestica* [L.] Borkh) growth and development. *Plant Molecular Biology*, **33**, 757–763.

Wendel, J., Cronn, R. & Johnston, J.S. (2002). Feast and famine in plant genomes. *Genetica*, **115**, 37–47.

Whitney, H. & Glover, B. (2007). Morphology and development of floral features recognised by pollinators. *Arthropod-Plant Interactions*, **1**, 147–158.

Whittall, J. & Hodges, S. (2007). Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature*, **447**, 706–709.

Wilkins, C. (2004). *Analysis of Genes Regulating Plant Perception of Cytokinins*. Ph.D. thesis, University of Cambridge, Department of Plant Sciences, UK.

Wilkinson, J. & Crawford, N. (1993). Identification and characterization of a Chlorate-resistant mutant of *Arabidopsis thaliana* with mutations in both Nitrate Reductase and Structural genes *NIA1* and *NIA2*. *Molecular and General Genetics*, **239**, 289–297.

Williams, R. (1998). Plant homeobox genes: many functions stem from a common motif. *BioEssays*, **20**, 280–282.

Williams-Carrier, R.E., Lie, Y.S., Hake, S. & Lemaux, P.G. (1997). Ectopic expression of the maize *kn1* gene phenocopies the *Hooded* mutant of barley. *Development*, **124**, 3737–45.

Xu, L. & Shen, W. (2008). Polycomb silencing of *KNOX* genes confines shoot stem cell niches in *Arabidopsis*. *Current Biology*, **18**, 1966–1971.

Yanai, O., Shani, E., Dolezal, K., Tarkowski, P., Sablowski, R., Sandberg, G., Samach, A. & Ori, N. (2005). Arabidopsis KNOXI proteins activate cytokinin biosynthesis. *Current Biology*, **15**, 1566–71.

Yu, H., Yang, S. & Goh, C. (2000). *DOH1*, a class 1 knox gene, is required for maintenance of the basic plant architecture and floral transition in the Orchid. *The Plant Cell*, **12**, 2143–2159.

Yuan, Q., Hill, J., Hsiao, J., Moffat, K., Ouyang, S., Cheng, Z., Jiang, J. & Buell, C. (2002). Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion. *Molecular Genetics and Genomics*, **267**, 713–720.

Zhang, L., Vision, T. & Gaut, B. (2002). Patterns of nucleotide substitution among simultaneously duplicated gene pairs in *Arabiopsis thaliana*. *Molecular Biology and Evolution*, **19**, 1464–1473.

Appendix A

List of common solutions

1. **Orchid compost (John Haggar)**: 2x botanic gardens compost, 1x Devils dyke soil, 1x potting compost, 0.5x fine grit, 1x super coarse perlite
2. **FAA fixative**: 85% of 70% ethanol, 10% formaldehyde, 5% glacial acetic acid
3. **Kew mix**: 53% IMS, 37% dH_2O, 5% formaldehyde, 5% glycerol
4. **Copenhagen mix**: 70% IMS, 28% dH_2O, 2% glycerol
5. **DNA extraction buffer**: 50mM EDTA, 0.1M Tris HCl pH 8.0, 0.1M NaCl 1% (*v/v*) SDS, autoclave
6. **TE buffer**: 10mM Tris HCl pH 8.0, 1mM EDTA, autoclave
7. **RNA extraction buffer**: 100mM NaCl, 10mM Tris HCl pH 8.0, 1mM EDTA, 1% (*v/v*) SDS, autoclave
8. **1X DNase buffer**: 10mM Tris HCl pH 7.5, 2.5mM $MgCl_2$, 0.5mM $CaCl_2$, autoclave
9. **10X TBE buffer**: 880mM Tris base, 880mM Boric acid, 40mM EDTA pH 8.0
10. **10X Orange-G loading buffer**: 0.5% (*w/v*) Orange-G dye, 10% (*w/v*) Ficoll Type 400, 100mM EDTA, autoclave
11. **6X Bromophenol blue/Xylene cyanole FF loading buffer**: 40% (*w/v*) Sucrose, 0.25% (*w/v*) Bromophenol blue, 0.25% (*w/v*) Xylene cyanole FF, store at 4°C in the dark
12. **SOC medium**: 2% (*w/v*) Tryptone, 0.5% (*w/v*) Yeast Extract, 20mM Glucose, 10mM NaCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 2.5mM KCl, autoclave

13. **LB agar**: 12g/L Bacto-Agar, 10g/L Bacto-Tryptone, 10g/L NaCl, 5g/L Bacto-Yeast extract, autoclave

14. **LB broth**: 10g/L Bacto-Tryptone, 10g/L NaCl, 5g/L Bacto-Yeast extract, autoclave

15. **Phi broth (pH 7.6)**: 20g/L Bacto-Tryptone, 5g/L Bacto-Yeast extract, 5g/L $MgSO_4$, pH to 7.6 with KOH, autoclave

16. **MS9 media (pH 5.9)**: 4.4g/L MS salts, 0.8% (*w/v*) Bacto agar, 2% (*w/v*) sucrose, 1μg/mL 6-benzylaminopurine (BAP), 0.5μg/mL indole-3-acetic acid (IAA), adjust to pH5.9 with KOH, autoclave

17. **MS0 media pH5.9**: 4.4g/L MS salts, 0.8% (*w/v*) Bacto agar, 3% (*w/v*) sucrose, adjust to pH5.9 with KOH, autoclave

18. **Transformation buffer I (TFB I)**: 30mM KAc, 50mM $MnCl_2$, 100mM RbCl, 10mM $CaCl_2$, 15% (*w/v*) Glycerol, pH to 5.8 with 0.2M Acetic acid, filter sterilize, prepare 100ml as necessary

19. **Transformation buffer II (TFB II)**: 10mM NaMOPS pH 7.0, 75mM $CaCl_2$, 10mM RbCl, 15% (*w/v*) Glycerol, pH to 6.5 with KOH, filter sterilize, store at 4°C

20. **Ribonuclease buffer P1**: 25mM Tris HCl pH 8.0, 10mM EDTA, 100μg/ml RNase A, autoclave prior to addition of RNase A, store at 4°C

21. **Lysis buffer P2**: 200mM NaOH, 1% (*w/v*) SDS, autoclave

22. **Precipitation buffer P3**: 3M KAc; pH adjusted to 6.5 with glacial acetic acid, autoclave, store at 4°C

23. **Salty-ethanol**: 95% EtOH, 5% 3M NaAc, pH 5.5 with acetic acid

24. **Denaturation solution**: 1.5M NaCl, 0.5M NaOH

25. **Neutralisation solution**: 1M Tris base, 1.5M NaCl

26. **20X SSC**: 3M NaCl, 300mM Sodium citrate

27. **Pre-hybridisation buffer**: 6X SSC, 0.1% (*w/v*) SDS, 0.02% (*w/v*) PVP, 0.02% (*w/v*) Ficoll Type 400

28. **Hybridisation buffer**: 3X SSC, 0.1% (*w/v*) SDS, 0.02% (*w/v*) PVP, 0.02% (*w/v*) Ficoll Type 400

29. **Wash solution (High/[Low stringency])**: 0.1X [2X] SSC, 0.5% (*w/v*) SDS

Appendix B

List of PCR primer sequences

The primers used in gene identification, cloning and diagnostics are described in the table below. The annealing temperature of each primer (Tm °C) is listed in the final column. Primer sequences are written in the 5′ to 3′ orientation.

Table B.1: List of oligonucleotides primers used

Name	Sequence	Temp (°C)
Vector		
m13F	GTAAAACGACGGCCAG	52
m13R	CAGGAAACAGCTATGAC	50
pGEM-T Easy F	TTAAGTTGGGTAACGCCAGG	57
pGEM-T Easy R	TGTGGAATTGTGAGCGGATA	58
CaMV35S F	CTTCGCAAGACCCTTCCTCT	59
CaMV 35S R	CGGGAAACTACTCACAC	52
T7	TAATACGACTCACTATAGGG	57
SP6	ATTTAGGTGACACTATAG	59
Degenerate PCR		
DQFM F1	CGGACCCGGAGCTGGAYSARTTYATGG	40-55
HYKW R1	CGGTCGGGTACGGCCAYTTRTARTG	40-55
WFIN R1	CCGCTGGTTGATGAACCARTTRTT	40-55
cDNA synthesis		
B26	GACTCGAGTCGACATCGACATCGATTT(16)	42
RACE PCR		
B25	GACTCGAGTCGACATCG	55
GeneRacer 5′RNA adapter	CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUAGAAA	–
GeneRacer oligo dT primer	GCTGTCAACGATACGCTACGTAACGGCATGACAGTGT(24)	–
GeneRacer 5′ RACE	CGACTGGAGCACGAGGACACTGA	74
GeneRacer 5′ nested RACE	GGACACTGACATGGACTGAAGGAGTA	78
GcpL3 3′RACE F1	CTGCTCCTTGACTGGTGGAC	61
GcpL3 3′RACE F2	TGGCCATATCCTACTGAAGCA	57
DfKNOX1 3′RACE F1	GGGCAGCCTCAAGCAAGAAT	59
DfKNOX1 3′RACE F2	GCTGCTCGACTGGTGGACTC	63
DfKNOX1 3′RACE F3	AAGCTTACTGCGAGATGCTGG	52
DfKNOX1 5′RACE R1	TTGAGGAGCGGGAGGAAAGGAAAGG	57
DfKNOX1 5′RACE R2	GCGTCAATGCGAGATAAGAAAAGCATGG	65
DfKNOX2 3′RACE F3	ATTCTTATTGCCGCGTTTTG	53
DfKNOX2 5′RACE R1	GAGGCGGCGGAAGAGGAAGAGATG	67
DfKNOX2 5′RACE R2	CACGCCAAAACGCGGCAATAAGAAT	63
DfKNOX3 3′RACE F1	CACTACTGCTGCTCCCTTCC	61
DfKNOX3 3′RACE F2	AGACGAGAAATGCGAAGGAG	57
DfKNOX3 3′RACE F3	AGGCATACTGCAACATGCTGG	59
DfKNOX3 3′RACE F3(long)	AAAAGAAGAAGAAAGGAAAACTGCCCAAAGAT	63
DfKNOX3 3′RACE F4	ACGGGGCTGGATCAGAAG	58
DfKNOX3 5′RACE R1	CCGAGGAGAGGAAGGGAGCAGCAGT	69
DfKNOX3 5′RACE R2	GCTGCTCCTTGTACTTCACCAGCATGT	66
DfKNOX4 3′RACE F1	CGCCTTTCTCAACACCTC	56
DfKNOX4 3′RACE F2	ATCTCCGATGAGGCTGTTG	56

Continued on next page...

Table B.1 – continued from previous page

Name	Sequence	Temp (°C)
DfKNOX4 3′RACE F3	ATGGAGATGTACTGCGATGTGC	60
DfKNOX4 3′RACE F4	TTGACTGGTGGACAGCTCAC	59
DfKNOX4 5′RACE R1	GTGTTGAGAAAGGCGGTGGCTTCGT	66
DfKNOX4 5′RACE R2	GGCTGCGCCAAATCTCTCCGATACT	66
DvKNOX1 3′RACE F1	GGATACCTGGGCAGTCTCAA	57
DvKNOX1 3′RACE F2	AAGAAGAGGAAGAAAGGGAAGC	58
DvKNOX4 3′RACE F1	GAAGAAGAAGAAGAAAGGGAAGC	58
DvKNOX4 3′RACE F2	AGGCTCGATCAGAAGCAGAT	56
LvHirz3′RACE F3	AGGCTTACTCTGAAATGTTGACCA	59
LvHirz3′RACE F4	ACAAGTGGCCATATCCTTCG	57
LvHirz5′RACE R1	GCTTCTGGTCGAGCCCCGTTGATT	61
LvHirz5′RACE R2	GCCACGAGCTGTACCAGGAGAGGGATGA	66
LvIna 5′RACE R1	GCCGATACTCGCAGCCGATGAACAA	61
LvIna 5′RACE R2	GAGGATGAGCCATGATTTTTGCCTTGA	58
Full-length primers		
AmHirz F1	ATGGAGGGTAGTGGTGGTGG	61
AmHirz R1	TCAAAGAAGTGCAGGCGAG	56
AmInv F1	ATGGAGGGTGGTGTGG	57
AmInv R1	TTAAAGAAGCGTTGGCAAGATATC	57
LvHirz F1	ATGGAGGGTGGCAATAGTAGTAGTAG	73
LvHirz R1	GAGAAGTGCAGGCGAGATATCC	76
LvIna F1	ATGAGTAGTAATAATAATGGA	65
LvIna R1	AAGAAACGATGGCGAGATGTC	74
DfKNOX2 F1	ATGAGAGGGAGGGAAAGGG	58
DfKNOX2 R1	GGGCCGCTGCTCCG	59
VIGS constructs		
NbPDS F1	GGCACTCAACTTTATAAACC	43
NbPDS R1	CTTCAGTTTTCTGTCAAACC	43
Restriction sites		
HindIII	CCCAAGCTTGGG	–
BamHI	CGCGGATCCGCG	–
PstI	AACTGCAGAACCAATGCATTGG	–
EcoRI	GGAATTCC	–
KpnI	CGGGGTACCCCGCTT	–
***In situ* hybridisation probes**		
AtHistoneH4 F1	GAGAGACAACATCCAAGGAATCA	53
AtHistoneH4 R1	GACAACATCCATGGCGGT	50
LvHirzprobe F1	TCAAAAGAGGCAGCACAAGA	50
LvHirzprobe R1	GGTGCCCCTATCTTTTGACA	52
LvInaprobe F1	ACCGAGTGGCAATTCATGTT	50
LvInaprobe F1	CACAAGCTTCCTCCAACCTC	54
Q/RT-PCR primers		
AmHirz RT F1	GGTAGTGGTGGTGGTGGTATG	61
AmHirz RT R1	CGATGGTAATGAGGATGAGC	57
LvHirz Q/RT F1	TGAAATGTTGACCAAATATGAGC	55
LvHirz Q/RT R1	GCAGTTTCCCCTTCTTCCTT	57
LvIna Q/RT F1	TCAAAGCCCTCACTGTCTCA	57
LvIna Q/RT R1	CAACGCCAACCTTTGAGATT	55
DfKNOX1 RT F1	TATCTCGCATTGACGCTCAG	57
DfKNOX1 RT R1	GGAATTCTTGCTTGAGGCTG	57
DfKNOX2 RT F1	CGTCCTTCCTTTCCAGCAT	57
DfKNOX2 RT R1	CCATTGGTAATGGCTGTTCC	57
DfKNOX3 RT F2	ATCTCCCACCCTCAATACCC	57
DfKNOX3 RT R2	CACTTTCTCCGTTTCCGATG	56
Df/DvKNOX4 RT F1	TTTCTCAACACCATGGAGATTCG	61
Df/DvKNOX4 RT R1	ATGGCCACTTGTAGTGAGCTG	57
DvKNOX4 RT F1	TTTCTTATCTCGCATTGATGCTC	59
DvKNOX4 RT R1	GGCCATTTGTAGTGCCGG	59
Control genes		
PACT F1	GGATTTGCTGGAGATGATG	56
PACT R1	AGCAATACCAGGGAACATGG	57
OrACT F1	GCCGTGCTTTCTCTTTATGC	57
OrACT R1	TGAGAGAGGGCTGGAAAAGA	57
Lv 18S F1	GCGGATGTTGCTTTTAGGAC	57
Lv 18S R1	AACCAGACAAATCGCTCCAC	57
TUA5 F1	GGTGTTCAATGCTGTTGGTG	54
TUA5 R1	CTTCAGCACCAACTTCTTCG	56
LvTUA5 F1	AGGTTGTCGGTTGATTACGG	56
LvTUA5 R1	CAGGTGCGTATGAGGAAAGC	57

Appendix C

DNA sequences used in phylogenetic analyses

C.1 DNA sequence data

This section contains DNA sequences, alignments and corresponding programming code implemented for phylogentic analysis using Bayesian Inference with MrBayesv3.1.2 (Huelsenbeck & Ronquist, 2001).

Table C.1: Published DNA sequences were obtained from GenBank and subsequently used in Bayesian phylogenetic analysis. This table details the accession numbers and corresponding species names of the genes characterised (section 4.2.4). Gene names were adopted from GenBank and/or the recent class 1 KNOX phylogenetic analysis by Jouannic *et al.* (2007).

Species	Gene name	Accession no.
Eudicot		
Antirrhinum majus	INVAGINATA	AY072735
	HIRZINA	AY072736.1
Arabidopsis thaliana	STM	NM 104916
	KNAT1	NM 116884
	KNAT2	NM 105719
	KNAT3	NM001036861
	KNAT4	NM 121144
	KNAT6	NM 180620
Brassica napus	BnHD1	Z29073
Brassica oleracea	BoSTM1	AF527947
Glycine max	sbh1	L13663
Helianthus annus	Haknot2	AY096803
	Hakn1	AY096802
Helianthus tuberosus	Htknot1	AJ519674
Ipomoea nil	InPKn1	AB015999
	InPKn2	AB016000
	InPKn3	AB016002
Solanum lycopersicum	TKN1	U32247
	TKN2	U76407
	TKN3	U76408
	TKN4	AF375968
	LeT6	AF000141
	LET12	AF000142
	THox2	U76410
Malus x domestica	MdKN11	Z71978

Continued on next page

Table C.1 – continued from previous page

Species	Gene name	Accession no.
	MdKN12	Z71979
	MdKNAP3	Z71980
Medicago truncatula	MtKn1	AF308454
Nicotiana tabacum	NTH1	AB025573
	NTH9	AB025713
	NTH15	AB004785
	NTH20	AB025714
	NTH22	AB025715
	NTH23	AB004797
Petunia x hybrida	PhSTM1	AY112704
Pisum sativum	Hop1	AF063307
Populus alba x Populus tremula	ARBORKNOX1	AY755413
Populus balsamifera x Populus deltoides	PtdKn2	AY684937
	PtdKn3	AY684938
Populus tomentosa	PtKNAT1	AY660748
Solanum tuberosum	POTH1	U65648
Streptocarpus dunnii	SdSTM1	AY655752
Streptocarpus rexii	SrSTM1	AY655753
Streptocarpus saxorum	SsSTM2	AY655754
Monocot		
Chasmanthium latifolium	CLKN1	DQ317421
Dendrobium grex Madame Thing-In	DOH1	AJ276389
Hordeum vulgare	Hvkn1	AF544045
	Hvh21	AF022390
Oryza sativa	OSH1	AC145380
	OSH3	AB071664
	OSH6	AB028883
	OSH15	AB016071
	OSH43	AB028884
	OSH45	D49704
	Oskn2	AF050180
	Oskn3	AF050181
Panicum miliaceum	PmKN1	DQ317418
Saccharum officinarum	SoKn1	AY781901
Triticum aestivum	TaKnox1b	AF224499
Zea mays	GNARLY1	AY312168
	KNOTTED	X61308
	liguleless3	AF100455
	liguleless4a	AF457118
	liguleless4b	AF457119
	RS1	L44133
Gymnosperm		
Picea abies	PaKn	AF063248
	PaKn2 (HBK3)	AF483277
	PaKn3 (HBK3)	AF483278
Picea mariana	PmSKN1	U90091
	PmSKN2	U90092
	PmKN4	AY680405
Pinus taeda	PtKN1	AY680402
	PtKN2	AY680403
	PtKN3	AY680404
Pteridophyte		
Ceratopteris richardii	CRKNOX1	AB043954
	CRKNOX2	AB043956
Lycophyte		
Selaginella kraussiana	SkKNOX1	AY667449
	SkKNOX2	AY667450
	SkKNOX3	AY667451
Setaria italica	SiKN1	DQ317420
Bryophyte		
Physcomitrella patens	MKN2	AF285147
	MKN4	AF284817
	MKN1-3	AF285148
Green algae		
Acetabularia acetabulum	AaKNOX1	AF170172

Table C.2: Unpublished DNA sequences were obtained from degenerate PCR (section 4.2.1) and subsequently used in Bayesian phylogenetic analysis. This table details the species names of the genes characterised (section 4.2.4). Gene names are arbitrary in most cases pertaining to the order in which different genes/gene fragments were isolated. Gene names of clear orthologues have been named in accordance with the characterised GenBank accession. Full-length genes are indicated by an asterisk (*). DNA sequences for each of the genes is available on-line (see Section C.2)

Species	Gene name
Eudicot	
Linaria vulgaris	LvHirz* (LvKN1)
	LvIna* (LvKN2)
	LvpL103
Monocot	
Dactylorhiza fuchsii	DfKNOX1* (DfKN1)
	DfKNOX2* (DfKN2)
	DfKNOX3 (DfKN3)
	DfKNOX4* (DfKN4)
Dactylorhiza incarnata	DipLB31
	DibpLB11
	DiapL32
	DipL47
	DipL30
Dactylorhiza viridis	DvKNOX1 (DvKN1)
	DvKNOX4 (DvKN4)
	DvpL105
Gymnadenia conopsea	GcpL3pL6
	GcpL24
	GcpLC26
	GcpL52
	GcpLB88
	GcpL89
Gymnadenia odoratissima	GopLA21
	GopLB21
	GopLC84
	GopL85
Gymnadenia rhellicanii	GrpL67
	GrpL68
Orchis anthropophora	OapL18
	OapL35
	OapL49
Orchis italica	OipLA15

Table C.3: A number of published DNA sequences were returned from BLAST queries but not used in subsequent phylogenetic analyses. The accession numbers and corresponding species names of the genes were obtained from GenBank

Species	Gene name	Accession no.
Eudicot		
Ipomoea batatas	IBKN3	AB283029
Kalanchoe daigremontiana	KdSTM-like	DQ674268
Malus x domestica		KNAP2
DT041379		
Monocot		
Dendrobium grex Madame Thong-In	OVG2	AF100326
Dendrobium nobile	DnSTM1	AY608889
Oryza sativa	OSH71	AB028885
	Os05g0129700	NM 001061081
	HOS9	AB007624
Ruscus aculeatus	RaSTM	AB007624

C.2 DNA alignment and Bayesian analysis

DNA alignments and program code used for phylogenetic analyses using Bayesian Inference may be downloaded from: http://www.mediafire.com/myfiles.php
(Login: msb44@cam.ac.uk, Password: S7e3uceh).

Files available on-line:

- BOXPhDthesis2010Class12HDalign.pdf - Total *KNOX* gene alignment, HD encoding domain only
- BOXPhDthesis2010Class12MD-HDalign.pdf - Total *KNOX* gene alignment, MD-HD encoding domains only
- BOXPhDthesis2010Class1HDalign.pdf - Class 1 *KNOX* gene alignment, HD encoding domain only
- BOXPhDthesis2010Class1MDHDalign.pdf - Class 1 *KNOX* gene alignment, MD-HD encoding domains only
- BOXPhDthesis2010UnpubGeneSeq.pdf - All *KNOX* gene DNA sequences identified in this work