Are changes at *ARP* and *KNOX* genes responsible for the evolution of leaf form in *Begonia* section Gireoudia?

Saima Umbreen

Doctor of Philosophy

University of Edinburgh

2011

Declaration

I declare that this thesis is composed of work carried out by myself unless otherwise acknowledged and that this thesis is my own composition. This thesis has not in whole or in part been previously presented for any other degree.

Saima Umbreen

2011

Acknowledgement

First of all I would like to thank Higher Education Commission Pakistan for providing me funds to carry out this research.

I am deeply indebted to my supervisor Dr. Catherine Kidner whose help, stimulating suggestions and encouragement helped me in all the time of research and writing of this thesis.

This project could not have been completed without my second supervisor Professor Andrew Hudson, who has encouraged and challenged me throughout my project.

I am extremely thankful to Professor Neelima Sinha from University of California Davis USA for allowing me to carry out a part of my project in her lab. This was indeed the most learning phase of my PhD. I am especially thankful to Tanya Kumar, Ravi Kumar, Mariane and Brad Townsely for their valuable suggestions and for the friendly atmosphere in lab.

I want to thank to the Royal Botanic Garden Edinburgh staff especially to Frieda Cristae for teaching me SEM.

I am thankful to other people at Edinburgh University for their moral support and encouragement during this project in no particular order Ralf Mueller, Byung-Ho Hwang, Fatima Dhamani, Mobina Shaukat Ali, Andrew Mathew, Rabia Amir, M. Usman and Jo Critchley.

I am extremely thankful to my viva examiners Dr. Justin Goodrich and Dr. Sinead Drea for their valuable suggestions.

Especially, I would like to give my thanks to my husband Imran whose patient love enabled me to complete this work. I am also grateful to my son, Ryaan for being a calm boy which helped me focusing on my thesis.

At last but not least I would like to thank my mother without whose efforts and prayers I would not have been successful in my PhD and in my life indeed.

Abstract

Leaf primordia initiation takes place at the flanks of SAM and then passes through common developmental stages. Very different final leaf shapes and sizes result from varying the timing and further patterning events within these developmental stages. Similar final leaf shapes may also result from very distinct early events. *Begonia* section Gireoudia is a recently radiated group of species with highly divergent leaf forms. I have used a classical genetic approach and candidate gene approach to explain the evolution of leaf form in this genus. These results suggest that convergent evolution of peltate leaves may be through changes at different loci. Key developmental regulators *KNOX* and *ARP* genes are reported to be involved in the evolution of leaf form in different species. I have shown that in at least one species *ARP* is linked to the evolution of peltate leaf form. In a second species there is no link between *STM*-like *KNOX* genes and leaf dissection. Estimates of the rate of evolution of *ARP* CDS showed that different domains of the genes are under different selection pressures. Myb domain2 of *ARP* genes is under positive selection and variable between two copies of *ARP* genes in *Begonia*. Results of complementation tests with *Begonia ARP* genes in *Arabidopsis* show that *ARPs* from *Begonia* are functionally equivalent to *Arabidopsis AS1* genes and one of the two *ARP* genes in *Begonia* may be a dominant negative. Expression analysis based on insitu hybridization in compound, peltate and simple leaved *Begonias* is described. There is no variation in expression patterns between peltate, non peltate or compound leaved *Begonia* species for *BARP1* and *KNB1* genes.

Abbreviations

ARP	*ASYMMETRIC LEAVES1* [*AS1*]/ *ROUGH SHEATH2* [*RS2*]/*PHANTISTICA* [*PHAN*] like genes
AN3	*ANGUSTIFOLIA3*
AS1	*ASYMMETRIC LEAVES1*
AS2	*ASYMMETRIC LEAVES2*
BC1	Backcross population1
bp	Base pairs
BP	*BREVIPEDICELLUS*
BSA	Bovine Serum Albumin
BLAST	Basic Local Alignment Search Tool
CaMV 35S	*Cauliflower Mosaic Virus 35S* constitutive promoter
CAPS	Cleft Amplified Fragment Polymorphisms
CDS	Coding sequences
CZ	Central zone
dH_2O	deionised water
DIG	Digoxigenin
DMSO	Dimethyl sulphoxide
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribo-nucleic acid
dPCR	degenerate PCR
dNTPs	Deoxynucleoside triphosphate mix

EtBr	Ethidium bromide
EDTA	Ethylene di-amino tetra-acetic acid
F1	The first filial Generation
FLO	*FLORICAULA* gene
GA	Gibberellic acid
HD	Homeodomain
HIRA	Histone regulatory homolog A
L1	The outermost meristem layer
L2	The layer subtending the L1
L3	The innermost cell layers of the meristem
LFY	*LEAFY* gene
NEB	New England Biolabs
NTE	Sodium Chloride, tris EDTA buffer
o/e	Over-expressed
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PTS	*PETROSELINUM*
PZ	The peripheral zone of the meristem
RT-PCR	Reverse transcription PCR
QTL	Quantitative trait locus
RNA	Ribo-nucleic acid
RZ	The rib zone of the meristem
SAM	Shoot apical meristem
SDS	Sodium dodecyl sulphate
SEM	Scanning electron microscopy/microscope

SkARP1	*ARP* gene ortholog in *Selaginella kraussiana*
SOC	Super optimal broth
SSC	Saline sodium citrate buffer
STM	*SHOOTMERISTEMLESS*
TFs	Transcription factors
TALE	Three amino acid loop extension
TBE	Tris/Boric acid / EDTA buffer
TE	Tris/EDTA buffer
TM	Melting temperature
TWEEN 20	Polyethylenesorbitan monolaurate
UV	Ultra violet light
WT	Wild type

Chapter 1

Introduction

1.1. The evolution of leaves

All life depends on plants fixing the energy from sunlight into chemical energy. On land most of this activity occurs in leaves. Leaves are optimised to intercept sunlight and absorb CO_2 with minimal water loss and materials cost. Surprisingly they are not synapomorphies of land plants but have evolved in six to eight separate lineages, million years after they move to land. Their delayed evolution may be due to the difficulty of cooling a structure designed to absorb sunlight. Laminas occur in the fossil record only after a dip in CO_2 levels resulted in a greater density of stomata, which allowed more efficient cooling (Beerling *et al.*, 2001).

Leaves exist in a huge range of forms and sizes. These are related to the range of climates and habitats that leaves operate in (Givnish, 1986). Some of this range is due to the different developmental origins of leaves in different lineages. For example in *Selaginella*, leaves are formed from two adjacent epidermal daughter cells and the leaf primordium is of 12-16 cells whereas in ferns leaves are formed from a single apical cell which divides to generate abaxial and adaxial daughter cells for the formation of a lamina (Harrison *et al.*, 2007; White and Turner, 1995). In seed plants the leaf primordium is relatively large structure of a large number of cells formed in the peripheral zone of shoot apical meristem (Kidner and Timmermans, 2010).

Most analysis of leaf developmental genetics has focused on Angiosperm leaves. Despite the great diversity of forms the early developmental processes are well conserved throughout Angiosperms (Kidner and Timmermans, 2010; Barkoulas *et al.*, 2008; Micol, 2009; Hay and Tsiantis, 2009; Kidner and Umbreen, 2010; Koening and Sinha, 2010).

1.2. The Shoot Apical Meristem (SAM)

The Shoot Apical Meristem (SAM) is a group of cells at the very tip of the shoot, which give rise to the whole shoot of the plant and lateral organs. Most SAMs are circular in outline when seen from above while they may be convex, flat or concave when seen from the side. The SAM has both radial and vertical structure. Considering radial structure, the meristem typically consists of; the peripheral zone (PZ), the basal zone or rib zone (RZ) and the central zone (CZ) (Figure 1). These different regions accomplish different essential tasks. The PZ is responsible for the formation of lateral organs, the RZ produces internode and the CZ is composed of stem cells and maintains the indeterminate nature of SAM by replenishing the cells recruited by the PZ and RZ (Bowman and Eshed, 2000).

In most of the angiosperms the SAM can also be divided into three clonally distinct layers (cells within these layers have characteristic fates); the epidermal layer (L1), the sub epidermal layer (L2) and the corpus (L3). The PZ, CZ and lateral organs produced all are composed of cells from these three layers (Figure 1). The developmental process is coordinated through the mutual communication of these three clonally distinctive layers (Vernoux *et al.*, 2010).

Several genes have been characterized by mutations that disrupt shoot apical meristem structure and/or activity. They include *SHOOTMERISTEMLESS* (*STM*), *WUSCHEL* (*WUS*), *CLAVATA1* (*CLV1*) and *CLAVATA3* (*CLV3*). *WUS* is a homeodomain protein of the WOX class (Haecker *et al.*, 2004). *WUS* expression at the early globular stage provides the first indication of shoot apical meristem development. WUS activity interferes with hormone signalling particularly with cytokinins (Gordon *et al.*, 2009). *STM* (Class-1 *KNOTTED*-like homeobox (*KNOX*) is expressed in the late globular embryo, and *CLV1* (a member of leucine-rich receptor kinases family (Clark *et al.*, 1997)), and *CLV3* (encodes a small secreted protein postulated to be the ligand for *CLV1* (Trotochaud *et al.*, 2000)) are expressed in the early heart-shaped embryo at the site of the presumptive SAM. *CLV* and *WUS* are important for regulating SAM size (Leyser and Furner, 1992; Clark *et al.*, 1993, 1995) and *STM* is required for SAM maintenance (Barton and Poeting, 1993).

In the post embryonic SAM, *CLV1* is expressed in the corpus of the central zone while *CLV3* is expressed in the tunica of the central zone (Fletcher *et al.*, 1999) and *WUS* is expressed in a small group of cells in the central domain. The binding of *CLV1* to *CLV3* limits the spread of *CLV3* (Lenhard and Laux, 2003). *CLV3* moves a short distance and causes the down regulation of *WUS* expression. *WUS* promotes the activity of its repressor *CLV3* and establishes a feed-back loop which is important to maintain a relatively constant cell number in SAM (Figure 2) (Brand *et al.*, 2000; Schoof *et al*., 2000).

The expression of *WUS*, *CLV1*, *CLV3* and *STM* is independent at the embryonic stage, but their expression is interdependent at post embryonic stages as complete loss of function *stm* mutants lack SAM and the expression of *CLV1*, *CLV3* and *WUS* is absent from the shoot apex of *stm* mutants. Similarly, *STM* expression is absent from terminated apex of *wus* mutants and *STM* expression spreads throughout the enlarged SAM of *clv3* mutants (Clark *et al.*, 1996).

1.3. Leaf Initiation

Leaves initiate on the flank of SAMs. The time course of leaf development is described in units of plastochrons which is the interval of time between the emergences of two successive leaf primordia. The protruding flank of SAM that precedes the emergence of leaf primordium is designated the *foliar buttress* (Plastochron 1). Leaf primordium initiation begins by the localization of growth activity at the apex of the foliar buttress but as the primordium arises growth expands over the apical surface incorporating additional tissues into it. The meristematic activity of the leaf apex ceases after primordium attains a certain length which varies widely between different species. For example in *Solanum tuberosum* the apical meristematic activity continues until the primordium reaches a length of 200 um, in *Nicotiana tabacum* with a length of 3 mm, and *Angelica archangelica* with a length of 15mm (Leyser and Day, 2003). There is also a considerable variation in the size relationship between a leaf primordium and the meristem that produces it (Furner and Pumfrey, 1992; Irish and Sussex, 1992; Leyser and Day, 2003).

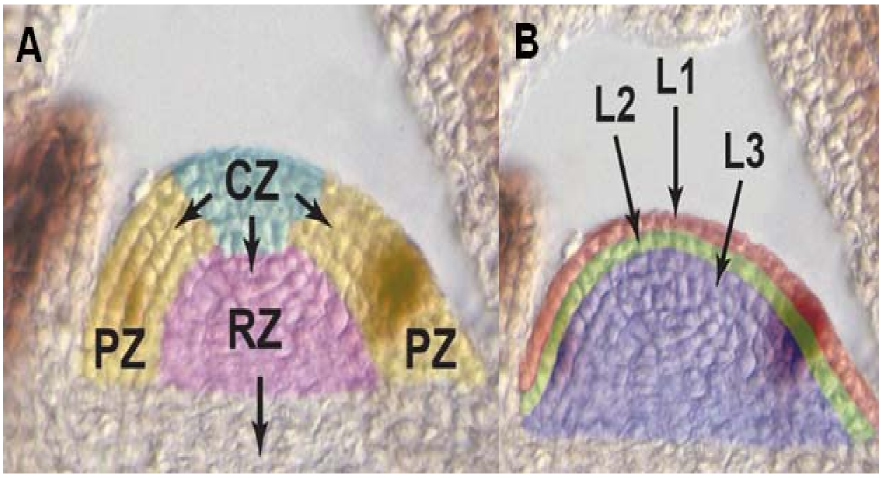


Figure 1 The shoot apical meristem.
(A) Division of SAM on the basis of cytoplasm density and cell division rate, (B) three distinctive layers of the SAM (Bowman and Eshed, 2000).

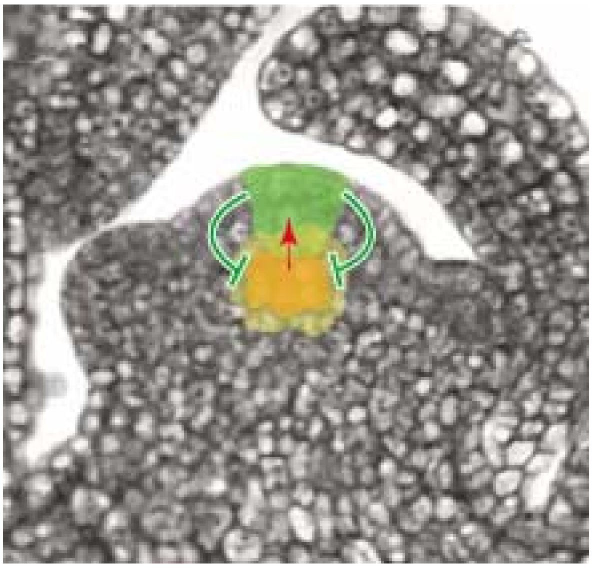


Figure 2 The gene expression patterns underlying meristem organization.
In the shoot *CLV3* (green) is expressed in the central zone in the upper layers. *CLV1* is expressed sub terminally (yellowish green) and overlaps with *WUS* expression (yellow). *WUS* promotes *CLV3* function in the cells above it (arrow) (Picture taken from Byrne *et al*., 2003).

For example in most eudicots the primordium is large and its initiation involves a sizeable portion of the meristem which bulges out and is recognized as a foliar buttress. However there is no foliar buttress stage in some monocotyledons because the primordium is small in size and the first emergence is recognized as a distinct leaf primordium.

Though leaf primordium initiation is characterized by an increase in the cell division at the shoot apex, the plane of cell division is of greater significance than the increase in the rate of cell division (Lyndon, 1998). Further, leaf development is not restricted simply to superficial regions of SAM but involves its full depth. Leaf primordium intiation in *Arabidopsis* is first visible by the periclinal cell division in the L2 layer of SAM (Medford *et al*., 1992). The adxial and abaxial epidermal cells arise from anticlinal division of the L1 layers and inner cells of the leaf originate through anticlinal and periclinal cell divisions from the L2 layer (Furner and Pumfrey, 1992).

The earliest indicator of leaf initiation is the down regulation of class 1 *KNOTTED LIKE* homeobox (*KNOX*) genes (detailed description is on page 7) such as *STM* at the sites of leaf initiation. *KNOX* genes maintain the SAM indeterminacy and their down regulation is important for the leaf to acquire determinacy (Jackson *et al*., 1994; Long *et al*., 1996). However transgenic up regulation of *KN1* in tobacco plants does not affect the normal leaf initiation process suggesting that down regulation of *KNOX* genes alone is not sufficient for normal leaf initiation (Sinha *et al*., 1993). Several genes are expressed at the sites of leaf initiation and *ARP* (*ASYMMETRIC LEAVES1* [*AS1*]/ *ROUGH SHEATH2* [*RS2*]/*PHANTISTICA* [*PHAN*]) genes is one of them ((Byrne *et al*., 2000; Timmermans *et al*., 1999; Waites and Hudson, 1995) and this gene is involved in the down regulation of *KNOX* genes. (*KNOX* and *ARP* genes are discussed in detail later in this chapter).

The presence of the indolic hormone auxin at sites of leaf initiation is critical. Auxin delivery is determined by the auxin efflux facilitator *PIN-FORMED1* (*PIN1*) (Heisler *et al*., 2005). It has been proposed that the auxin and *ARP*/*KNOX* leaf development pathways may be interconnected. At the SAM *PIN1* and *KNOX* genes are expressed in nearly complementary patterns. *KNOX* and auxin activities promote leaf–meristem delimitation by acting in a feedback mechanism. Compromising auxin signalling or *PIN1* activity results in misexpression of *KNOX* genes in leaves and loss of *KNOX* genes activity partially suppresses the inability of *pin1* mutants to generate lateral organs (Hay *et al*., 2006, Hay *et al*., 2010). EXPANSINs (extracellular proteins found in plant cell wall) are expressed at the leaf initiation sites and the application of beads coated with EXPANSIN on SAM generate ectopic leaves supporting the involvement of cell wall loosening mechanism in the leaf initiation process (Flemming *et al*., 1997; Reinhardt *et al*., 1998).

1.4. Establishment of leaf sub domains

Leaves differ from the SAM in two aspects; a leaf is determinate and has a limited potential for development while the SAM is indeterminate; secondly, a leaf is asymmetrical. Three axes of asymmetry are the proximal-distal, medial-lateral and abaxial-adaxial. The proximal-distal axis is defined by the attached (proximal) and free (distal) ends. Asymmetry in leaf development along this axis is inherent and results in proximal petiole and distal blade. The medial-lateral axis is from the midrib to the edge of the leaf blade. The adaxial-abaxial axis runs from the top side of the leaf (adaxial) to the bottom side (abaxial). The adaxial/abaxial axis is acquired in the early stages of leaf initiation. The adaxial or top surface is specialized for light capture and abaxial or bottom for gas exchange (Figure 3).

The adaxial side of the leaf is physiologically and developmentally different from abaxial side. In most of the higher plants xylem is towards the adaxial side and the phloem is towards abaxial side. Palisade mesophyll cells are located towards adaxial side and spongy mesophyll cells are located towards abaxial side. Abaxial-adaxial polarity is of primary importance in leaf patterning. As well as specifying the correct cell types for each side, it provides signals for proper lamina growth (Waites and Hudson, 1995). *PHANTASTICA* (*PHAN*) genes are required for establishing the adaxial domain of the leaf as *phan* mutants generate needle like or cup shaped leaves with reduced adaxial domains (Waites and Hudson, 1995; Waites *et al*., 1998).

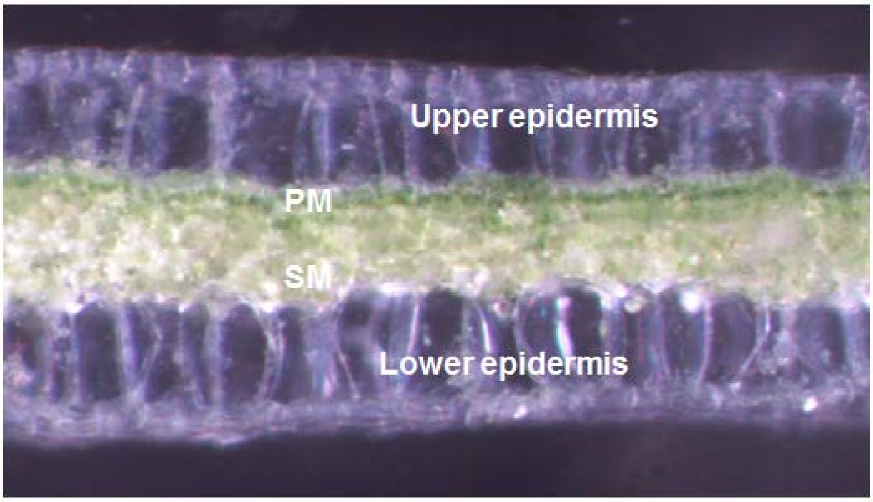


Figure 3 Transverse section of *B. stigmosa* leaf.
Transverse section through *B. stigmosa* leaf. The adaxial and abaxial sides of the leaf have distinct cell types. Palisade mesophyll (PM) cells are present on adaxial side and spongy (SM) mesophyll are present on abaxial side.

However *PHAN* is not important for the leaf initiation processes because the leaf primordia of *phan* mutants resemble wild type at emergence. The *lbl1* (*Leaf Bladless1*) gene encodes a key component in the *trans*-acting small interfering RNA (ta-siRNA) biogenesis pathway which specifies the adaxial cell identity within leaves and leaf like lateral organs in maize (Timmermans *et al*., 1998; Nogueira *et al*., 2007). The *lbl1* also plays a direct or indirect role in the down regulation of *homeobox* gene, *KNOTTED1* during leaf development (Timmermans *et al*., 1998). The *lbl1* and *phan* mutants show that juxtaposition of adaxial and abaxial cell types is essential for lamina formation in both monocots and eudicots.

Homeodomain-leucine zipper (*HD-ZIP*) family genes; *REVOLUTA* (*REV*), *PHABULOSA (PHB*) and *PHAVOLUTA (PHV)* also regulate adaxial-abaxial polarity. The semi-dominant *phabulosa1-D* mutant forms adaxialized leaves with an additional axillary meristems around their bases (McConnell and Barton, 1998), *phavoluta* mutants show very similar phenotypes to that of *phabulosa* mutant (McConnel *et al*., 2001) and *REVOLUTA* plays an important role in the meristematic activity of SAM and leaf primordia (Talbert *et al*., 1995) and is also involved in patterning vascular tissues (Zhong *et al.,* 1999). The three *HD-ZIPIII* genes are post transcriptionally regulated by micro RNA 165/166 (miRNA 165/166) (Kidner and Martienssen**,** 2004; Chitwood *et al*., 2007**).**

Four *LITTLE ZIPPER* (*ZPR1-4*) proteins also regulate the activity of *HD-ZIPIII* genes through a negative feedback loop (Kim *et al*., 2008). *HD ZIPIII* activity upregulated the ZPRs and ZPR over-expression phenotypes in *Arabidopsis* resemble those seen when *HD-ZIPIII* expression is reduced. The ZPR proteins form heterodimers with the HD ZIPIII proteins and prevent it from binding with DNA (Wenkel *et al*., 2007).

KANADI genes are important for abaxial leaf identity regulation and ectopic expression of *KANADI* produces ectopic abaxial cell fate (Kerstetter *et al*., 2001). *HD-ZIPIII* genes and *KANADI* genes are expressed in complementary patterns and interact antagonistically (Kerstetter *et al*., 2001; Eshed *et al*., 2004; Candela *et al*., 2008; Zhang *et al*., 2009). *HD-ZIPIII* genes are ectopically expressed in loss of function *kan1 kan2* mutants whereas overexpression of *KANADI* genes leads to loss of *HD-ZIPIII* genes expression with complete abaxialization of leaf blade (Eshed *et al*., 2004; Kerstetter *et al*., 2001; Candela *et al*., 2008). Conversely over expression of *HD-ZIPIII* genes results in a decrease in *KANADI* genes expression.

KANADI genes may also interact antagonistically with the *AS* (*Asymmetric leaves* genes) pathway as *KANADI* proteins down regulate the expression of *AS2* on the abaxial side of the leaf and ectopic expression of *AS2* genes may result in down regulation of *KANADI* genes (Wu *et al*., 2008). Loss of *AS2* in the abaxial domain is important for restricting the *AS1*/*AS2* interaction to the adaxial side of the leaf only for promoting adaxial cell fate (Kidner and Timmermans, 2010). A balance between the expression of adaxial and abaxial cell determinants is crucial for generating and maintaining the boundary between adaxial and abaxial domains in the developing leaf which is promoted by the antagonistic interactions of the KAN proteins with *AS2* and the *HD-ZIPIII* genes (Kidner and Timmermans, 2010).

YABBY genes (transcription factors) promote the lamina growth through the differentiation of abaxial tissues. Loss and gain of *YABBY* gene function dramatically affects the plant morphology. Loss of *YABBY* function result in loss of abaxial cell types; conversely gain of *YABBY* function results in abaxial cell types differentiating in adaxial positions (Golz *et al*., 2004; Siegfried *et al*., 1999). Figure 4 describes the network of genes involved in leaf polarity establishment.

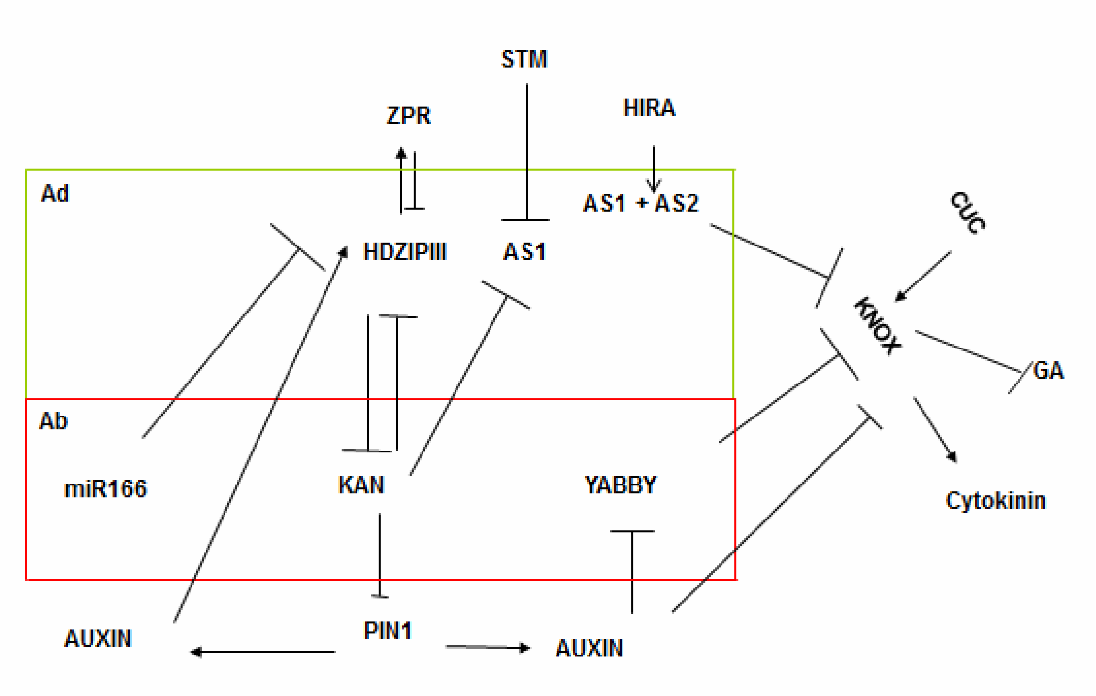


Figure 4 Pathways involved in the regulation of leaf polarity.
Class 1 *KNOX* genes up regulate the cytokinin in SAM which promotes SAM activity and down regulate the Gibberellic acid (GA) in SAM and confines its activity to the leaf primordia. *STM* keeps the *AS1* genes in off state in SAM. Class 1 *KNOX* genes are repressed at leaf initiation sites and *ARP* genes maintain the repression of *KNOX* genes in most of the simple leaved species through a novel epigenetic mechanism in which *ARP* genes make a complex with *AS2* genes through an interaction with HIRA (histone regulatory homolog A), which directly binds to the promoter of *BP* (*BREVIPEDICELLUS*) like *KNOX* genes and maintain *KNOX* gene silencing. *CUC* genes are expressed at the SAM-leaf boundaries and may be unregulated *KNOX* genes in compound leaved species. Adaxial abaxial polarity is maintained by the expression of domain specific genes and antagonistic relationship of these different genes.

1.5. The genetics of leaf form variation

Leaf form affects the plant's response to temperature, water availability, light, wind and herbivores/pathogens (Kidner and Umbreen, 2010). The many factors that affect leaf function suggest that many adaptive peaks for leaf form could exist in any one habitat (Vogel, 2009). It has also resulted in many genera with widely divergent leaf forms, such as *Ipomoea*, *Tropaloeum*, *Violats*, *Solanum* and *Begonia* (Landon *et al*., 1978; McLellan, 1990; Bright *et al*., 2008; Whaley and Whaley, 1942; Kimura *et al*., 2008). These groups and other model species have been used to study the genetics underlying leaf form diversity.

Early studies of the genetics of leaf form showed that it does not usually segregate as a Mendelian trait and that heterosis is common (Melville, 1960; Gottleib, 1986). This may be related to the difficulty of quantifying leaf form. The cases where one or two loci have been identified are mostly easily scored cases of entire versus divided leaves. Table 1 shows a list of these.

1.6. Candidate genes for leaf form variation

Expression studies and functional analysis using forward and reverse genetics have identified a small number of gene families which are associated with variation in leaf form between species (table 2). These key components of leaf form variation and their interactions (Figure 4) have been described in many recent papers (Kidner and Timmermans, 2010; Barkoulas *et al*., 2008; Micol, 2009; Hay and Tsiantis, 2009; Kidner and Umbreen, 2010; Koening and Sinha, 2010).

Table 1 Leaf traits whose variance between or within natural populations/species is controlled by few loci (Kidner and Umbreen, 2010).

Dominant Trait	**Dominant Trait**	**No. of Loci**	**Species**	**Variation level**	**Reference**
Lobed leaves	Dominant to entire leaves	1	*Lactuca sariola*	Intraspecific	Durst, 1929
Lobed leaves	Dominant to entire leaves	2	*Lactuca sativa*	Intraspecific	Durst, 1929
Lobed leaves	Dominant to entire leaves	1	*Carthamus flavescens X C. tinctorius*	Interspecific	Imrie and Knowely, 1970
Lobed leaves	Dominant to entire leaves	1	*Lactua graminifolia X L. canadensis*	Interspecific	Whitaker, 1944
3 leaflets	Semi dominant to 5-7 leaflets	Few	*Potentilla glandulosa*	Intraspecific	Clausin *et al*., 1947
Highly dissected leaves	Semi dominant to less dissected leaves	1	*Solanum cheesemaniae X S. galapagenese*	Intraspecific	Kimura *et al*., 2008
Lobed leaves	Dominant to entire leaves	2	*Capsella bursa pastorilis*	Intraspecific	Shull, 1909
Three-Lobed leaves	Dominant to entire leaves	1	*Ipomoea bederacea*	Interspecific	Elmore, 1986
Deeply-Lobed leaves	Dominant to weekly lobed leaves	1	*Crepis tectorum*	Interspecific	Andersson, 1991, 1994
Peltately compound leaves	Dominant to entire leaves	1	*Vitis spp*	Interspecific	Peter and Prins, 2008
Shallow serrations	Dominant to deep serrations	1	*Urtica pilifera X U. dioica*	Interspecific	Correns, 1928
Rounded leaves	Dominant to acute leaves	1	*Tropaeolum majus X Tropaeolum peltophorum*	Interspecific	Whaley and Whaley, 1942
Entire, orbicular	Dominant to five-lobed	1	*Tropaeolum majus X Tropaeolum peltophorum*	Interspecific	Whaley and Whaley, 1942

Table 2 Candidate genes in leaf form.

Name	Function	Role	Reference
KNOX	Homeobox TF	Required and sufficient for divided lamina though continued cell division.	Barathan *et al*., 2000; Kimura *et al*., 2008; Hay and Tsinaits 2006
ARP	Myb TF	Expression associated with lamina outgrowth.	Waites *et al*., 1998; Kim *et al*., 2003; Golz *et al*., 2004
UNI	TF	Required and sufficient for a divided lamina though continued cell division in Legumes.	Hofer *et al*., 2001
CUC	*NAM* TF	Required and sufficient for divided leaf margins.	Blein *et al*., 2008
YABBY	TF	Expression associated with lamina outgrowth. Genetically associated with lamina outgrowth and medio-lateral polarity	Gleissberg *et al*., 2005; Yamaguchi *et al*., 2010

1.6.1. *KNOX* genes

Knotted like homeobox (*KNOX*) genes belong to the TALE (three amino acid loop extension) family of homeodomains. KNOX proteins consist of a MEINOX domain, GSE box, ELK domain and a Homeodomain. The MEINOX domain is closely related to MEIS (Myeloid ecotropic viral integration site) proteins in humans (Burglin, 1997). The MEINOX domain is split into *KNOX*1 and *KNOX*2 sub domains and may be involved in protein dimerization and interaction with BELL (BEL-like homeodomains) transcription factors (Bellaoui *et al*., 1998; Burglin, 1998; Nagasaki *et al*., 2001; Bhatt *et al*., 2004; Chen *et al*., 2003; Cole *et al*., 2006; Muller *et al*., 2001; Smith *et al*., 2002). The GSE domain is a protein degradation signal and the ELK domain is responsible for nuclear localization (Cole *et al*., 2006). The Homeodomain is involved in DNA binding (Gehring *et al*., 1994; Treisman *et al*., 1989). Figure 5 shows the position of different domains of *STM*- like *KNOX* proteins in *Begonia*.

There are two families of *KNOX* genes in plants which differ in structure (class 2 has an intron in the ELK domain which is absent in class 1) and expression pattern, and probably have different developmental roles (Kerstetter *et al*., 1994; Kerstetter and Hake, 1997; Bharathan *et al*., 1999; Reiser *et al*., 2000). Both class 1 and class 2 *KNOX* genes have been amplified from Bryophytes (Champagne *et al*., 2001), Ferns, Gymnosperms (Sundas-Larsson *et al*., 1998) and Angiosperms. *KNOX* genes from *Acetabularia acetabulum* (green algae) and *Chlamydomonas* have features of both class1 and class 2 *KNOX* genes (Serikawa and Mandoli, 1999; Hake *et al*., 2004). The duplication of *KNOX* genes occurred after the formation of the land plant lineage 500 MYA, but before the divergence of the Bryophytes from the Euphyls 400 MYA (Hake *et al*., 2004). There are multiple class 1 and class 2 *KNOX* genes in flowering plants (Figure 6) (Champagne *et al*., 2001; Bharathan *et al*., 1999; Reiser *et al*., 2000; Mukherjee *et al*., 2009).

KNOX genes are involved in meristem formation and maintenance, internode elongation and patterning within SAM to position lateral organ primordia in *Arabidopsis* and maize (Barton and Poething 1993, Jackson *et al*., 1994; Endrizzi *et al*., 1996; Long *et al*., 1996; Vollbrecht *et al*., 2000, Floyd and Bowman, 2006). Class1 *KNOX* genes prevent the differentiation of cells in the SAM by excluding the activity of the growth factor gibberellin (Sakamoto *et al*., 2001) and inducing the production of cytokinin (Frugis *et al*., 2001). The production of cytokinin in the meristem promotes SAM activity while gibberellin exclusion from meristem confines its activity to the leaf primordia where it promotes leaf growth (Shani *et al*., 2006; Hay & Tsiantis, 2006).

KNOX genes are down regulated at the leaf initiation sites (Harevan *et al*., 1996; Hay & Tsiantis, 2006; Bharathan *et al*., 2001; Shani *et al*., 2009). In simple leaved species such as in *Arabidopsis thaliana*, maize (*Zea mays*), rice (*Oryza sativa*), tobacco (*Nicotiana tobacum*) and snapdragon (*Antirrhinum majus*) *KNOX* genes stay off throughout leaf development, however in compound leaved species *KNOX* expression returns as the primordia grows out and leaflets begin to develop

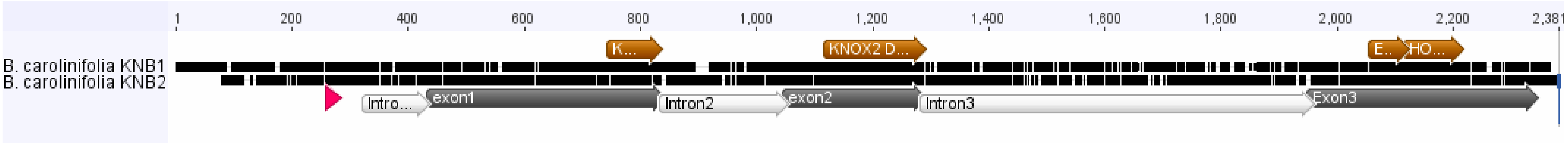


Figure 5 Domains of *STM*-like *KNOX* genes in *B. carolineifolia*.
The brown arrows indicate the *KNOX*1, *KNOX*2, ELK and Homeodomains of *KNOX* genes, Light grey arrows indicates introns and dark grey arrows indicates exons, Pink arrow represent K-BOX, a conserved promoter elements of *STM*-like *KNOX* genes.

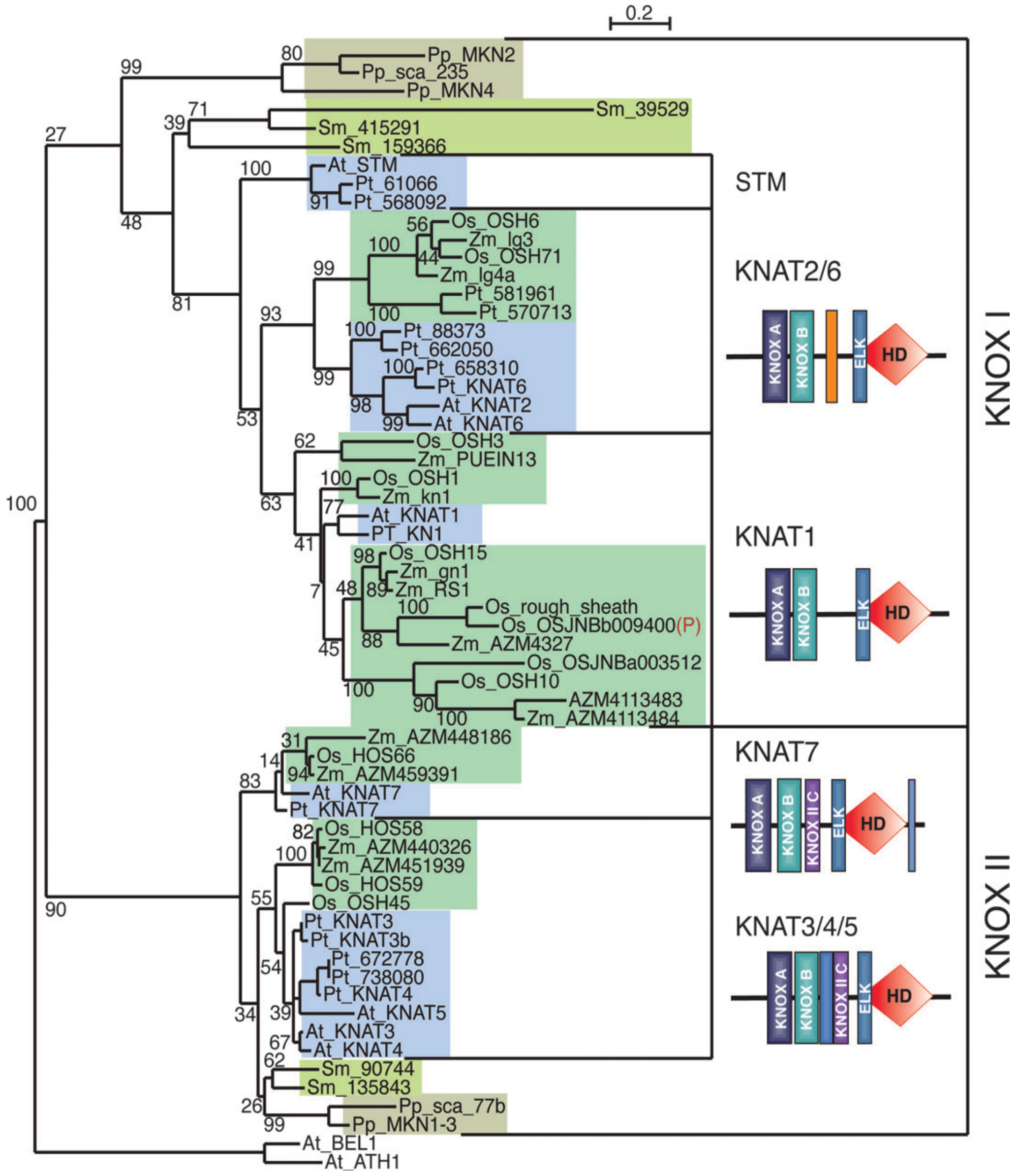


Figure 6 Phylogenetic relationships of *KNOX* genes in land plants (Mukherjee *et al*., 2009).

Maximum Likelihood tree obtained using the homeodomain and codomain sequence of KNOX proteins showing Class1 and Class2 *KNOX* genes. Each of these classes can be further subdivided into two subfamilies having members conserved in both monocots (lightgreen boxes) and eudicots (light-blue boxes). Selaginella (Sm) proteins are shown in yellowish green–colored boxes, whereas moss (Pp) proteins are shown inside the gold-colored boxes. The tree has been rooted with BELL class protein representatives (Mukherjee *et al*., 2009).

(Harevan *et al*., 1996; Hay & Tsiantis, 2006; Bharathan *et al*., 2001; Shani *et al*., 2009). This pattern is seen even in species where secondary growth results in an entire leaf developing from a dissected primordium. In *Lepidium oleraceum KNOX* genes are expressed in the complex leaf primordium which undergoes secondary morphogenesis to form simple leaves (Bharathan *et al*., 2002).

Expression of *KNOX* in the leaf is even seen in species where compoundness is conferred by cell death rather than by differential growth. *Elaeis guineensis* has pinnately compound leaves formed as when the lamina subdivides along folds called plications fairly late in development. A class1 *KNOX* gene, *EgKNOX1* is expressed at plications as well as in the meristem. Transgenic experiments showed that *EgKNOX1* was orthologous to *STM* in *Arabidopsis* (Jouannic *et al*., 2007).

Class1 *KNOX* expression in leaves is also seen in unusual growth forms where the leaf is indeterminate. *Welwitschia mirabilis* has two photosynthetic leaves which continue to grow from a basal meristem for 400 to 1500 years. *KNOX* genes are initially down regulated in the leaves but subsequently their expression increases in the basal regions of the leaves indicating a role in the basal meristem of the leaf (Pham and Sinha, 2003).

Class1 *KNOX* genes are also expressed in the blades of *Streptocarpus* species which grow continuously from basal meristems and produce new leaves or inflorescences from groove meristems (Harrison *et al*., 2005; Nishi *et al*., 2010). Similarly in Podostemoideae which lacks shoot apical meristems and generates new leaves from the base of older leaves, *STM* is expressed in the initiating leaf/bract primordium and becomes restricted to the basal part of older leaf/bract primordia. These results suggest that *KNOX* genes are creating an indeterminate niche at the base of older leaves similar to SAM (Katayama *et al*., 2010).

As well as being associated with compound leaves, *KNOX* expression in the leaf primordia is sufficient to create a more divided leaf. Ectopic expression of *KNOX* in simple leaves species leads to continued cell division in the leaf lamina producing knots of cells, or lobes (Muller *et al*., 2006).

Taraxacum officinale, Web. (Dandelion) has deeply lobed simple leaves which form from secondary morphogenesis of complex leaf primordia which express *KNOX* genes. Constitutive expression of barley *KNOX* genes (*bkn1* and *bkn3)* in addition to the normal endogenous Dandelion *KNOX* genes converts the dandelion simple leaves into compound leaves. On the basis of this the author proposed that the evolutionary step from simple to compound leaves could occur through increased *KNOX1* gene expression (Muller *et al*., 2006).

Changes in *KNOX* genes are also responsible for species-level differences in leaf form in the native tomatoes of Galapagos Islands. A single base pair deletion in the promoter (1266 bp upstream of the open reading frame) of *KNOX* like gene *PETROSELINUM* (*PTS*) up regulated the expression of the *KNOX* genes in the leaves of highly dissected *Solanum galapagense* in comparison to its expression levels in the less dissected sister species *Solanum cheesmaniae* (Kimura *et al*., 2008).

Differential expression of *KNOX* genes between pinnately compound leaved *Cardamine hirsuta* and simple leaved *Arabidopsis thaliana* corresponds to the natural variation in the leaf form of these two closely related species. Promoter swap experiments demonstrate that the difference in the expression pattern of *KNOX* genes between these two species is driven by the variation in the promoter regions. The promoters of the class 1 *KNOX* genes from *A. thaliana* (*SHOOTMERISTEMLESS* and *BREVIPEDICELLUS*) were attached to reporter genes and introduced into both *A. thaliana* and *Cardamine hirsuta*. The reciprocal experiment was set up with the promoters of the corresponding *Cardamine hirsuta* genes. This allowed the comparison of the expression patterns of the *Cardamine hirsuta* promoters with those of the *A. thaliana* promoters for each gene in both species. The *A. thaliana* promoters drove the expression in the meristem only; conversely the *Cardamine hirsuta* promoters drove expression in both the meristem and in the leaves. Thus, the promoter provides the information necessary to determine *KNOX* genes expression in leaves and generate different leaf forms (Hay & Tsiantis, 2006).

The K-box, a conserved non coding sequence (CNS) of the *STM* promoter controls the repression of *KNOX* genes in many simple leaved species. This conserved element is present in a region between 105-271bp upstream of the translation start codon of simple leaved species (asparagus, yucca, grape, poplar, cotton, *Arabidopsis*, tobacco and snapdragon) and compound leaved species (palm, acacia, tomato and ash) (Figure 7). A GUS reporter gene was fused with the promoters of *STM* genes from each species with or without K-box and introduced into *Arabidopsis* and tobacco. The K-box was unable to repress *STM* at incipient leaf primordia and has no role in *STM* expression in the SAM but regulates the persistent repression and/or the reestablishment of *STM* expression in the developing leaves (Uchida *et al.*, 2007).

KNOX genes are good candidates for studying the natural diversity in leaf form as they have proved to be the key genes controlling natural variation in leaf form between *Arabidopsis* and *Cardamine hirsuta* and in *Solanum galapagense* (Hay & Tsiantis, 2006; Kimura *et al.*, 2008). However class 1 and class 2 *KNOX* gene duplication occurred 400-500 MYA followed by many subsequent duplications in monocots and eudicots independently which could lead to uncertainty in establishing orthology. For example the *SHOOTMERISTEMLESS* (*STM*) gene of *Arabidopsis* has many similarities with that of maize *KN1* gene. Both *STM* and *KN1* genes are significantly similar at the protein level and both are expressed specifically in meristems and are down-regulated in the P0 in *Arabidopsis* and maize respectively. However *STM* and *KN1* are not orthologs by phylogenetic studies as they both resolve into different clades (Hake *et al.*, 2004). So far orthologs of *STM* have been only established in eudicots while orthologs of *kn1* are found in monocots only. This indicates that all of the *KNOX* genes may not have been identified in different lineages.

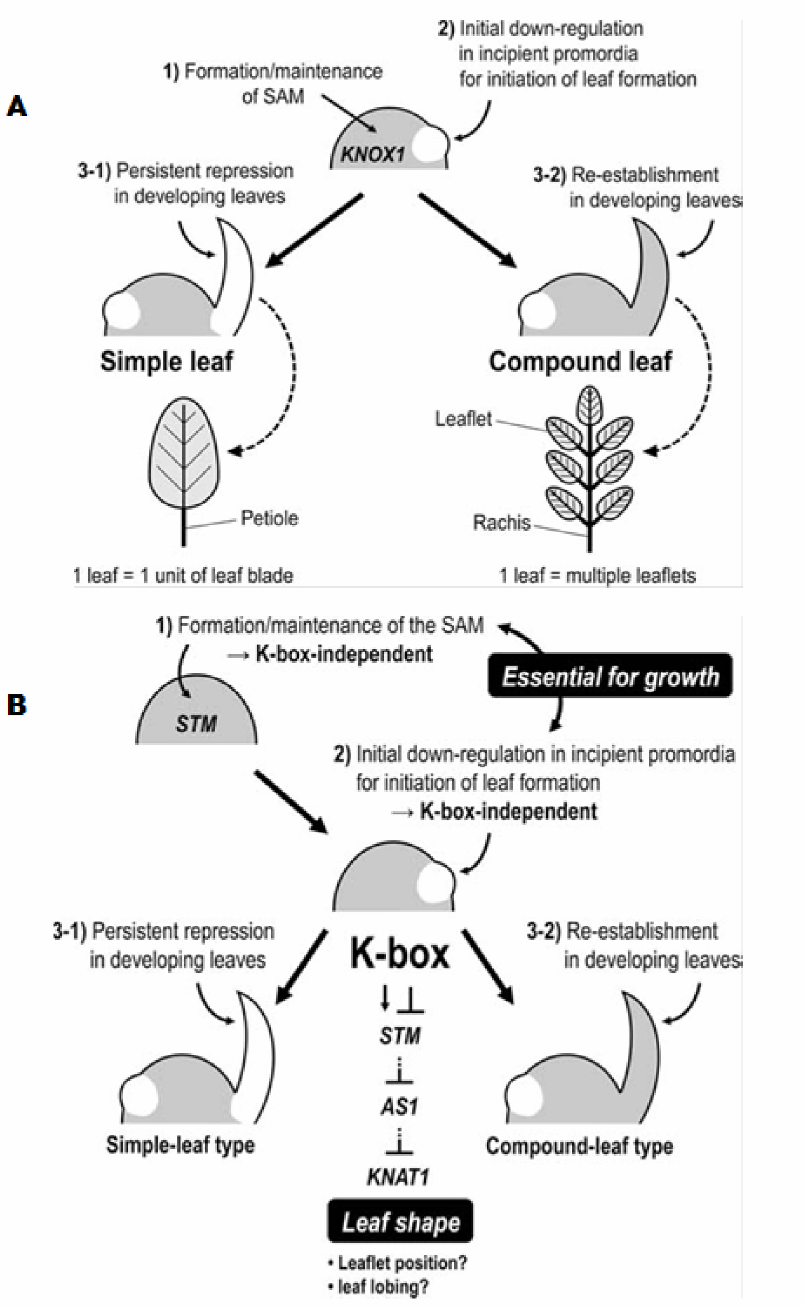


Figure 7 The role of *KNOX* genes in generating variation in leaf form.
(A) Regulation of *KNOTTED1*-LIKE HOMEOBOX (*KNOX1*) genes in leaf development. (B) K-BOX functions in the regulation of *STM* gene in leaf formation (Uchida *et al*., 2009).

1.6.2. *ARP* genes

The transition from indeterminate SAM to determinate leaf primordium requires extensive reprogramming of gene expression. One pathway of reprogramming of gene expression during leaf development involves *ARP* (*ASYMMETRIC LEAVES1* [*AS1*]/ *ROUGH SHEATH2* [*RS2*]/*PHANTISTICA* [*PHAN*]) genes (Byrne *et al*., 2000; Timmermans *et al*., 1999; Waites and Hudson, 1995). *ARP* genes encode MYB transcription factors, a family of transcription factors found throughout eukaryotes but particularly numerous in Angiosperms (Waites *et al*., 1998; Theodoris *et al*., 2003). MYB proteins from animals contain three repeats (R1, R2 and R3) while MYB DNA binding domains from plants generally consist of two imperfect repeats (R2 and R3). Some MYB proteins from fungi also consist of two repeats (Martin and PazAres, 1997).

The genetic structure of *ARP* genes is well conserved amongst Angiosperms and consists of two exons separated by an intron in the 5′ untranslated regions of the gene. *ARP* proteins are distinguished from the rest of this large family by a distinctive C-terminal domain which is required for homodimerization (Theodoris *et al*., 2003). Figure 8 shows the position of different domains in *Begonia heracleifolia BARP1* genes. A major function of *ARP* genes in simple leaved species is repression of *KNOX* activity in the leaves, confining its activity to the meristem (Byrne *et al*., 2000; Schneeberger *et al*., 1998; Tsiantis *et al*., 1999; Ori *et al*., 2000; Timmermans *et al*., 1999). *ARP* genes keep *KNOX* genes in a repressed state by making a complex with *AS2* (*ASYMMETRIC LEAVES2*) which binds directly to the regulatory motifs present at two sites in the promoter of *KNOX* genes. This acts to recruit HIRA (histone regulatory homolog A), a predicted histone chaperone involved in heterochromatic and euchromatic gene silencing, blocking the activity of *KNOX* genes (Phelps-Durr *et al*., 2005; Guo *et al*., 2008).

In the shoot apex of *Arabidopsis*, maize, *Antirrhinum* (Byrne *et al*., 2000; Timmermans *et al*., 1999;Waites and Hudson, 1995) and tobacco (McHale &

Koning, 2004) the balance between indeterminate and determinate growth is regulated by *KNOX-ARP* interactions. In *Arabidopsis* the relationship between *ARP* genes and *KNOX* genes is one of mutual inhibition. *ARP* genes are turned off in the meristem by the *KNOX* gene *STM*. Without *STM* function, *AS1* transcripts are found in the meristem (Byrne *et al*., 2000). Similarly *PSKN1* (*KNOX* genes) and *CRISPA* (*ARP* ortholog in pea) are expressed in complementary domains in *Pisum sativum* (Tattersall *et al*., 2005) and *HIRZINA* and *PHAN* in *Antirrhinum* (Waites *et al*., 1998). However, in a wide range of compound leaved plants *ARP* genes are expressed in the shoot apical meristem and it has been suggested that this is related to the changed interactions between *ARP* and *KNOX* genes in the compound leaf primordium (Kim *et al*., 2003). The relationship between *ARP* in the SAM and compound leaves is not clear as in the compound leaved *Cardamine hirsuta ARP* is not expressed in the SAM (Hay & Tsiantis, 2006).

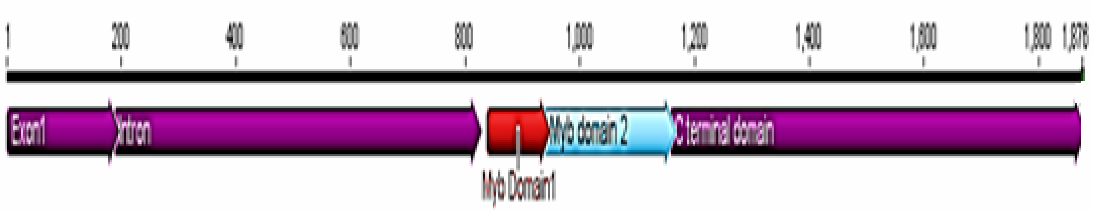


Figure 8 Schematic representation of *BARP1* gene.
Map of *ARP* genes from *B. heracleifolia* (*BARP1 her*) showing the position of different protein domains, intron and exons.

Expression of both KNOX and ARP proteins in the meristem and leaf of the compound leaved fern *Osmunda*, from a lineage which evolved leaves independently to seed plants, suggests independent recruitment of the same mechanisms of leaf development (Harrison *et al*., 2005). *Sellaginella kraussiana* is from a third lineage and its leaves are developmentally very different from the leaves of seed plants. As they are derived from the epidermis of the meristem and have a single vascular trace they are termed microphylls. The expression patterns of *ARP* and *KNOX* orthologs in their meristem and leaf primordia are consistent with a mutually exclusive *KNOX-ARP* interaction in microphylls. The *SkARP1* gene can also repress the *KNOX* genes in *Arabidopsis* illustrating that it is functionally equivalent to *AS1* of *Arabidopsis*

(Harrison *et al*., 2005). This supports a very ancient origin for the *ARP-KNOX* interaction.

Two *PHAN* like genes (*LjPHANa* and *LjPHANb*) have been found in *Lotus japonicus* as a result of gene duplication. Both *LjPHANa* and *LjPHANb* proteins are 76% similar to each other, and 82.0% and 74.2% similar to *PHAN* orthologs in pea while 77.4% and 71.9% similar to *PHAN* homologs in *Medicago trucatula*. Both *Lotus japonicus PHAN* like genes are functional. Over-expression of either *PHAN* gene from *Lotus japonicus* in *Arabidopsis* produces narrow leaves with longer petioles which resembles the phenotypes of 35S *AS1*. The differences in expression patterns suggests that both *LjPHANa* and *LjPHANb* regulate adaxial identity of compound leaf primordia while adaxial identity of leaflet primordia is regulated by *LjPHANb* only (Luo *et al*., 2005).

Ruschioideae is the largest sub family of Aizoaceae which is popular for its remarkable natural diversity in leaf *form*. Ruschioideae species are highly similar at the chloroplast level which led the authors conclude that adaptive radiation in this group has happened approximately 1.5 MYA at a rapid pace (Klak *et al*., 2003). Ihlenfeldt (1994) had hypothesized that this rapid and recent radiation must be due to the diversity in developmental genes. The duplication and subsequent loss of *ARP* genes from some of selected species of Ruschioideae correlates with the diversification of plant forms in the Ruschioideae (Illing *et al*., 2009). As only one *ARP* has been reported to regulate the multiple *KNOX* genes in most plant species (Byrne *et al*., 2002), the duplicated *ARP* genes may have caused the explosive radiation in Ruschioideae while regulating the expression of *KNOX* genes. This led the authors to hypothesize that duplicated *ARP* genes in Ruschioideae may have achieved this while making different complexes and setting up different kinds of interactions than normally seen in different species (Illing *et al*., 2009).

Although the function of *ARP* genes appears to be strongly conserved throughout plants their role in leaf development may vary. The mutant phenotypes of *ARP* genes in different species suggest each plays a slightly different role in the genetic networks regulating dorsoventral polarity, proximal distal polarity and blade outgrowth.

Asymmetric leaves1 (*as1*) mutants in *Arabidopsis* produce leaves which are short and round with occasional lobes and showed no obvious adaxial/abaxial polarity defects (Reidei, 1965; Ori *et al*., 2000; Semiarti *et al*., 2001, Byrne *et al*., 2000). Dorsoventral defects have been seen in combination with other mutations such as *AS2* and *AN3* (*ANGUSTIFOLIA3*) (Horiguchi, 2011 a, b) indicating that in *Arabidopsis* the dorsoventral role of *ARP* genes is largely redundant with other pathways. The *rough sheath2* (*rs2*) mutant in maize has proximodistal defects as proximal features of the sheath, ligule and auricle are displaced distally into the leaf blade (Schneeberger *et al*., 1998). Despite these differences the function of the gene is conserved between the two species as *RS2* rescues *as1* (Theodoris *et al*., 2003). The last of the original trio, *Phantastica* (*Phan*), in *Antirrhinum* has a clear dorsoventral effect as severe phenotypes are fully abaxialised, weaker ones abaxialised only proximally and weakest phenotypes have wider blades with patches of ectopic outgrowth surrounding abaxial patches on the adaxial surface (Waites and Hudson, 1995).

The role of *ARP* genes has also been functionally characterised in tobacco, tomato, pea and *Cardamine hirsuta* (McHale & Koning, 2004; Kim *et al*., 2003; Hay and Tsiantis, 2006; Tattersall *et al*., 2005). *NSPHAN* is required in tobacco for development of the palisade layer, to promote adaxialisation in adult leaves (but not juvenile) and to repress ectopic leaf blade via GA (Gibberellic acid) regulation in juvenile leaves (but not adult) (McHale & Koning, 2004). The interaction of key developmental genes with phase change is also seen for other genes and other species (Wu *et al*., 2008).

Down regulation of *LePHAN* by antisense suppression in tomato resulted in a range of phenotypes. For example needle leaves were observed in plants which lack *LePHAN* expression while cup-shaped or palmate compound leaves have *LePHAN* expression confined to the distal region of leaf primordia only (Kim *et al*., 2003).

CRISPA is the pea *ARP* ortholog. Its expression patterns are similar to simple leaved *Arabidopsis* rather than resembling tomato. Loss of *CRISPA* function affects polarity in the leaflet, but not in the main rachis. *cri* mutants produce two types of ectopic laminae; ectopic laminae on leaflets and stipules and ectopic stipule laminae on the petiole-rachis axis. *CRISPA* down regulates the expression of *KNOX* like genes *PSKN2* in pea to establish a boundary for stipule formation at the base of leaf (Tattersall *et al*., 2005).

In the initial characterisation of *PHAN* it was suggested that lamina outgrowth occurred along the boundary between *PHAN* and no-*PHAN* (Waites and Hudson, 1998). The production of a region of activity and no-activity must be due to interactions with polar genes such as *AS2* (Kidner and Timmermans, 2010). However in other species *ARP* expression patterns do predict where lamina will grow out (Kim *et al*., 2003, Luo *et al*., 2005). Variation in expression patterns of *ARP* genes could be behind some of the variation between peltate and non peltate species and regulate the number and placement of leaflets in compound leaved-species (Kim *et al*., 2003a; Kim *et al*., 2003b). However, evidence for this is currently based on expression patterns and anitsense experiments and lacks the genetic data that supports a role for *KNOX* genes in natural variation in leaf dissection.

1.6.3. *LEAFY* genes

Expression of *KNOX* genes in the leaf primordia is sufficient for generating leaflets in a range of compound leaved plants (Hay and Tsiantis, 2006, Shani *et al*., 2009; Bharathan *et al*.,2002) with the exception of compound leaves of legumes where orthologs of *LEAFY* (*LFY*) and *FLORICAULA* (*FLO*) from *Arabidopsis thaliana* and *Antirrhinum majus* respectively are required for compound leaf formation (Hofer *et al*. 1997; Gourlay *et al*., 2000; Hofer *et al*., 2001; Champagne *et al*., 2007). Mutants of *single leaflets1* (*sg1*) (*LFY*/*FLO* orthologs in *Medicago truncatala*) turned the compound leaves of *Medicago truncatala* into simple leaves owing to the lack of ability to generate lateral leaflet primordia and also caused defects in flower development (Wang *et al*., 2008). *LFY* from *Arabidopsis* was able to rescue the compound leaf and floral defects of *sg1* mutants whereas *SG1* was able to rescue floral defects of mutant *lfy* genes in *Arabidopsis* but it has failed to convert the simple leaf of *Arabidopsis* into compound leaf. This indicates that the ability of *SG1* to control compound leaf formation in *Medicago truncatala* is due to its promoter activity. The difference in expression pattern of *KNOX* genes between pinnately compound leaved *Cardamine hirsuta* and simple leaved *Arabidopsis thaliana* corresponds to the natural variation in the leaf form and is driven by the variation in the promoter region of the *KNOX* genes (Hay and Tsiantis, 2006). It is temptating to hypothesize that *KNOX* and *SG1* promoters may have some common elements for controlling the leaflet formation across different species.

Conversely the function of *LFY/FLO* orthologues in controling compound leaf formation in other legumes is not very strong. For example loss of function mutants of *LFY/FLO* orthologues reduced the number of compound leaves in a *Lotus japonicus* to a varying degree but in *Glycine max* the reduction in leaflet numbers was very low (Dong *et al*., 2005; Champagne *et al*., 2007). However *LEAFY* could be a strong candidate for compound leaf formation in species where *KNOX* genes lack variation to generate differential leaf forms.

1.6.4. *CUC* genes

The boundaries between plant organs are controlled by genes of *CUC/NAM* family {*CUP-SHAPED COTYLEDON1*(*CUC1*), *CUC2*, *CUC3* and the petunia *NO APICAL MERISTEM* (*NAM*)} which all encode transcription factors of the NAC-domain class (Takada *et al*., 2001; Aida *et al*., 1997; Souer *et al*., 1996; Weir *et al*., 2004; Hibara *et al*., 2006 a, b; Blein *et al*., 2008; Adam *et al*., 2011). *CUC* genes activity reduces the growth between organs, forming boundaries and it is correlated with low auxin levels (Aida *et al*. 2002; Furutani *et al*. 2004). Mutations in these genes generate a plant with fused organs due to defects in the establishment of several boundaries. *NAM/CUC3* genes are expressed at the distal boundaries of leaf rachis where leaflets are formed in *Solanam lycopersicon*, *Cardamine hirsuta* and *Pisum sativum* and also regulate the expression of *KNOX/UFO* like genes through a feed forward regulatory loop during compound leaf formation in these species (Blein *et al*., 2008).

1.6.5. *YABBY* genes

YABBY genes are important for setting up polarity as well as establishing an adaxial/abaxial juxtaposition for the proper formation of leaf lamina (Bowman 2000; Golz *et al*., 2004; Eshed *et al*., 2004). In *Arabidopsis* leaves *YABBY* genes are expressed on the abaxial side of the leaves whereas in maize leaves they are expressed adaxially (Juarez *et al*., 2004 a, b; Golz *et al*., 2004). This suggests that the function of *YABBY* genes for controlling leaf polarity may has diverged between *Arabidopsis* and maize. However the change of expression domain maintains the distinction between *YABBY* and no-*YABBY* at the point of lamina outgrowth. Besides their role in polarity and laminae formation they also have been associated with the development of peltate leaf form in *Tropaeolum majus* (Gleissberg *et al*., 2005). However *Tropaeolum* can not be used to test the hypothesis that variation in the expression of *YABBY* genes is responsible for the evolution of peltate leaf form as all the species of *Tropaeolum* are peltate. However *Begonia* can be a good system to test this hypothesis as this genus has a large number of sister species with peltate and non peltate leaf form in section Gireoudia.

1.7. Candidate genes or QTL approaches to identifying the genes underlying species-level variation

In the candidate gene approach, a gene is selected whose role for controlling the desired trait has been established on the basis of mutational analysis, biochemical interactions studies and expression pattern analysis in model plants. This approach has both strengths and limitations.

A candidate gene approach is hypothesis driven, cost effective and phenotype specific. However cloning of candidate genes from non-model organisms is a tedious job and candidate gene selection can miss some other important loci controlling the trait of interest. A candidate gene approach can be used for simple traits or well characterized pathways like leaf development or flower development which have been studied widely through natural or induced mutants and genetic analysis (Saïdou *et al*., 2009; Kelly *et al*., 2003; González-Martínez *et al*., 2007; Hay and Tsiantis, 2006; Shani *et al*., 2009; Kim *et al*., 2003). Conversely for complex traits the whole genome is potentially a candidate. Furthermore choosing a candidate gene for studies of the genetic basis of natural variation is tricky as most of the knowledge of candidate gene functions is based on the comparison of severe mutants with wild type plants.

In QTL studies a large number of markers are analysed for association with the phenotype (Thornsberry *et al*., 2001; Olsen *et al*., 2004; Zhao *et al*., 2007; Wei *et al*., 2006; Gonzalez-Martinez *et al*., 2006, 2007; Casa *et al*., 2008; Kraakman *et al*., 2006). For performing QTL analysis previous knowledge of candidate genes is not necessary. However it is a costly technique as it requires high density genotyping with molecular markers to efficiently identify the number of loci controlling traits of interest.

Over all a candidate gene approach is better to choose in a scenario where the trait of interest has already been studied in different species and a genetic link has already been established through functional and expression data. Further candidate gene studies in non-model organisms can generate information about nucleotide diversity and species relatedness and this knowledge can be built through genome sequencing and high density genotyping.

Plant leaf form has been a focus of plant developmental biology and Class1 *KNOX* and *ARP* genes have been shown to control the simple, compound and peltate leaf form across different species and are the strongest candidates for natural variation in leaf form.

1.8. *Begonia* as a model to study the genetics of leaf form

To establish whether key developmental genes are involved in species-level variation in leaf form the ideal study system would allow genetic association studies, expression analysis and transgenic work. A combination of these features with leaf forms that range from compound to entire and both peltate and basifixed was not available so we have chosen a group which offers good genetics, a wide range of phenotypes and the possibility of expression analysis and transgenics.

The genus *Begonia* belongs to the angiosperm family Begoniaceae in order Cucurbitales. It is one of the ten largest Angiosperm genera with over 1,500 species found throughout the tropics. Phylogenetic studies indicate that the most basal *Begonia* species are African, from which both Asian and American *Begonia*s are derived (Forrest *et al*., 2003). Parallel radiations have occurred in the South East Asia and in the New World. Although the floral form is similar throughout the genus the vegetative form varies widely.

The typical *Begonia* is a low growing herb with heart shaped asymmetric leaves scrabbling over rocks or riverbanks. However, some *Begonia* grow up to 4m tall with huge compound leaves over 50 cm across (*B. luxurians*), others have grass-like tufts of short, linear leaves (*B. bogneri*). Peltate species have evolved at least eight times and compound leaves at least four times. Chromosome number varies widely within the genus, constraining the possibilities for genetic work (Forrest, 2000). We have picked section Gireoudia as a group with invariant chromosome number (2n=28) and a long history of horticultural hybridisation supporting our plans for genetic analysis. This section includes five peltate leaved species and two compound leaved species. Phylogenies of this group are difficult as its recent radiation means there are few phylogeneticaly informative sites, but support for independent evolution of peltate and compound leaves in each case is strong (Nicola Harrison, unpublished data).

Using this group of species I aim to answer the following questions

1- Is leaf form a simple genetic trait?
2- Is sequence variation in coding sequence and promoters of *ARP* and *KNOX* associated with variation in leaf form?
3- Is *ARP* from *Begonia* functionally equivalent to *ARP* in *Arabidopsis*?
4- Is expression variation in *ARP* and *KNOX* associated with variation in leaf form?

5- Are *ARP* and *KNOX* genes genetically associated with leaf form variation between species?

To do this I used F1 (First generation seeds/plants that come from a cross of distinctly different parental types) and, where possible F2 (Seeds/plants resulting from a self or cross pollination of F1) or F1BC1 (Seeds/Plants resulting from a cross of F1 with either of the parents) populations from crosses between compound, simple, peltate or basifixed species to determine dominance, reveal whether any are simple Mendelian traits and to provide mapping populations (Chapter 3). *ARP* and *KNOX* genes were cloned using degenerate PCR and TAIL PCR and the sequences examined for possible functional variations and signs of selection (Chapter 4). Establishing the function of *ARP* genes in *Begonia* requires the development of a method of gene silencing or transformation. Attempts at this are described in Chapter 5. To determine the function of *Begonia ARP* genes expression vectors were constructed and transformed into *as1* mutants in *Arabidopsis* (Chapter 5). Mapping populations were genotyped to test linkage between leaf form and these genes (Chapter 7). Chapter 8 summarises my finding and discusses how far we have come in being able to answer these questions.

Chapter 2

Material and Methods

2.1. Plant materials used and growth conditions

All plant material for this study was provided by the Royal Botanic Garden of Edinburgh *Begonia* collections where they grew in the glass houses in tropical climatic conditions with approximate day temperature of 28˚C, night temperature of 20˚C and a relative humidity of 70 %.

The *Begonia* seeds were sown on a mix composed of 50% grow bark + 50% John Innes compost which had been sterilized with boiling water and allowed to cool in sealed plastic bags. The seedlings of approximately eight weeks old were transferred to a sieved tropical mix (70ltrs Melcourt potting bark + 70ltrs Melcourt propagation bark + 70ltrs Melcourt grow bark + 45ltrs John Innes no.1 + 30ltrs perlite and 500g osmocote). Seedlings were left on a warm spray bench in the open air for four to six weeks to acclimatize before they were transferred to an unsieved tropical mix. *Begonias* were also propagated from leaf cuttings and were grown on a mix of 50% perlite + 50% propagation bark on the spray bench. A list of species used, their accession numbers and leaf form is given in Table 3.

Table 3 *Begonia* species, their accession numbers and leaf forms.

Species	**Accession number**	**Leaf form**
Begonia nelumbiifolia	19791880	Peltate
Begonia conchifolia	20042082	Peltate
Begonia peltata	2004078	Peltate
Begonia kellemanii	20030642	Peltate
Begonia heracleifolia	20042080	Non peltate
Begonia plebeja	20051406	Non peltate
Begonia sericoneura	20051412	Non peltate
Begonia lindlyana	20051412	Non peltate
Begonia thiemei	20042079	Compound
Begonia carolineifolia	20042077	Compound
Begonia multinervia	20051411	Non peltate
Begonia pruniata	20070430	Peltate
Begonia mazae	20070792	Non peltate

2.2. Production of hybrids and backcross populations

Pollen ripeness is checked by confirming that the stamens shed pollen when brushed. The stamens are then picked and brushed several times against a fresh looking stigma with clear papillae. The crosses were labeled with plastic tags tied to the pedicle with thread and the details noted. The crosses were checked after a couple of days. Successful crosses will have dropped their tepals and unsuccessful crosses may have fallen off entirely. After approximately two to three weeks the carpel dried out and were removed and stored in glassine bags in desiccators at 4 ^{o}C. List of hybrids and back crosses is given in table 4.

Table 4 *Begonia* section Gireoudia hybrids and back crosses

Cross	**Cross number**	**No. of plants**
B. nelumbiifolia x *B. heracleifolia*	CKB44	4
B. conchifolia x *B. plebeja*	CKB137	8
B. carolinifolia x *B. heracleifolia*	CKB66	9
B. conchifolia x *B. sericoneura*	CKB197	3
CKB44 x *B. nelumbiifolia*	CKB250	8
CKB137.6 x *B. conchifolia*	ARB312	129
CKB137.1 x *B. conchifolia* AFLP genotyping showed this population to be a likly mixed sib cross from CKB137	ARB302	158
CKB137.1 x *B. plebeja*	BOB360	276
CKB137.1 x *B. plebeja*	ARB314	120
CKB197.2 x *B. conchifolia*	ARB520	87
CKB197.1 x *B. sericoneura*	ARB210	83
CKB66.7 x *B. heracleifolia*	ARB444	40

2.3. Molecular techniques

2.3.1. DNA extraction

Total genomic DNA was extracted from fresh vegetative buds (approximately 100mg) or silica dried vegetative/floral buds (approximately 20mg) using a QIAGEN DNAeasy extraction protocol. Fresh vegetative buds were ground in liquid nitrogen with a micropestle in a micro centrifuge tube and silica dried vegetative/floral buds were disrupted with Tissuelyser system. 400 µl of Buffer AP1 and 4 µl RNase stock solution was added to the ground material and vortexed vigorously. The mixture was incubated for 30-50 min at 65 ^{o}C. 130 µl of buffer AP2 was added to the lysate, mixed well and incubated for 5 min on ice. The samples were than centrifuged at 14,000 rpm for 5 minutes. A QIAshredder Mini spin column (lilac) was placed in a 2 ml collection tube and lysate was added into it and centrifuged at 14000 rpm for 2 min. The flow through was transferred into a new tube without disturbing cell pellet. To this 1.5X volume of Buffer AP3 was added, mixed and 650 µl of the mixture was pipetted into DNeasy Mini spin column placed in 2 ml collection tube and centrifuged for 1 min at 8000 rpm. This step was repeated with the remaining flow through from previous step. Flow through and collection tube was discarded. The sample was washed twice with 500 µl buffer AW by centrifuging for 1 minute at 8000 rpm, than centrifuged for 2 min at 14000 rpm to dry the membrane. The DNeasy Mini spin column was transferred to 2 ml micro centrifuge tube and 100 µl buffer AE was added directly onto the DNeasy membrane incubated for 5 minutes at room temperature and than centrifuged for 1 min at 8000 rpm to elute the DNA. The last step was repeated once. DNA was stored at -20°C.

2.3.2. PCR amplification for *Begonia ARP/KNOX* genes

2.3.2.1. Primer designing

Primers were designed using Primer 3 run from within the Geneious program. The primers are listed in appendix A. These primers were used in various combinations to amplify *BARP* genes, *KNBs* and *CUCs*.

2.3.2.2. Gradient / General Polymerase Chain Reaction (PCR)

Total genomic DNA was isolated from 13 Gireoudia species, their hybrids and back crosses given in table 3 and 4. The DNA obtained from these species was used as a template in PCR amplification using primers given in appendix A. The reaction was set up in a final volume of 20/50 µl. Each individual reaction mixture consists of 200 µM of each dNTP, 1 µM of each primer, 10-50 ng of template DNA, 1 x Standard Taq Buffer, 2.5 mM $MgCl_2$ and 1 U of Taq DNA polymerase. Each PCR reaction was performed in a single PCR tube. Reactions were carried out using the Thermo Hybaid gradient PCR with an annealing temperature gradient ranging from 50-60°C and the rest of the conditions were the same as in Table 5.

Table 5 Reaction conditions for Standard Taq Polymerase

PCR step	Temperature and time	No. of cycles
Initial Denaturation	94°C for 2 min	1x
Denaturation Annealing Extension	94°C for 30 sec Tm °C for 1:30 sec 72°C for 2 min	39x
Final Extension Cooling	72°C for 10 min 4°C Forever	1x

(Tm °C depends on the annealing temperature of the primer pair). Optimum annealing temperature of the primer pair was determined by performing a Gradient PCR.

2.3.2.3. TAIL PCR

Thermal asymmetric interlaced PCR is used for isolation of target DNA segments adjacent to known sequences as described by (Liu *et al*., 1995) with some modifications. This PCR used a set of nested sequence specific primers together with a short arbitrary primer of low melting temperatures. For efficient function of specific primers the annealing steps were carried out at different temperatures.

Targeted products were amplified preferentially over non-specific sequences by interspersing high and reduced stringency cycles. The primers used in TAIL PCRs are listed in appendix A. The reactions were set up in a final volume of 20 µl for TAIL1 and TAIL2 PCR; and 50 µl for TAIL3 PCR. DNA templates used in TAIL1 PCRs were 10-50ng, a 1/50 dilution of TAIL1 PCR products were made and 1µl of each diluted sample were used in TAIL2 PCR reactions as templates and a 1/10 dilution of TAIL2 PCR products were made and 2.5 µl of each diluted sample were used in TAIL3 PCR reaction as template while the remaining reaction ingredients were the same for all three stages of TAIL PCRs. Each individual reaction mixture consists of 200 µM of each dNTP, 0.5 µM of each primer, 1 x Standard Taq Buffer, 2.5 mM MgCl2 and 1 U of Taq DNA polymerase. Reactions were carried out using the Thermo Hybaid gradient PCR and reaction conditions for TAIL1, TAIL2 and TAIL3 are described in table 6, 7 and 8.

2.3.2.4. Infusion PCR

Infusion PCR (Clontech) was carried out to fuse the ends of the PCR fragment to the homologous ends of a linearized vector. During primer design 15 bp extensions to both forward and reverse PCR primers were added that precisely match the ends of the linearized vector. When the vector is combined with the PCR fragment, the In-Fusion® enzyme (Clontech) converts the double-stranded extensions into single-stranded DNA and fuses these regions to the corresponding ends of the linearized vector. I carried out the infusion reactions in 2:1 ratio for insert (infusion PCR amplified fragment) to vector (double digested vector of interest). The final volumes used for this reaction were calculated using online tool “Infusion molar ratio calculator”. Total 10 µl of insert and vector were put into infusion dry down pellet, mixed well and incubated fifteen minutes at 37 ^{o}C and fifteen minutes at 54 ^{o}C. 1/40 volume of infusion reaction was transformed into chemically competent TOP10 cells. Colony PCR followed by sequencing was used to select positive clones.

Table 6 Reaction conditions for TAIL PCR 1

Temperature	Time	No. of cycles
94˚C	1 min	1x
94˚C	10 sec	4x
62˚C	1 min	
72˚C	2.30 min	
94˚C	10 sec	14x
25˚C	3 min	
0.2˚C /sec to 72		
72˚C	2.30 min	
94˚C	10 sec	
69˚C	1 min	
72˚C	2.30 min	
94˚C	10 sec	
68˚C	1 min	
72˚C	2.30 min	
94˚C	10 sec	
44˚C	1 min	
72˚C	2.30 min	
72˚C	5 min	
4˚C	for ever	

Table 7 Reaction conditions for TAIL PCR 2

Temperature	**Time**	**No. of cycles**
94˚C	10 sec	11x
64˚C	1 min	
72˚C	2.30 min	
94˚C	10 sec	
64˚C	1 min	
72˚C	2.30 min	
94˚C	10 sec	
44˚C	1 min	
2˚C	2.30 min	
72˚C	5 min	
4˚C	for ever	

Table 8 Reaction conditions for TAIL PCR 3

Temperature	**Time**	**No. of cycles**
94˚C	15 sec	19x
44˚C	1 min	
72˚C	2.30 min	
72˚C	5 minutes	
4˚C	for ever	

Size standards used were

- 1Kb ladder NEB (New England Biolabs) (10Kb-500bp, ~1kb intervals between bands)
- 100bp ladder NEB (1.5Kb – 100bp, 100bp intervals)
- HyperLadder™ I (10Kb– 100bp, 200/500bp intervals)

Some of the PCR amplified products were used for cloning into plasmid vector (pGEM@T vector) while some were used directly for sequencing.

2.3.4. PCR clean up

The PCR products were cleaned up using a GFX^{TM} PCR Purification kit (GE Healthcare). 100 µl of capture buffer was added to the GFX column placed in collection tube for approximately 20 µl of product to be purified. DNA solution was added to the GFX column and mixed thoroughly by pipetting 4-6 times following centrifugation for 30 s at full speed on a bench centrifuge. Flow through was discarded and columns were washed with 500 µl wash buffer than centrifuging for 30 s at full speed. The collection tube was discarded and DNA was eluted using 10 µl elution buffer with one minute incubation on ice follwed by centrifuging at full speed for 1 min. The purified DNA fragments were either cloned or directly sequenced.

2.3.5. Gel Extraction of PCR Products

The desired bands were excised with a sharp, clean scalpel whilst viewing the gel on a long wave UV transilluminator (340nm). The excess gel was trimmed away to minimize agarose. The gel slice was weighed in a 2 ml Eppendorf tube and 3 gel volumes of capture buffer were added. The gel mixture was incubated at 65°C for 10 -30 minutes, vortexed for every 2-3 minutes during incubation to dissolve the gel. Dissolved gel mixture was added directly onto a GFX Column followed by an incubation of 1 minute at room temperature. The tube was centrifuged at 13,000 rpm for 1 minute. The flow-through liquid was discarded after centrifugation and columns were washed with 500 µl wash buffer while centrifuging for 30 s at 13000 rpm. The collection tube was discarded and DNA was eluted in 2 ml eppendorfs with 10-20 µl elution buffer following a 1 min incubation on ice and centrifuging at full speed for 1 min. The purified DNA fragments were either cloned or directly sequenced.

2.3.6. Cloning of PCR products

2.3.6.1. Ligation Reaction

Purified PCR products were ligated into pGEM-T easy vector (Promega) (Figure 9). The ligation reactions were set up in a total volume of 5 µl containing a molar ratio of ~31 insert to vector DNA, 2.5 µl of 2X Rapid ligation buffer, 0.5 µl of pGEM@T easy vector and 0.5 µl of T4DNA Ligase. The ligation reaction was incubated at 4˚C over night after which the ligation mixture was directly used for transformation.

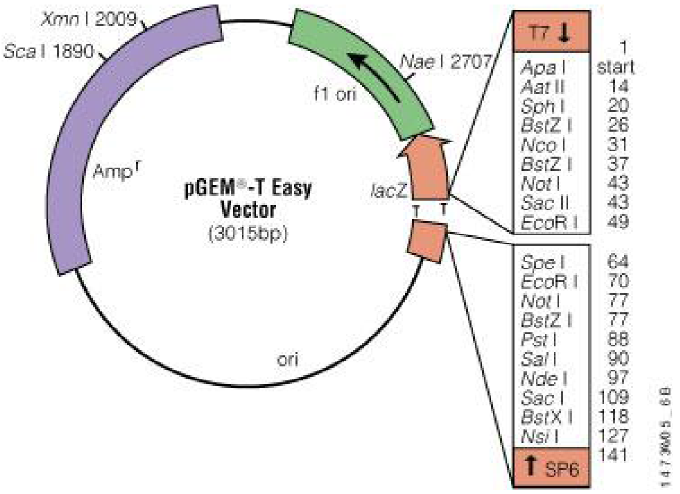


Figure 9 pGEM T-Easy vector restriction map and sequence reference points.

(www.promega.com).

2.3.6.2. Transformation

Transformation of ligation mixture into competent E.coli Top10/JM109 cells was performed by the heat shock method. An aliquot of frozen *E. coli* chemically competent cells was thawed on ice at least 5 minutes before transformation. 25 µl of competent cells were used for a single transformation reaction. 2.5 µl of ligation mixture was added to the cells directly and incubated on ice for 20-30 minutes. Following incubation the cells were than heat shocked at 42°C for 30 sec and immediately placed on ice for 2 minutes. 500 µl of room temperature SOC (Super optimal broth) media was added to the cell suspension followed by incubation at 37°C in an orbital shaker (150 rpm) for 1 hour. 100 ml and 400 ml of each transformed vial were spread onto duplicate LB agar plates containing Ampicillin (100ug/µl) and Bluetech as selection markers. The plates were kept at 37^oC for about 12-18 hours. White colonies represent recombinant colonies while blue colonies are self-ligated. The size of insert was checked by colony PCR and the bacterial strain was maintained as streaks on LB Amp (100µg/µl) plates.

2.3.7. Colony PCR

Colony PCR was carried out using universal M13 forward and M13 reverse primers. PCR reaction was set up in a final volume of 20 µl consisting of 200 µM of each dNTP, 0.5 µM of each primer, 1 x Standard Taq Buffer, 2.5 mM MgCl2 and 1 U of Taq DNA polymerase. A white tip was touched to the colony to pick it up, dipped in the PCR reaction and tapped in order to release some cells into PCR reaction. The tip was stored in 100 µl Amp (100µg/µl) in closed eppendorf at 4 ^oC. PCR conditions for colony PCR are described in table 9. The PCR product was purified and sequenced. Also plasmids of positive clones were isolated.

Table 9 Reaction conditions for Colony PCR

PCR step	**Temperature and time**	**No. of cycles**
Initial Denaturation	94°C for 5 min	1x
Denaturation Annealing Extension	94°C for 30 sec 54°C for 130 sec 72°C for 2 min	35x
Final Extension Cooling	72°C for 10 min 4°C Forever	1x

2.3.8. Plasmid isolation

Plasmid DNA from transformed colonies was extracted using the QIAprep Spin Miniprep Kit according to manufacturer instructions. List of plasmids isolated during my project are listed in appendix A.

2.3.9. Long term storage of Plasmid DNA

Once the correct clones were identified, glycerol stocks were made by purifying the colony. The original colony was streaked out for a single colony on LB plates containing 50 mg/ml kanamycin. A single colony was isolated and inoculated into 3 ml of LB containing 50 mg/ml kanamycin and grown until the culture reached stationary phase. 850 µl of culture was mixed with 150 µl of sterile glycerol in a cryovial, followed by dipping in liquid nitrogen for 2 minutes and finally stored at -80°C.

2.3.3. Agarose Gel Electrophoresis

Agarose gels of appropriate concentration (generally 1% or 2%) were cast using agarose melted in 0.5x Tris-borate EDTA (TBE) buffer in a microwave oven. The molten gel was allowed to cool to 55-60°C and gels were supplemented with ethidium bromide at a final concentration of 0.5µg/ml. The molten agarose was than poured into a gel tray with its edges taped and an appropriate comb or combs inserted. After the bubbles were removed the gel tray containing the agarose was than left at room temperature until it had properly set. The gel tray was than submerged in a gel tank containing 0.5X TBE buffer. The wells were loaded with DNA samples after mixing them with an appropriate volume of 6x gel loading dye. After loading appropriate size markers and samples, gels were run at 90-120 V for 40 min to 1 hour. The gels were than visualized through the U.V. transilluminator and photographed using a video copy processor.

2.3.10. Sequencing

Sequencing reactions were set up with purified PCR product or with plasmid. For plasmids a heat treatment for 1 minute prior to making the reaction mixture was used as recommended in the PCR profile (Alex & Hollingsworth, 2006). The sequencing reaction was set up in a final volume of 10µl containing 4 µl of dH2O, 4 µl of DTCS Quickstar mix, 1 µl of 10 µM Primer and 1 µl of Template DNA (10ng). Sequencing reactions were run in the PCR thermocycler with the conditions described in table 10.

Table 10 Sequencing PCR Profile

Temperature	Time	No. of cycles
96˚C 50˚C 60˚C	20 sec 20 sec 4 min	35
4˚C	for ever	

2.3.11. Sequencing PCR Clean Up

The stop solution was made using 0.4 µl of 0.5 M EDTA, 1.6 µl of Sigma water, 2 µl of 3M NaOAc pH 5.2 and 1 µl of Glycogen. The volume of sequencing PCR reaction was raised to 20 µl with dH2O and than added into 0.5ml tubes containing stop solution. 60 µl of 95-100% ethanol was added in the same tube and mixed thoroughly following 15 minutes centrifugation at 13000 rpm. The pellet was washed with 70% ethanol twice at 13000rpm for 5minutes. Than the pellet was dried using vacuum dryer for 2-5minutes until no traces of alcohol were left. Each pellet was resuspended in 35 µl of Sample Loading Solution (SLS) and mixed well. Finally sequencing reactions were run and analysed on CEQ 8000 genetic analysis systems by gene pool at the University of Edinburgh or by Royal Botanic Gardens Edinburgh.

2.3.12. Sequencing analysis and alignment

Sequences were imported into Geneious software for editing and assembly of consensus sequences. They were than aligned into a sequence matrix and were manually adjusted before analysis.

2.3.13. Gateway Cloning

2.3.13.1. Primer designing for directional TOPO cloning

To enable directional cloning CACC sequence was inserted at the 5′ end of the forward PCR primer as CACC base pair with the overhang sequence GTGG in pENTR™ / D TOPO® vector (Figure 10) (Xu and Lee *et al.*, 2008). To ensure that PCR product clones directionally with high efficiency, the reverse PCR primer must not be complementary to the overhang sequence GTGG at the 5′ end. There was no stop codon in the gene of interest. The stop codon at the end of C-terminal of gene was removed with a codon for an innocuous amino acid i.e. lysine. The primers used in pENTR™ / D TOPO® cloning are described in table 11.

Table 11 Primers used for TOPO cloning

Gene	F Primer	R Primer	Tm	Extension Time
AS1/*AS1* modified	CACCATGAAAGAGAGACAACG	CTTGGGGCGGTCTAATCTGC	60˚C	2 min
BARP1	CACCATGAAAGATAGGCAGCG	CGGCCATTAGAATCAACGACTC	61˚C	2 min

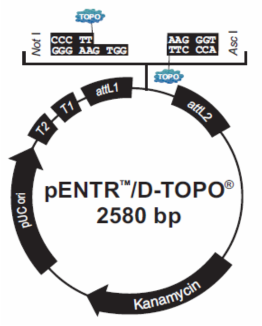


Figure 10 pENTR™ / D TOPO® vector map (http//tools.invitrogen.com)

2.3.13.2. Expand High Fidelity Polymerase chain reaction (PCR)

Expand High Fidelity PCR was performed with Expand High Fidelity PLUS PCR System (Roche). The reaction was set up in a final volume of 50 µl containing 200 µM of each dNTP, 0.4 µM of each primer, 2 µg of plasmid, 1 x Expand HiFiPLUS reaction buffer with MgCl2 and 2.5 U of Expand HiFiPLUS enzyme. Reactions were carried out using the Thermo Hybaid gradient PCR. Expand High Fidelity PCR profile is described in table 12.

The PCR products were checked on 2% agarose gel as described in 2.3.3. Desired PCR products were either purified or gel extracted according to the protocols described in 2.3.4 and 2.3.5 respectively.

Table 12 Reaction conditions for Expand HiFiPLUS Enzyme

PCR step	Temperature and time	No. of cycles
Initial Denaturation	94°C for 2 min	1x
Denaturation Annealing Extension	94°C for 30 sec 60°C/61°C for 130 sec 72°C for 2 min	10x

Denaturation Annealing Extension	94°C for 30 sec 60°C/61°C for 130 sec 72°C for 2 min + 5sec	20x
Final Extension Cooling	72°C for 10 min 4°C Forever	1x

2.3.13.3. TOPO® Cloning Reaction

The molar ratio of PCR product TOPO® vector used in the TOPO® reaction was 2:1. The reaction was set up using 1 µl of salt solution, 1 µl of TOPO® vector, 2-3 µl of purified PCR product and sterile water was added to a final volume of 6 µl. The reaction was gently mixed and incubated for 5 minutes at room temperature. 2 µl of TOPO® reaction was transformed into chemically competent TOP10 cells according to the procedure described in 2.3.6.2 except LB agar plates had 50mg/ml kanamycin.

Positive clones were confirmed by colony PCR (description in 2.3.7) using gene specific primers followed by sequencing. Restriction analysis was performed in parallel. After confirming clones by sequencing, 5 colonies were picked and cultured overnight in 300ml LB media containing 50mg/ml kanamycine. The plasmids were isolated and digested.

2.3.13.4. Double Digestion of Plasmid DNA

The restriction enzymes used were all from NEB UK. The double digestion was planned using NEB double digest finder tool (http//www.neb.com). The restriction analysis was performed in a total volume of 20 µl consisting of 7 µl of plasmid DNA, 1 µl of each restriction enzyme, 2 µl of enzymes specific buffer, 2 µl of 1X BSA and 7 µl of sterile water. The reactions were incubated at 37°C for 3-4 hours following analysis by gel electrophoreses. Glycerol stocks of positive clones were made as described in 2.3.9 and stored at -80°C.

2.3.13.5. Construction of modified *AS1*, *BARP2* like vectors

Modified *AS1* with an 18bp deletion in Myb domain 2 like *BARP2* was synthesized by Synthesizer Company DNA 2.0. The plasmids were rescued from filter paper according to manufacturer's instructions. The yield was too low to use in subsequent experiments. The plasmid was retransformed into TOP10 cells and used for infusion PCR.

2.3.13.6. LR Recombination Reaction

LR Recombination Reaction was carried out using Gateway® LR Clonase ™ II Enzyme mix for constructing vectors for transgenic work described in chapter 5 (Invitrogen). LR Clonase™ II enzyme mix was thawed for about 2 minutes on ice and vortexed for 2 seconds twice. 150ng of entry clone and 150 ng of destination vector (Figure 11) were mixed with TE buffer, pH 8.0 to final volume of 8 µl. To this 2 µl of LR Clonase™ II enzyme mix was added following a brief vortex twice and micro centrifuged briefly. The reaction was incubated at 25°C for 1 hour and terminated by adding 1 µl of the Proteinase K solution to each sample while incubating at 37°C for 10 minutes. 1 µl of each LR clonase reaction was transformed into chemically competent cells TOP10 according to the procedure described in 2.3.6.2 except LB agar plates had 50µg/ML Kanamycin and 50µg/ML Hygromycin. Transformants were screened for positive clones with double digestion followed by sequencing.

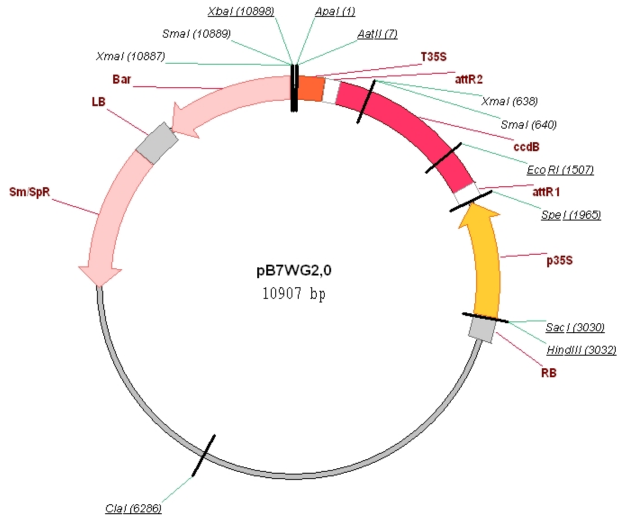


Figure 11 Destination vector (Gateway vector with 35S Promoter) (Karimi *et al*., 2002)

2.3.15. Transformation of *Arabidopsis* by floral dipping

The transformation of *Arabidopsis* was done according to the floral dipping method of Clough and Bent (1998) using *Agrobacterium* strain 3101. The plants were grown for 6 weeks under the long day conditions (22°C, 70% humidity and 18 hours of light conditions) before actual transformation. After 6 weeks the emerging bolts were clipped to induce the growth of secondary bolts. Two days prior to plant transformation the *Agrobacterium* strain carrying the appropriate binary vector was grown in 5 ml of LB with appropriate antibiotics for 16-18 hours at 28°C. 2-3 ml of the overnight grown culture was than inoculated into 400 ml of LB with appropriate antibiotics and grown until OD600=0.8 or higher. The bacterial suspension was than spun down and resuspended in 5% sucrose (freshly prepared). Silwett L-77 was than added to the resuspended culture at a final concentration of 0.05% and the *Agrobacterium* suspension was transferred to a 1000 ml plastic beaker for dipping plants. The aerial parts of flowering *Arabidopsis* were than dipped into *Agrobacterium* solution for 20 sec with gentle agitation. The same suspension was used to dip 10-12 plants. After dipping, the plants were than covered with plastic sleeves for 24 hours to maintain high humidity. After 24 hours the plastic sleeves were opened and the plants were watered and grown normally. The dried siliques from the transformed plants were than harvested after 2-3 weeks.

2.3.16. Screening of *Arabidopsis* transformants

The dried siliques from the transformed plants were harvested and plated on ½ MS media containing 50 µg/ml of Kanamycin. The Petri dishes were than cold treated for 4 days in the dark to break dormancy and than transferred to the growth room with 16/8hr photoperiod at 23°C. After 2 weeks, Kanamycin resistant plants with true leaves and extended root system were than transferred into 9 cm pots.

2.3.17. Southern hybridization

2.3.17.1. DNA Probes for Hybridization

Probe templates were amplified from plasmid (pSU19.2) using Taq Polymerase (Bioline), Universal M13 F and M13 R primers under the conditions described in 2.3.2.2. The PCR products were digested with *Dra*1 to isolate the C-terminal domain of the *BARP* gene. The desired fragment was gel purified according to the protocol described in 2.3.3 and cloned again. The labeled probes were amplified using these templates and the same PCR conditions, but replacing unlabelled dATP with 32P labeled ATP.

2.3.17.2. Digestion of DNA

10 µg DNA was extracted from *B. nelumbiifolia*, *B theimei* and *B. sericoneura*, silica dried floral buds (approximately 20mg) using Quigen Kit according to the protocol described in 2.3.1. Each species DNA was digested either with an enzyme that does cut *BARP* gene or with an enzyme that does not cut the *BARP* gene. *B. nelumbiifolia* and *B. sericoneura* DNA were digested with *Dra*1 (cuts *BARP* gene) and *Xba*1 (does not cut *BARP* gene), and *B theimei* DNA was digested with *Pst*1 (cuts *BARP* gene) and *Xba*1 (does not cut *BARP* gene). The digestions

were carried out in a total volume of 100 µl using 2 µl of enzyme, 5 µl of enzyme specific buffer, 2 µl of 0.1 M Spermidine and 91 µl of DNA. The digestion reactions were incubated at 37°C overnight.

Note 1µl of enzyme was added at the start of digestion and another 1 µl was added after 1hour of incubation at 37°C.

2.3.17.3. Ethanol precipitation

At the conclusion of digestion, DNA was Ethanol precipitated. 0.1 Volume of 3M NaOAc and 2 volumes of 100% EtOH was added to the DNA and incubated at -20 °C overnight. The precipitated DNA was spun at 13000 rpm for 5 minutes. The pellet was rinsed with 1ml of 70% EtOH while centrifuging at high speed for 5 minutes. The DNA pellet was dried in a speed vacuum for 4 minutes. The DNA was resuspended in 15 µl distilled water.

2.3.17.4. Electrophoresis and Blotting

The digested fragments were separated by overnight electrophoresis on 0.8% agarose-TAE at 20V. The gel was photographed including a florescence ruler. The gel was placed in a container slightly larger than gel itself and 250 ml of Denaturation solution (see appendix C.1) was added and incubated for 15 minutes on a shaker twice. The Denaturation solution was decanted off and gel was washed with tape water followed by an incubation of 30 minutes on a shaker with 250 ml of neutralization solution (see appendix C.1) twice. The gel was transferred to the Hybond N transfer membrane (Amersham) by downward blotting using 10X SSC (Saline sodium citrate buffer) (see appendix C.1) as a transfer buffer. The gel was blotted for 24 hours after which the membrane was washed in 2X SSC and left to air dry. The membrane was laid with a DNA side up on saran wrap and DNA was fixed on the membrane by crosslinking it in the Stratalinker followed by baking at 80 °C for 3 hours.

2.3.17.5. Hybridizations

The membrane was prehybridized by gentle shaking at 65 °C for 15 minutes in hybridization solution (see appendix C.1) with Herring sperm DNA (50mg/ml) which had been denatured for 10 minutes at 100 °C in a sealed plastic wrap. The membrane was than hybridized to PCR generated 32P labeled ATP probes at 65 °C overnight. The membrane was washed in 2X SSC, 0.1 % SDS (2 x 5 minutes), 1X SSC, 0.1 % SDS (15 minutes), 0.1X SSC, 0.1 % SDS (2 x 10 minutes) at 65 °C , than autoradiographed for 3days at -80 °C using Kodak X-Ray film.

2.3.18. In Situ Hybridization

Protocol for vegetative buds fixation was optimized for *Begonia* species. 1X PBS (200 ML) was made from 10X PBS stock solution and pH was adjusted to 11. The solution was heated upto 60-70 ˚C. 8g Paraformaldehyde (4%) was added to it and solution was moved onto ice. Once cooled to 4 ˚C, pH was adjusted to 7 by using H_2SO_4 (added drop by drop). 2 ML 10% Triton X-100 and 20 ML DMSO was added to it. Vegetative buds were left in the fixative over night under vacuum.

Note : The recipes of all solutions used in In Situ Hybridization are presented in appendix C.2.

2.3.18.1. Probe Transcription

To synthesize the probes (C- terminal domain) for in situ hybridization, DNA templates were amplified by PCR from genomic DNA using primers incorporating the T7 RNA polymerase, SP6 RNA polymerase and T3 RNA polymerase binding sites. The PCR was carried out as described in 2.3.13.2. The primers sequences and PCR conditions are described in table 13. The PCR products were ethanol precipitated as described in 2.3.17.3 and were used as templates for probe transcription.

For making probes ~800 ng of PCR products/reaction were used as a template and mixed with 2.5 μL of 10X transcription buffer (Roche), 1 μL of RNAsin Plus, 2.5 μL of 10X DIG (Digoxigenin) RNA labelling mixture (Roche), 2 μL of T7/T3/SP6 RNA polymerase (Roche) and final volume is made 25 μL per reaction by adding RNase-free H_2O. All reagents were mixed well and incubated both the sense and antisense tubes at 37°C for 4 h followed by testing 1μl transcribed sense and antisense probes on gel. Than 2 μL of RQ1 DNAse and 75 μL of RNase-free H_2O was added to both probes and incubated at 37°C for 30 min. 1 μL of each probe was checked on a 1% agarose gel. An RNase-free gel is not required at this step. The transcribed probes were ethanol precipitated as described in 2.3.17.3 and the pellets were resuspended in 11 μL of RNase-free H_2O. 1 μL of RNAsin was added to both the sense and antisense RNA probes and stored at –80°C until slides are ready for hybridization.

2.3.18.2. Probe Hybridization (Day 1)

First, on day 1 of probe hybridization 150 mL of 1X proteinase K buffer (see appendix C.2) was prepared in a 250-mL flask and incubated at 37°C until needed in a later step. The slides containing tissue samples of interest were placed in the slide holder and incubated in the series of histoclear (2X Histoclear each for 10 minutes) and different strengths of ethanol (100%, 95%, 90%, 80%, 60% and 30% for two minutes in each solution) followed by two minutes incubation in H2O. The ethanol solutions were saved to be used in later steps. The slides were incubated in 2X SSC for twenty minutes, 0.2 M HCl for twenty minutes and two times in H2O for two minutes each. 15 µL of proteinase K was added from the 10-mg/mL stock to the 150 mL of 1X proteinase K buffer made at start and slides were transferred into it for forty five minutes at 37°C. During slide incubation 1X PBS, 0.2% glycine, 4% PFA (Paraformaldehyde) and 0.1 M TEA (Triethanolamine) were made. Slides were than transferred into 0.2% glycine and two times in 1X PBS for two minutes each followed by an incubation in 4% PFA for ten minutes and again two times in 1X PBS for five minutes each this time.

Table 13 Primers and conditions used to generate probes for In Situ Hybridization

Primer Names	**Primer Sequences**	**Annealing temperature (^oC)**	**Extension time (min)**
Sense T3 *BARP*	AATTAACCCTCACTAAAGGGTCCA TGTGTTCAAGAAACTTACTC	54	2
Sense SP6 *BARP*	CGATTTACCTCACACTATAGAACC ATGTGTTCAAGAAACTTACTC	54	2
Antisense T7 *BARP*	TAATACGACTCACTATAGGGAAGC AGCAGAGAGAACAAAAGG	54	2
Sense T3 *KNB1*	AATTAACCCTCACTAAAGGGTGCC ATTAGGTTGTGAGGAAGGTGA	55	2
Sense SP6 *KNB1*	CGATTTACCTCACACTATAGAATG CCATTAGGTTGTGAGGAAGGTGA	55	2
Antisense T7 *KNB1*	TAATACGACTCACTATAGGGTGTC GAGTCCACCAATCCAGCAA	55	2
Sense T3 *KNB2*	AATTAACCCTCACTAAAGGGGAGC AGGTTGCGAGGGGG	55	2
Sense SP6 *KNB2*	CGATTTACCTCACACTATAGAAGA GCAGGTTGCGAGGGGG	55	2
Antisense T7 *KNB2*	TAATACGACTCACTATAGGGTGCC TCGCCTCTTTCGGCAG	55	2

Slides were incubated in stirred TEA (Triethanolamine) solution in vacuum hood and 750 µL of acetic anhydride was added during stirring using a 1-mL glass pipette. The slides were than washed with 1X PBS twice for five minutes each in the hood followed by washing with ethanol series (100%, 95%, 90%, 80%, 60% and 30% for thirty seconds in each solution). A square bioassay tray was prepared by covering it completely with aluminum foil and placing three to four paper towels wet with 100% ethanol inside the dish and four 1-mL pipettes on the bottom of tray to hold the slides in the horizontal position until probes are ready for doing hybridization. 1000ng, 800 ng, 700 ng, 500 ng and 300 ng of each RNA probe was mixed with 50% formamide solution to a final volume of 40 µL and mixed well by avoiding making bubbles followed by heating at 100°C for 5 minutes. The samples are moved on ice until needed. Hybridization solution was prepared as described in appendix C.2 and kept at 85°C. Meantime the hybridization dish is prepared by placing three to four folded towels wet with RNase-free water in the bottom of a glass and 1-mL pipettes were set up to hold slides in the horizontal position. 160 µL of hybridization solution kept at 85°C was mixed with each RNA probe on ice and applied to slides one at a time and slides were covered with a HybriSlip coverslip with the help of forceps. It is crucial to avoid making bubbles at this stage. The dish was sealed thoroughly with plastic wrap and incubated at 52°C-55°C overnight in a hybridization oven.

2.3.18.3. Washes (Day 2)

On the 2nd day slides were removed from hybridization oven carefully and washed with 0.2X SSC followed by incubation in 0.2XSSC in the hybridization oven on a rocker at 55°C for 1 h twice. NTE, BM (Block Master) solution and block solution 2 were made (see appendix C.2 for solutions composition) during slide incubation to use in subsequent steps. 100 mL NTE was prewarmed at 37°C until needed and 100 µL of RNase A (20 mg/mL) should be added just before using this solution. The slides were transferred into NTE solution for five minutes twice followed by 30 minutes incubation in prewarmed NTE solution at 37°C. The slides were than again washed in NTE solution for five minutes twice. The slides were again washed in prewarmed 0.2X SSC in the hybridization oven on a rocker at 55°C for 1 h followed by two washes with1X PBS for ten minutes each. The slides were incubated in BM solution twice for 45 minutes and 30 minutes at room temperature. The slides were than transferred to block solution 2 for 45 minutes at room temperature. During slide incubation the anti-DIG-AP antibody was made by diluting the antibody 1:1250 in block solution 2 (8 µL of antibody from stock in 10 mL of block solution 2). The slide chamber was prepared by placing three to four paper towels soaked in water in the square bioassay dish and four 1-mL pipettes on the bottom to hold slides in the horizontal position. Slides were placed in a square bioassay dish and 400 µL of antibody solution applied to each slide and incubated for 2 hours at room temperature. Slides were than washed twice with block solution 2 for 20 minutes followed by overnight incubation in block 2 solutions.

2.3.18.4. Developing (Day 3)

The slides were washed with block solution 2 for 20 minutes followed by two washes in Buffer C for fifteen minutes and ten minutes. 24 µL of 1 M levamisole was added to the alkaline phosphatase substrate solution for a total volume of 30 mL and kept in a Coplin jar covered with aluminum foil. The slides were incubated in this solution until signal developed (which depends on specific gene, probe concentration, tissue type and species) which was 2.5-3.5 hours for *KNB* genes and 1-2 hours for *BARP1* genes in *Begonia*. Once signals are fully developed the reaction was stopped by dipping slides in TE buffer for two minutes followed by washing with water for two minutes. The slides were dehydrated through an ascending graded series of ethanol (30%, 50%, 80%, 90%, 95%, 100% and 100% for 30 seconds in each solution) and twice into fresh Histoclear for 2 minutes each. Finally slides were mounted in Permount and dried in the hood before taking photographs.

Chapter 3

Genetics of leaf form in *Begonia*

3.1. Introduction

Genetic analysis of model organisms has identified a number of candidate genes that may have contributed to the evolution of phenotypic differences between species (Bharathan *et al.*, 2002; Colosimo *et al.*, 2005; Sucena *et al.*, 2003). This has led to the hypothesis that similar genetic changes may have given rise to the existing interspecific variation in *Begonia*. This hypothesis can be tested by comparing the regulation and function of these candidate genes among species. Such comparisons can be carried out by interspecific genetic analysis, association mapping in crosses and expression analysis.

Leaves are the fundamental organs of land plants and their photosynthetic activity is the basis for most life on earth. Despite their uniform function - to capture light and exchange carbon dioxide for oxygen and water - there is considerable variation in leaf form. Some of this variation is a plastic response to the environment (Billings and Mooney 1968). This is particularly the case for variation in size whereas shape variation tends to be more independent of the environment and more heritable (Dickinson *et al.*, 1987, McLellan, 2000).

In some taxa heterophylly can induce marked phenotypic differences in leaf form (contrasting environmental conditions during development). An example of this is the leaves of emergent and submerged shoots of *Ranunculus flabellaris* (Bostrack and Millington, 1962; Young *et al.*, 1995). Leaf form variation within individuals may also occur regardless of environmental conditions, as part of the normal developmental pattern, notably among sequential leaf positions on a stem (leaf heteroblasty, Greyson *et al.*, 1982; e.g. Gossypium, Hammond, 1941, Stephens, 1945; Nicotiana, Paxman, 1956).

Genetically determined variation in leaf form occurs between many species and sub-species or varieties of plants that can be hybridized with each other, so is feasible to analyze the genetics of naturally occurring variation in leaf form. The genetics of natural leaf form diversity have been described in several comparative studies of leaf development (Kaplan, 1970; Merrill, 1979; Whaley and Whaley, 1942; Gurevitch, 1988; Kimura *et al.*, 2008; Hay & Tsiantis, 2006, Harevan *et al.*, 1996; Takahiro *et al.*, 2010).

Crosses between closely related species *Mimulus lewisii* and *M. cardinalis* have identified genetic loci of large effect that modify the floral architecture (Bradshaw *et al.*, 1995); indicating the role of one major QTL for the differences between flowers of *Mimulus lewisii* and *M. cardinalis* (Bradshaw *et al.*, 1998). It was shown later that *YELLOW UPPER* (*YUP*) which regulates the presence or absence of yellow carotenoid pigments in the petals of pink-flowered *Mimulus lewisii* and its red-flowered sister species *M. cardinalis* is the said major locus. This *YELLOW UPPER* (*YUP*) strengthens reproductive isolation in these two species as the difference in flower colour has strong effect on visitation of the flowers by different pollinators (Bradshaw and Schemske, 2003). Similarly crosses between the plants of *Aquilegia formosa* and *Aquilegia pubescens* revealed the genetic basis for reproductive isolation between these two species was due to differences in the length and colour of spurs affecting the types of pollinators visiting the plants (Miller, 1981; Hodges *et al.*, 2002).

Leaf form is an excellent system to study the genetic basis of morphological diversity as leaf form variation occurs at every hierarchical level within and between individuals, populations, and taxa. West and Noble (1984), Dickinson (1983, 1986) and Dickinson and Phipps (1984) have studied leaf variation within diverse genera. Among species within genera leaf variation was studied by Landon and Blum, 1978;

Parker *et al*., 1979, 1981; El- Gazzar, 1980; Phillips, 1983; Phipps, 1983; Parker and Maze, 1984, and in taxonomically diverse collections (Hill, 1980).

Many genes that effect leaf morphogenesis have been characterized in *Arabidopsis thaliana* and other species through modification of their expression by mutation or over expression (Kidner and Timmermans, 2010; Kidner and Umbreen, 2010; Koening and Sinha, 2010). Genetic evidence from a range of species supports a 1/2 locus model for leaf dissection as shown in table 1.

Multiple independent origins of compound leaves through plant evolution have been attributed to the differential expression of *KNOX* genes (Bharatan *et al*., 2002; Gleissberg, 2004). *KNOX* gene expression in the leaf primordium is correlated with the development of deeply lobed leaves and even simple leaves with toothed margins (Bharathan *et al*., 2002). One elegant example of the role of *KNOX* genes in the evolution of leaf form is shown by Hay and Tsiantis (2006) where the differences in leaf morphology between *Arabidopsis* and its relative *Cardamine hirsuta* is due to variation in the regulation of *KNOX* genes. In another example over-expression of *KNOX* genes is responsible for the increased complexity of compound leaves in *Solanum galapagense* in comparison to its less dissected sister species *Solanum cheesmaniae* (Kimura *et al*., 2008).

A second key developmental regulator which has been shown to be important in species level differences is the *YABBY* and related gene. Interspecific hybrids analysis in two closely related Juncus species, *Juncus prismatocarpus*, which has flattened unifacial leaves, and *Juncus wallichianus*, which has cylindrical unifacial leaves, has provided the decisive evidence for the role of *DROOPING LEAF* (*DL*), a member of the *YABBY* gene family for regulating flatness in the unifacial leaf blade of *Juncus prismatocarpus* (Yamaguchi *et al*., 2010).

Studies have shown a role for *ARP* genes in controlling the variation in leaf form while controlling the extant and the position at which the lamina grows out (Kim *et al*., 2003a; Kim *et al*., 2003b). However peltateness has not been studied as a genetic trait. Peltate/cordate leaves occur mostly in the woodland understories and are thought to be an adaptation to the low light conditions as they are optimized for unidirectional light (Givnish, 1986). They are the best mechanical solution to holding out a 'flat plate on a stalk' as the peltate shape held out the leaf on erect petioles to capture maximum light with minimum of expensive supportive tissues (Givnish, 1986).

Peltate leaves have probably evolved at least eight times and compound leaves at least three time in *Begonia* genus (Forrest, 2000). The replication of the evolution of peltate and compound leaves makes this a good system to study the genetics of morphological variation. We have used F1 and F1BC populations to determine the genetic behaviour of these two traits in *Begonia*. My hypothesis is that relatively few loci will control these traits in each species, and that those loci could be *ARP, KNOX, YABBY* or *CUC* family members.

3.2. Leaf Development in *Begonia*

An understanding of the mechanism of leaf form determination in any species requires the knowledge of early morphogenesis of leaf development. Nearly all species of *Begonia* have asymmetrical leaves (Smith *et al*., 1986) and this asymmetry is established soon after the initiation of leaf primordia. The leaf blade maintains this asymmetry throughout development (McLellan, 1990; Barabe *et al*., 2007). In *Begonia* the leaf primordium is initiated as a bump which ultimately occupies a significant portion of the shoot apex. A pair of stipules is present at the base of each leaf. Both stipules nearly enclose the next leaf primordium that is initiated on the SAM apex (Figure 12). The time between the initiations of two primordia is reported to be between 15 and 20 days (Barabe *et al*., 2007). The difference in shape of simple, peltate and compound leaves is apparent from the P2-P3 stage of leaf development in *Begonia* section Gireoudia (Figure 13). There is a deep notch in peltate leaves at P1. This area of the leaf grows more than surrounding regions, filling in the notch as the leaf grows. At maturity the leaf is fully peltate. This suggests that peltateness increases with the age of the leaf. This may indicate that

Begonia leaves have tendancy to develop lamina outgrowths at the blade/petiole attachments points as the leaf matures.

Neelima Sinha's lab has shown that among 25 angiosperm families, 289 species with pinnate compound and 153 species with non peltately palmate compound leaves have distinct ab-adaxial symmetry in petioles. The petioles of 56 species with peltately palmate compound leaves show radial symmetry with vascular bundles arranged in a ring. This suggests that petiole abaxialization is correlated with peltate leaves whether entire or compound (Kim *et al*., 2003).

Sections of peltate, non peltate and compound leaves from fourteen species of *Begonia* section Gireoudia species showed abaxialization of the petioles (Figure 14). As we expected to see a clear adaxial domain in the petioles of non peltate simple and compound leaved species, I looked closely at the petiole attachment points of non peltate *Begonia* species and found a laminar outgrowth in all the species (Figure 15 and 16) which have abaxialized petioles (Figure 14). *Begonia* species which lack laminar out growth had an adaxial domain in their petioles (Figure 14) which is consistent with the previous finding (Kim *et al*., 2003).

3.3. Quantification of leaf form

I quantified leaf form in *Begonia* section Gireoudia by taking the measurements shown in Figure 17 and derived summary formulae to describe the different aspects of leaf form. The ratio of width (B) to length (A) describes the circularity of leaf and peltateness is equal to the distance from petiole attachment to the leaf boundary (D) divided by half of the leaf diameter (A). Eccentricity of peltateness is the distance of petiole attachment to the centre of the leaf (C) divided by half of the leaf diameter (A) and the ratio of notch length (E) to leaf diameter determines the notchiness (A) (Figure 17). On the basis of this quantification I have described the variation of peltateness within and between plants of the same species and interspecific variation for peltateness in *Begonia* section Gireoudia in a later section of this chapter.

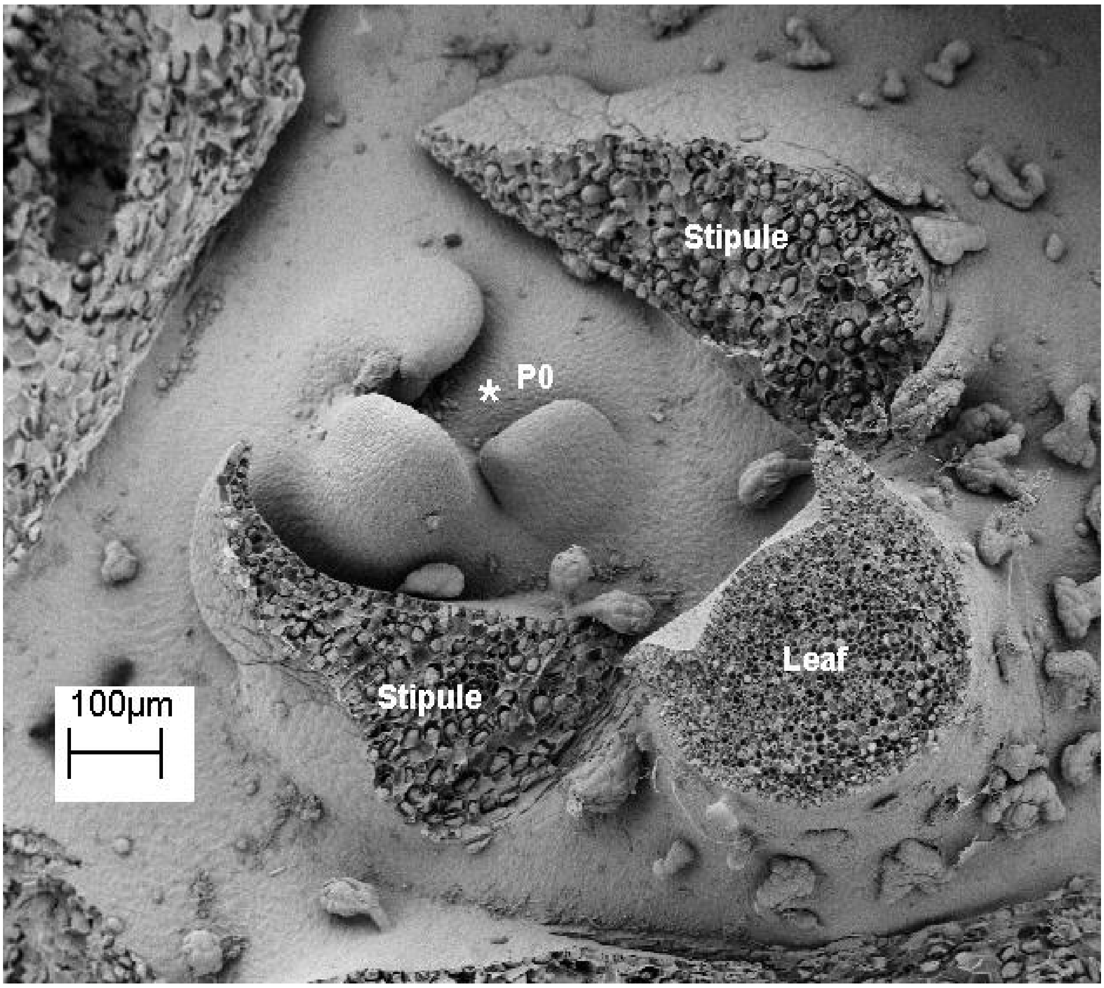


Figure 12. SEM micrograph of the vegetative meristem of *Begonia herbacea*.
***Begonia herbacea* has flat apices unobscured by trichomes and easy to visualize in the SEM. Star marks the position of SAM and P0 the region from which the next leaf primordium will develop. The P1 leaf has developed quite extensively and its paired stipules encircled the meristem. The leaf has become flattened and started to curve**

over, indicating dorsoventral differentiation has occurred. The P2 leaf and stipules have been removed. Photo is taken by C. Kidner.

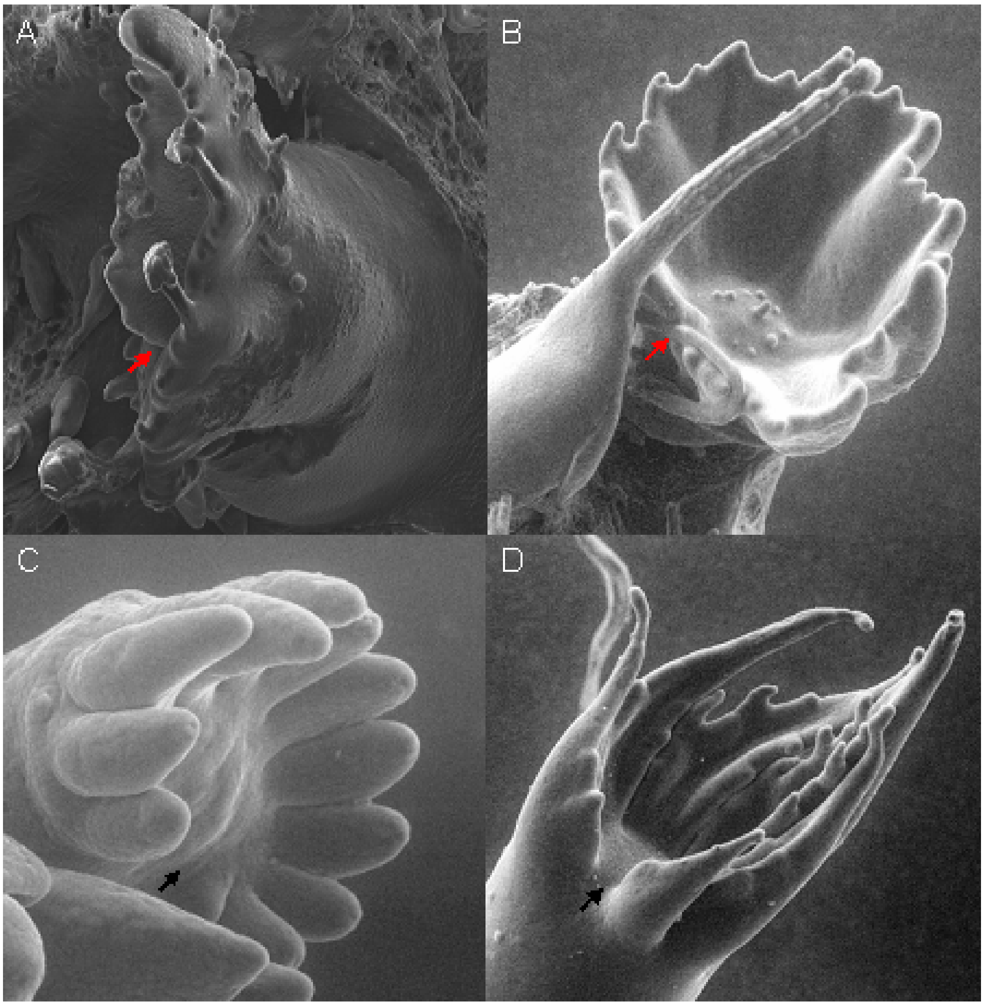


Figure 13 SEM micrograph showing differences in lamina growth in peltate and non peltate leaves during early leaf development.

(A*)* Peltate leaved *B. peltata*, (B) *B. pruniata*; non peltate leaved (C) *B. luxurians* and (D) *B. carolineifolia*. The red arrows in B and C indicate a clear notch in peltate leaves at P1 stage which is replaced with laminae tissue as leaves mature. Black arrows indicate the lack of laminae tissues at P1 of *B. luxurians* and (D) *B. carolineifolia*. (Pictures B, C and D are taken by C. Kidner).

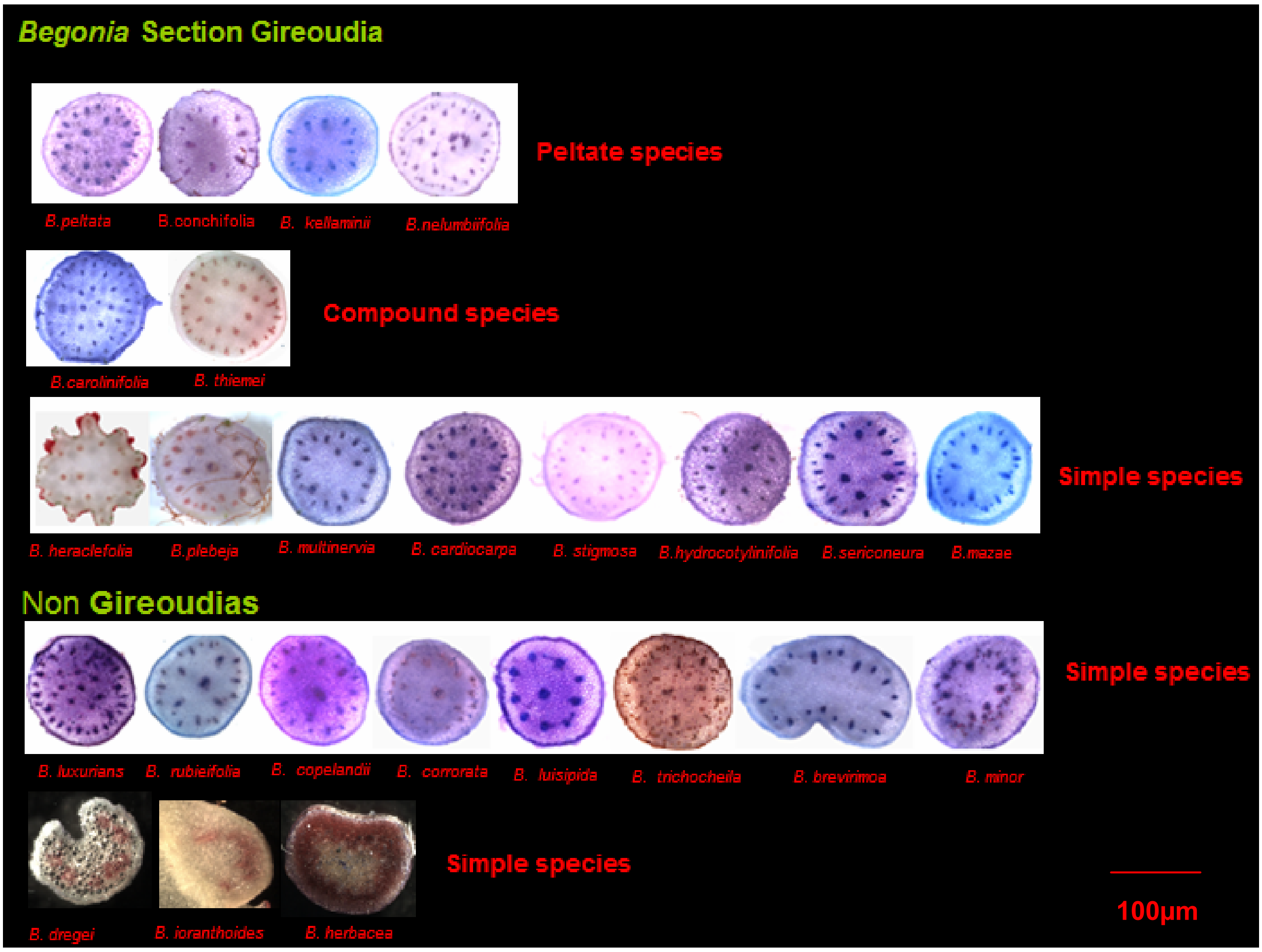


Figure 14 Petioles scetions of *Begonia* species. All *Begonia* section Gireaoudia species have radial petioles with vascular bundles arranged in a circle. Scale bar is 100 µm. Some non Gireoudia species (*B. brevirimoa, B. dregei*, *B. ioranthoides* and *B. herbacea*) have dosoventrarlity in their petioles.

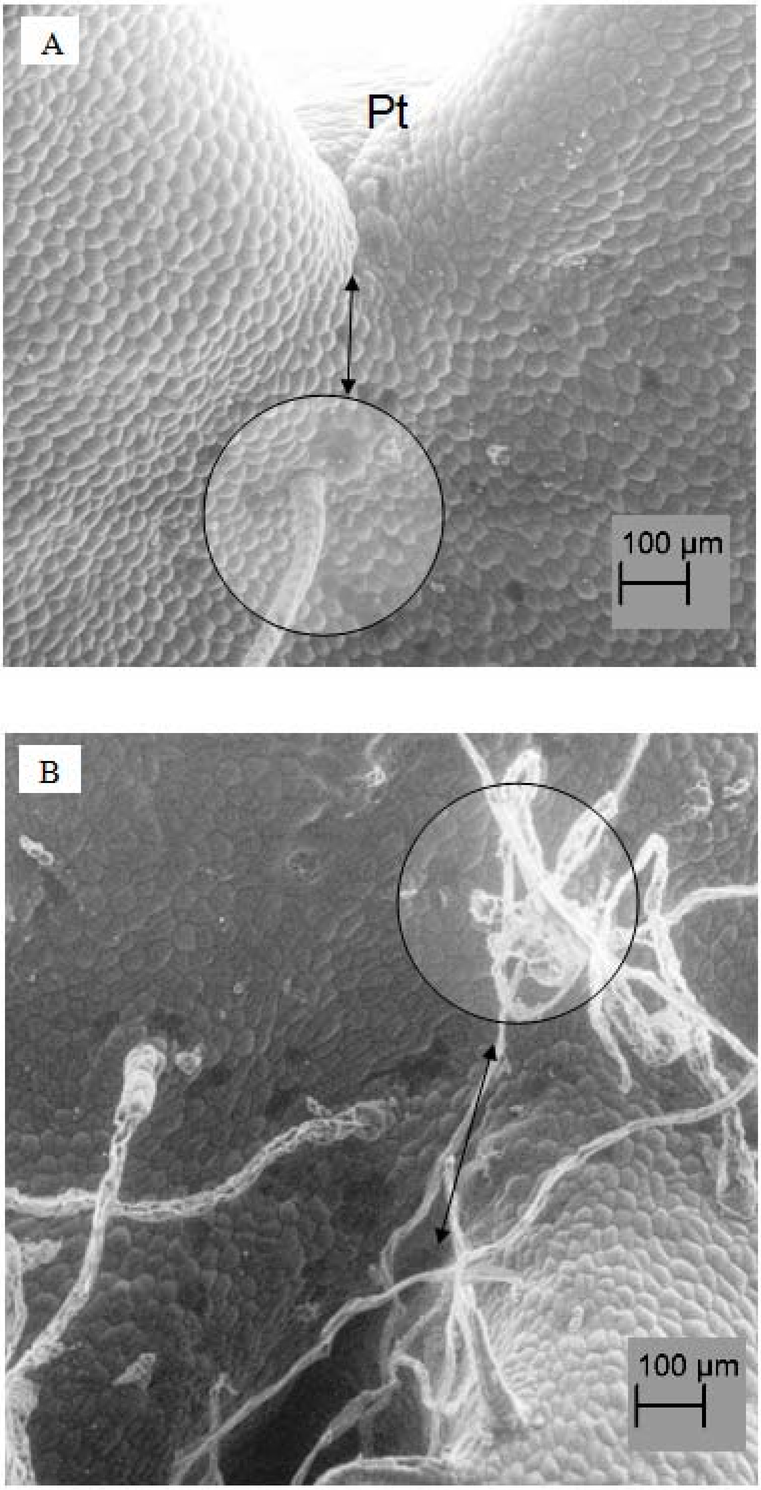


Figure 15 SEM showing outgrowths at lamina-petiole attachments points in non peltate *Begonia* species.
(A) *B. mazae* (B) *B. plebeja*. Pt represents Petiole. Circle shows the approximate area of petiole attachment, Arrow represents the area of outgrowth at lamina-petiole attachments points.

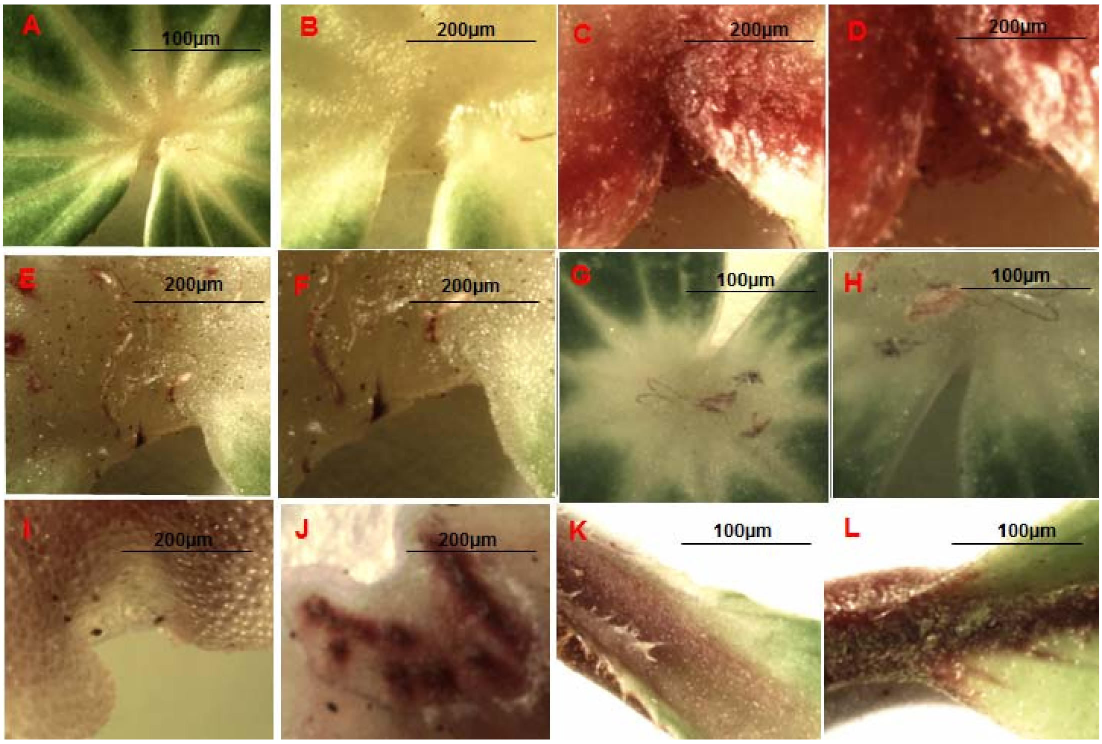


Figure 16 Lamina outgrowths in non peltate *Begonia* species.

Compound microscopy images for *B. mazae* (A & B); B. plebeja (C & D); *B. heracleifolia* (E & F); *B. stigmosa* (G & H); *B. dregei* (I & J); *B. herbacea* (K) and *B. ioranthoides* (L). Gireoudia species A-H indicate the lamina growth at the leaf- petiole attachment points in non peltate species. I-L show lack of lamina growth in non peltate *Begonia* species.

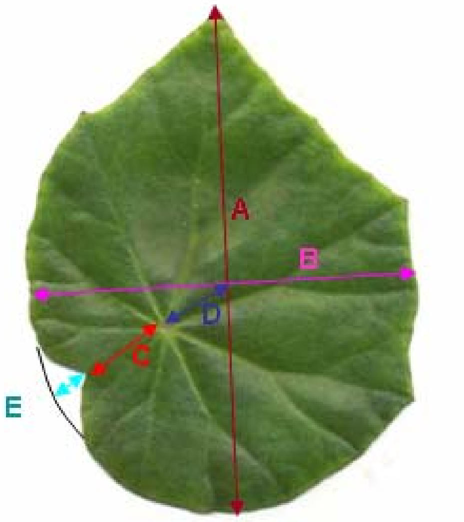


Figure 17 Quantification of leaf form in *Begonia* section Gireoudia species

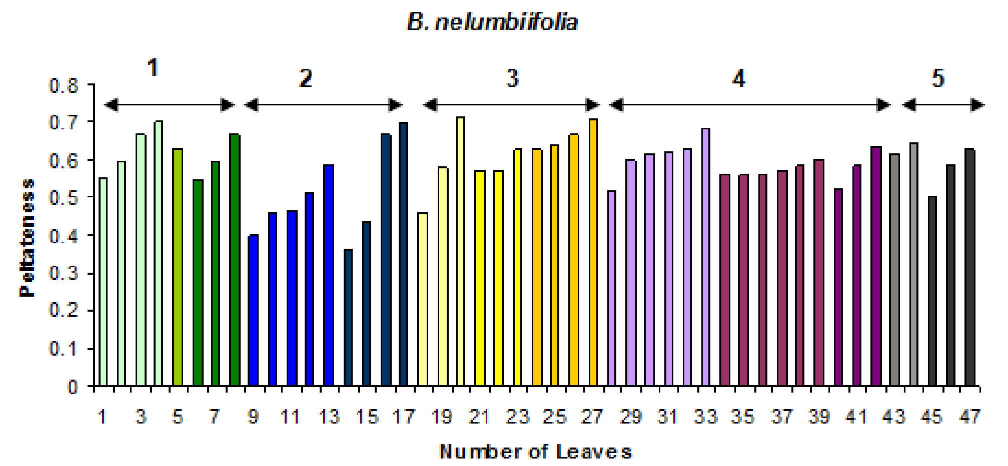
3.4. Within and between plant variation for peltateness

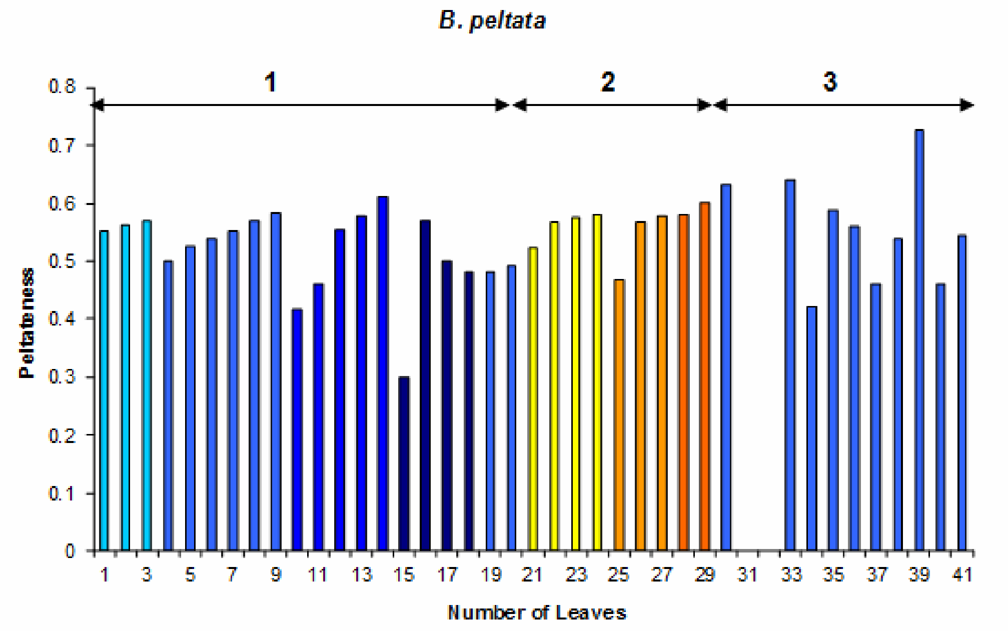
Peltateness increases with the age of leaves in *B. peltata, B. nelumbiifolia, B. conchifolia, B. pruniata* and *B. kellemanii*. The younger leaves at the end of each rhizome are not yet fully expanded and are not yet fully peltate. All fully expanded leaves show high peltateness (Figure 18).

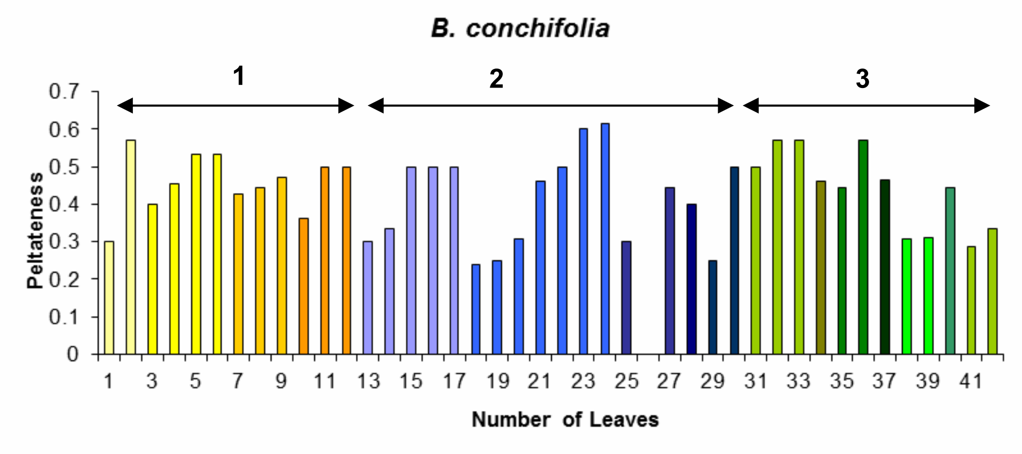
Once sources of within-individual variation in leaf form have been accounted for, other patterns of variation may be revealed. Peltateness varies between genetically identical clones of the same species kept in different greenhouses which reveals the influence of environment on the expression of peltateness (Figure 19). The analysis of variance (ANOVA) is carried out in PAST (PAleontological STatistics) and the heritability is calculated by dividing between groups variation (between different species) with the sum of between groups variation and within groups variation (between genetically identical clones of the each species. The heritability value for *B. peltata* is 99.95%, *B. nelumbiifolia* is 77.75%, *B. conchifolia* is 25%, *B. pruniata* is 34% and *B. kellemanii* is 75%. This suggests that non genetic factors have a larger contribution for peltateness variance among plants of *B. conchifolia* and *B. pruniata*.

3.5. Inter specific variation for peltateness

Peltateness varies a great deal among different species of *Begonia* (Figure 19). A one way ANOVA (analysis of variance) was carried out using Excel to analyse the differences between and within species. Peltateness differ significantly between the species ($F_{4,10}$=38.52, $P < 0.05$); the two subscripted numbers after F indicate the degrees of freedom for between group and within group variance.







Continued on next page……………….

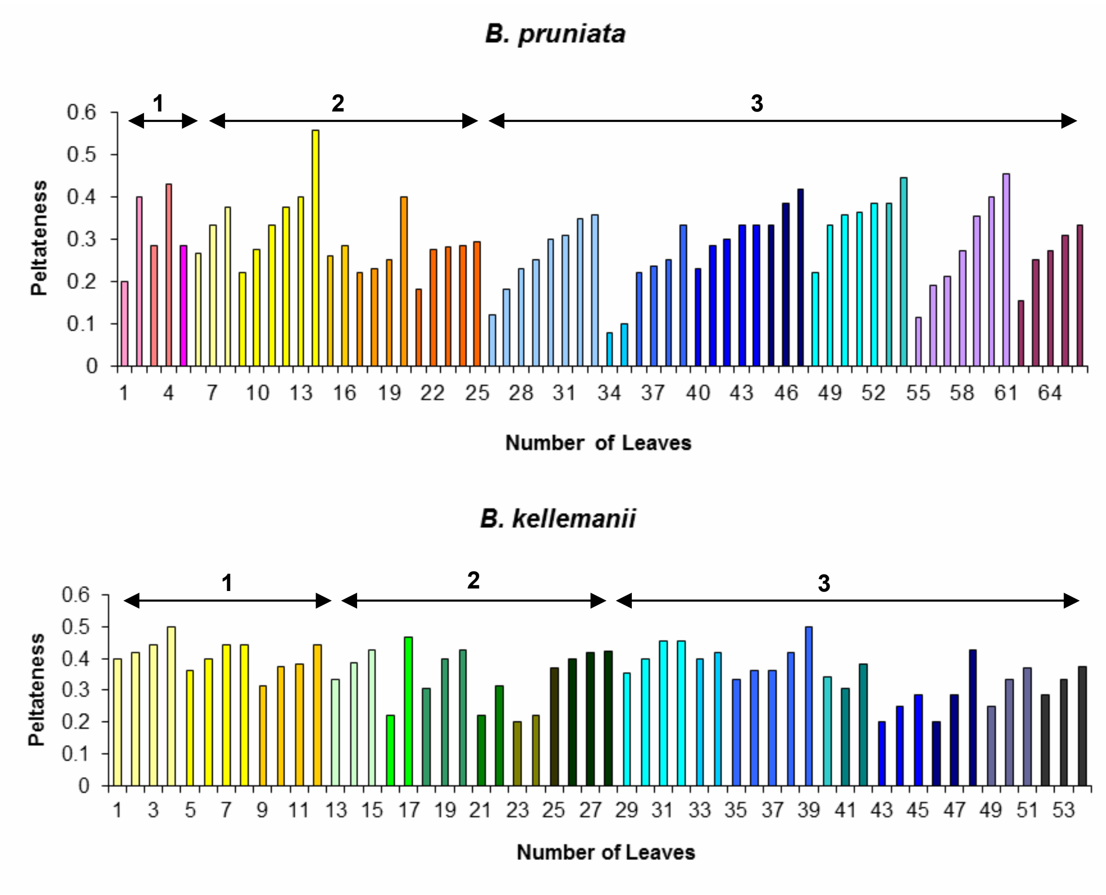


Figure 18 Peltateness increases with the age of the leaf.
Arrows above the graphs unite the measurements from the sample plant and different colours are used for the measurements from different rhizomes. Leaves on each rhizome are in ascending order from bottom of the rhizome towards the tip. There is clear trend of increase in peltateness as the leaves matures along each rhizome of the sample plants.

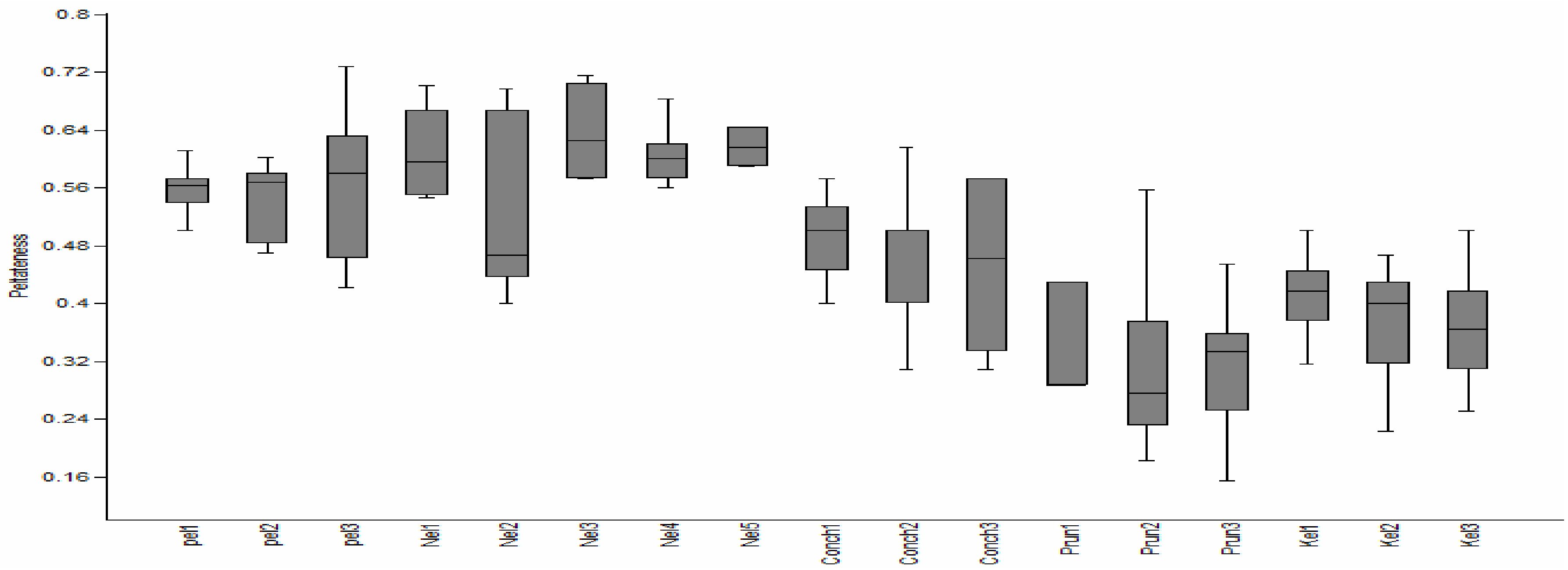


Figure 19 Environmental and interspecific variation in peltateness in *Begonia* section Gireoudia.
The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median. The plants are clones and variation between them is due to non-genetic factors which is strong but variation between different species is very obvious. Pel, Nel, Conch, Prun and Kel are the abbreviation for *B. peltata*, *B. nelumbiifolia*, *B. conchifolia*, *B. pruniata* and *B. kellemanii* respectively.

3.5.1. Peltateness Vs circularity and notchiness

Various aspects of plant morphology are genetically correlated which can explain the genetic basis for the coordinated evolution of characters (Davis, 2001; Conner, 2002; Ungerer *et al*., 2002). Pleiotropy (the association of more than one phenotypic characteristic with a single genotype), and linkage (when independent genes that determine different traits are inherited together because they are located near each other in the genome) are two possible mechanisms for these correlations (Falconer and Mackay, 1996).

There is a no correlation between the leaf circularity and peltateness (r=0.28, p (uncorr) = 9.64) in *Begonia* section Gireoudia species (Figure 20). At RBGE only one of the peltate *Begonia* species (*B. pruniata)* has a notch in mature leaves and there is a negative correlation between notchiness and leaf peltateness r = 0.5, p (uncorr) = 4.755) (Figure 21).

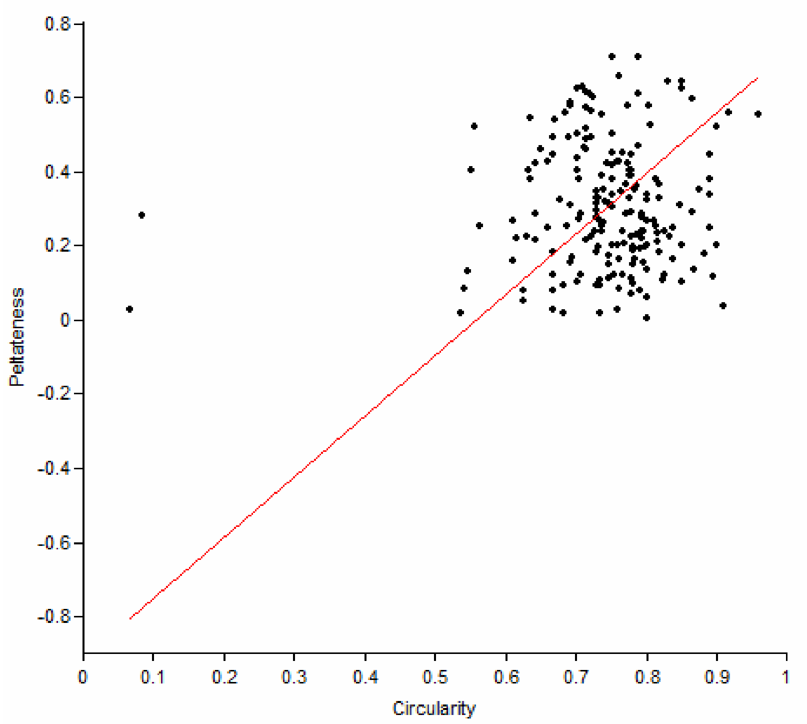


Figure 20 Peltateness is not related to the circularity.
Graph showing a positive association between peltateness and circularity between different peltate species of *Begonia* section Gireoudia.

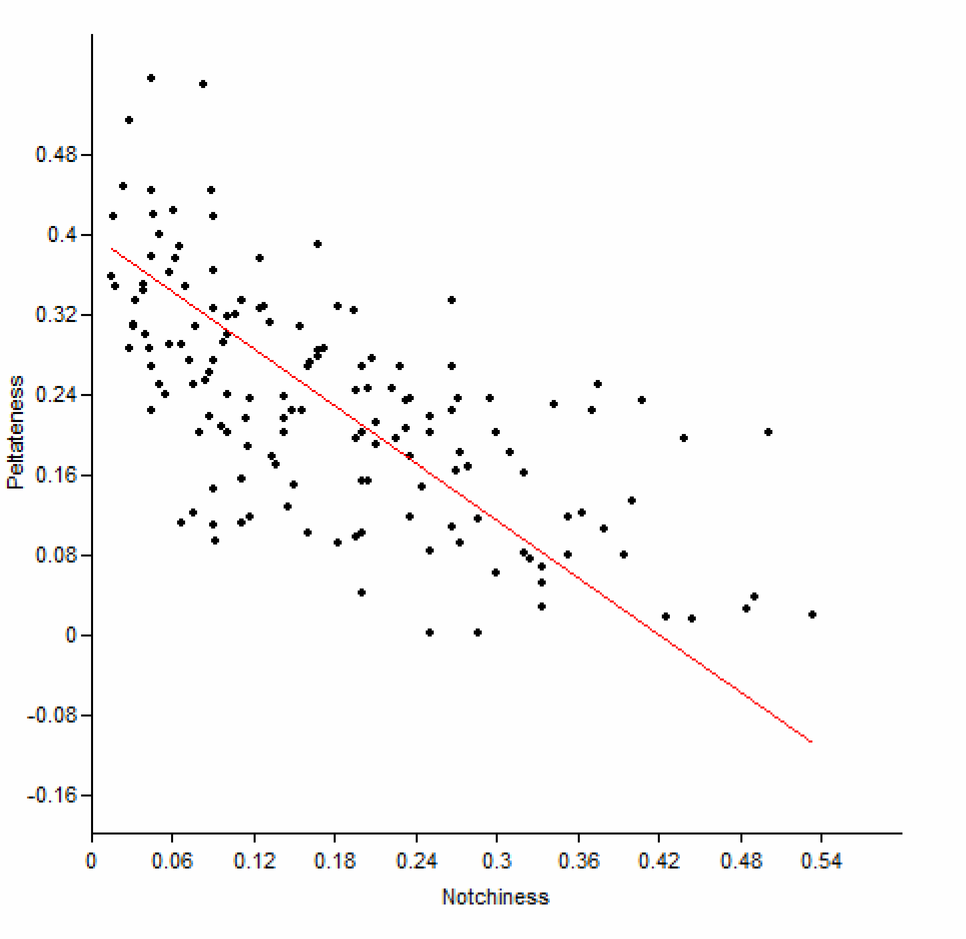


Figure 21 Peltateness is related to the notchiness.
Graph showing a negative association between peltateness and notchiness between different peltate species of *Begonia* section Gireoudia.

3.6. Genetics of peltateness

I used genetic analysis to see if peltateness is a simple genetic trait that behaves the same in each peltate species. F2 and F1BC1 progeny of peltate × non peltate interspecific crosses were generated to carry out genetic analysis. We expect peltateness to behave the same way regardless of which peltate species it comes from if it is controlled by the same locus. However if changes at different loci are involved then the F1 of peltate species could be non peltate and peltateness may behave differently in each F1 cross.

3.6.1. Complementation tests and development of F1 hybrids in *Begonia*

Complementation tests were carried out with peltate species (*B. nelumbiifolia, B. conchifolia, B. peltata* and *B. kellemanii*) to investigate whether the same loci were involved each time peltateness evolved. The peltate F1 hybrids were obtained from crosses between peltate species. It was concluded on the basis of these tests that peltateness does not complement (Figure 22). This indicates that same gene may be controlling peltateness in each species or it could be due to the peltate phenotype being dominant in each cross.

To study the dominance of peltateness and determine if it behaves as a single Mendelian trait ten non peltate species (*B. heracleifolia, B. plebeja, B. cardiocarpa, B. lindlyana, B. carolineifolia, B. multinervia, B. sericoneura, B. mazae, B. stigmosa* and *B. theimei*) were crossed with five peltate species (*B. nelumbiifolia, B. conchifolia, B. peltata, B. pruniata* and *B. kellemanii*). Not all the F1s were fertile. Analysis of behaviour of peltateness in those that were fertile gives clear insight into the genetics of the trait. Reciprocal crosses gave the same results indicating lack of involvement of cytoplasmic factors in leaf form determination.

Peltateness behaved differently in all F1s of these crosses (Figure 23). Peltateness behaves recessively in fifteen crosses and dominantly in the other twelve crosses. This suggests that either peltateness is the result of different genetic changes in each species or it could be due to the modifiers (it may be epigenetic factors) from the non peltate species in each F1. Peltateness is a juvenile trait in many of the F1 hybrids with peltate parents. Leaf form changes during the development of F1 hybrids. Some F1 hybrids were peltate at juvenile stage and later on changed into non peltate leaves and vice versa (Figure 24). Further F1 hybrids showed different degree of peltateness. The leaf form varied from semi peltate leaves to eccentrically peltate leaves. This suggested that peltateness is not a discrete trait in F1 hybrids and indicates the involvement of many genes in controlling peltateness.

3.6.2. Peltateness in *B. nelumbiifolia* × *B. heracleifolia*

The F1 hybrid of *B. nelumbiifolia* × *B. heracleifolia* (peltate and non peltate species respectively) is non peltate which indicates the recessive behaviour of peltateness in this case (Figure 25). Among F1 hybrids two out of ten juvenile plants were non peltate whereas other eight were eccentrically peltate. It was observed that juvenile eccentrically peltate plants later developed non peltate leaves. This indicated that there is a change in expression of some genes during the developmental process which resulted in a change in leaf form. In adult plants, therefore peltateness was recessive.

A back cross of *B. nelumbiifolia x B. heracleifolia* with one of the parents *B. nelumbiifolia* showed the segregation of some characters and all have peltate leaves (Figure 26 and 27). If a single recessive gene controls peltateness then we expect fifty percent to be eccentrically peltate as juveniles. The fact that all progeny (eight plants) which survived to adulthood are peltate shows that a single recessive gene is not a good explanation for the evolution of peltateness in *B. nelumbiifolia*. Instead it indicates the presence of several possibly semi dominant loci in *B. nelumbiifolia*.

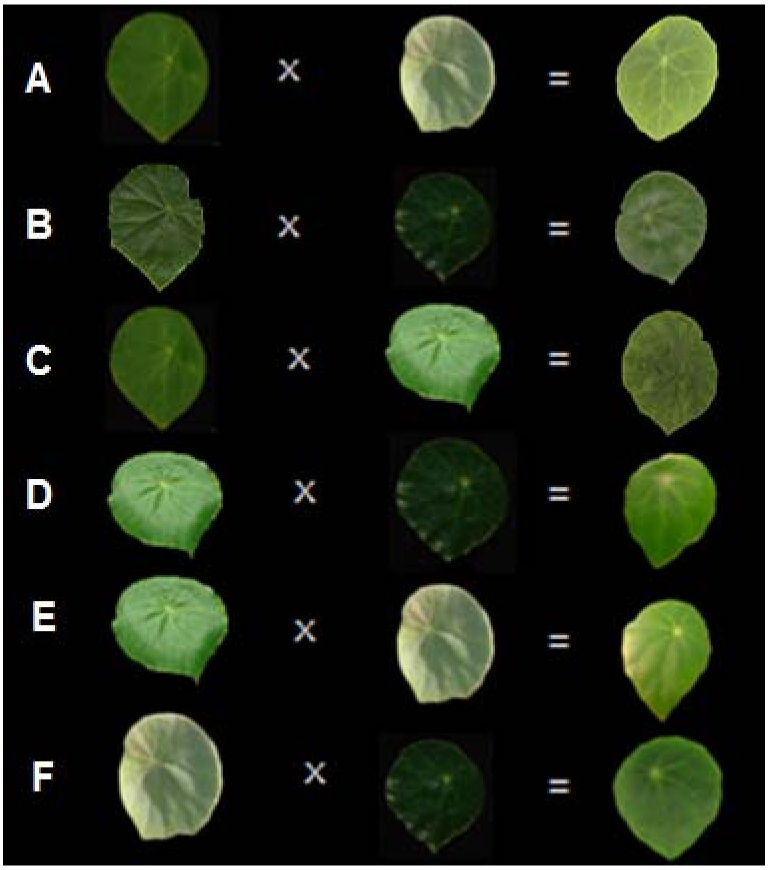


Figure 22 Peltateness does not complement in *Begonia* section Gireoudia.
Leaf form in F1s from crosses between *B. kellemanii* and *B. peltata* (A) *B. pruniata* and *B. conchifolia* (B) Cross between *B. kellemanii* and *B. nelumbiifolia* (C) Cross between *B. nelumbiifolia* and *B. conchifolia* (D) Cross between *B. nelumbiifolia* and *B. peltata* (E) Cross between *B. peltata* and *B. conchifolia* (F)

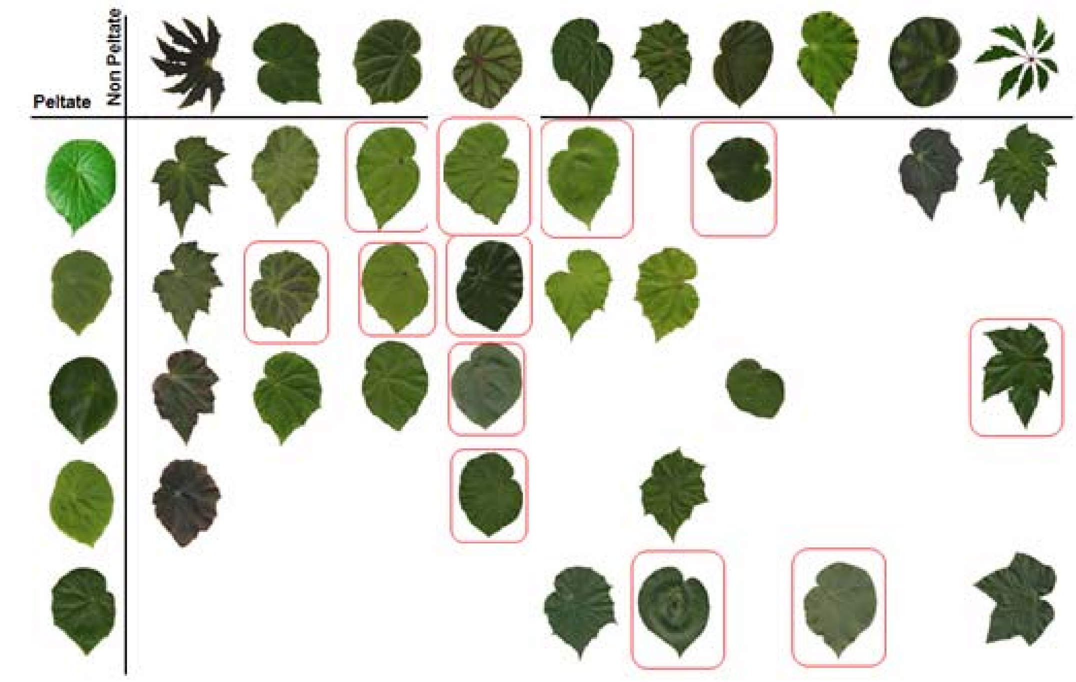


Figure 23 Leaf phenotypes in crosses between non peltate and peltate species of *Begonia*.
The red boxes represent peltate F1 hybrids. Peltateness is dominant trait in 10/27 F1s and recessive in others. This indicates that peltateness is a complex trait and may be controlled by more than one locus.

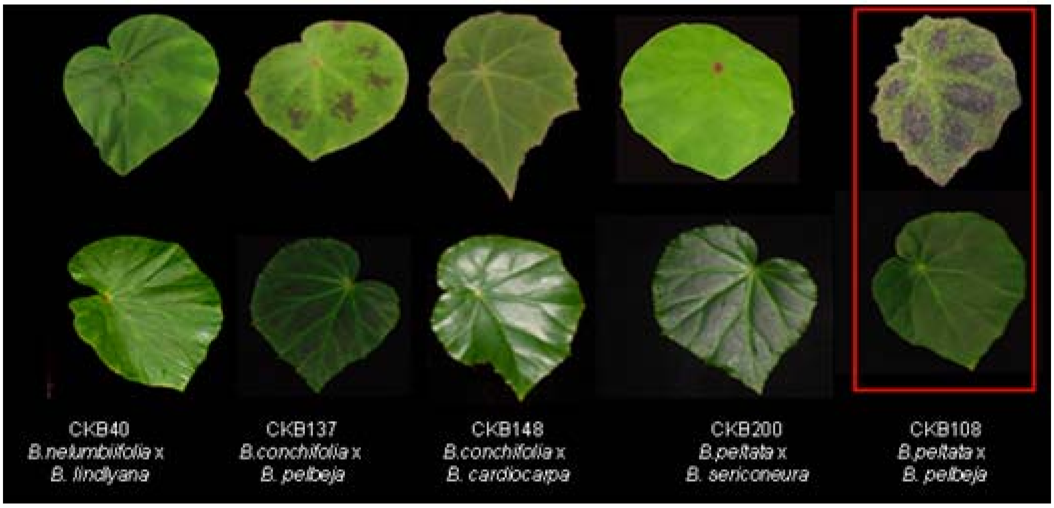


Figure 24 Changes in leaf form during development.
Top leaves are younger (approximately P5) and bottom ones are older (approximately P10). A change from peltate to non peltate leaves occur in the first four leaves. The red box represents the change of non peltate into peltate leaf form.

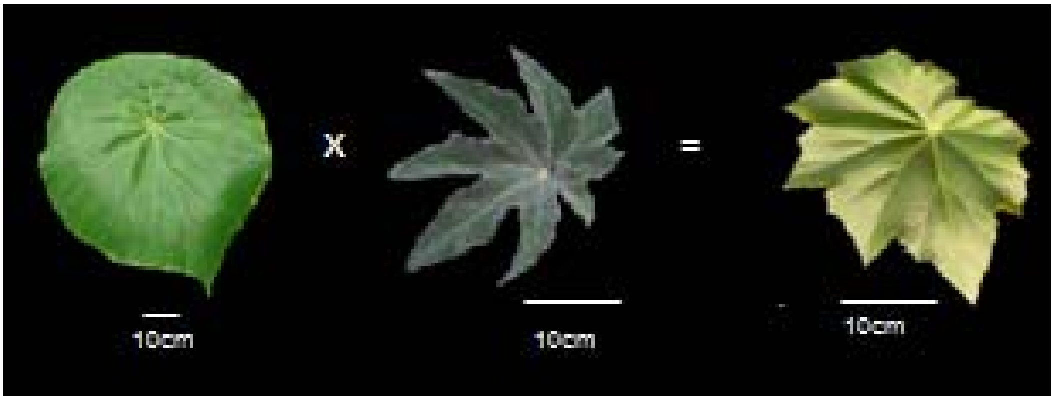


Figure 25 Cross between *B. nelumbiifolia* x *B. heracleifolia* yields a non peltate F1 hybrid (CKB44).
This suggests that peltateness is a recessive trait. The leaf size is indicated with an approximate scale.

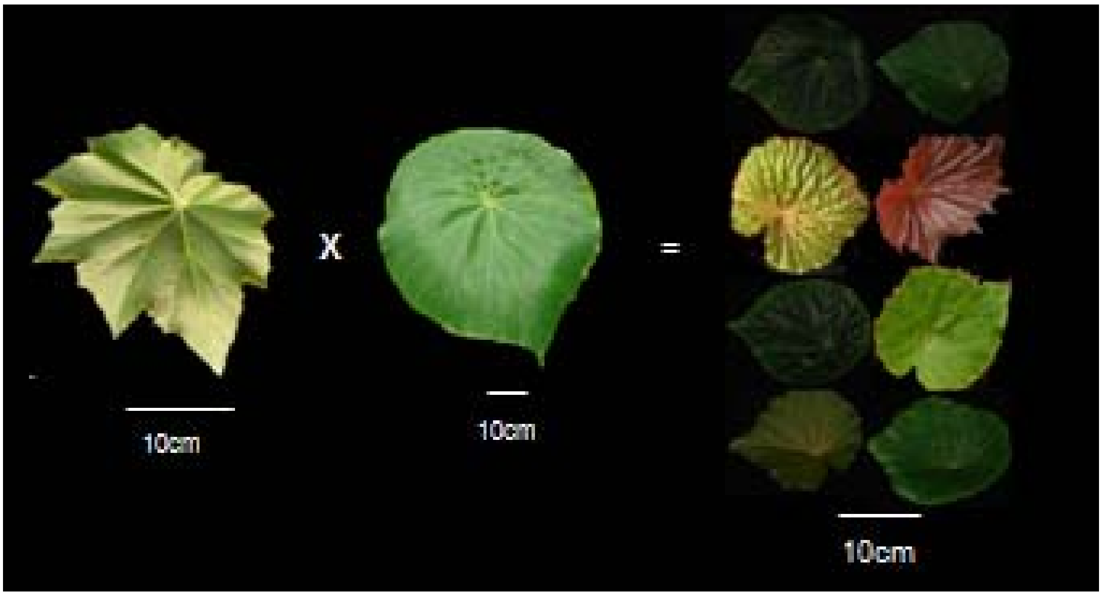


Figure 26 Cross between CKB44 and *B. nelumbiifolia* yields plants with peltate leaves. This indicates that peltateness is a complex trait and may be controlled by more than one locus.

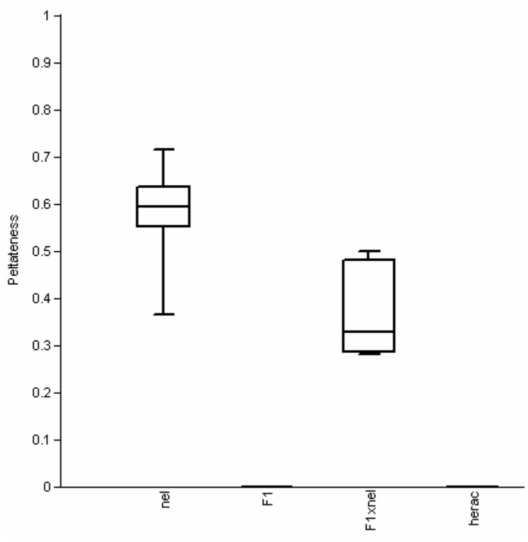


Figure 27 Graph showing the variability of peltateness in *B. nelumbiifolia*.
The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median. Peltateness variation among individuals of parent's *B. nelumbiifolia* and *B. heracleifolia*, F1 (CKB44) and Backcross progeny of 8 plants (ARB236).

3.6.3. Peltateness in *B. conchifolia* × *B. plebeja*

A cross of *B. plebeja* (non peltate) and *B. conchifolia* (peltate) generated an F1 progeny (CKB137) which was at first peltate but changed into non peltate during development. This F1 hybrid had blotches, a red eye and red margins on the adaxial side of the leaf like *B. plebeja.*

In 2007 we backcrossed CKB137.8 to *B. conchifolia* to generate a population of ARB 302. Because juvenile leaves can be variable in the F1 so I scored the mature leaf phenotypes of ARB 302. The ratio of peltate: non peltate leaves was not significant from 1:1 (Table 14), consistant with the involvement of a single gene/locus and a dominant non peltate allele from *B. plebeja*. The histogram of fraction of peltate leaves per F1BC1 (ARB 302) individual (Figure 28) shows a bimodal distribution suggesting that peltateness is mainly controlled by a single gene. The presence of variable proportions of peltate leaves on the third class of plants indicates that some other genetic effects are modifying the phenotype of peltateness. This could be either a number of less important loci, allelic variation or due to the epigenetic effect (Kalisz and Purugganan, 2004; Grant-Downton and Dickenson, 2006).

Figure 28 Fraction peltateness per F1BC1 individuals in ARB 302.

ARB 302 was generated with different F1 hybrids and can not be used for making a genetic map of *Begonia*. Therefore we generated another backcross population with *B. conchifolia* from a single F1 hybrid (ARB 312). ARB 312 generated 132 plants which were all peltate at maturity. I quantified peltateness in ARB 312 as described in section 3.3. The distributaion of peltateness in ARB 312 indicates the presence of more than one locus for controlling the trait (Figure 29, 30).

We also generated a backcross to *B. plebeja* and named it BOB 360. BOB 360 segregated into 274 non peltate and 22 peltate plants. If a single gene is controlling the peltate leaf form we expect 1:1 peltate to non peltate plants, if two genes are controlling the trait, we expect 3:1 segregation ratio and for 3 genes we expect 7:1. The fact that we got 12:1 segregation ratio for non peltate to peltate plants indicates the involvement of four or more loci for controlling the peltate leaf form. The ratio of non peltate to peltate leaves was not significant from 16:1 (Table 15) suggesting the involvement of four or more loci for controlling the peltate leaf form in this case.

Table 14 Number of peltate and non peltate individuals in ARB 302

	Peltate	**Non peltate**	**Total**
Observed	**83**	**75**	**158**
Expected	**79**	**79**	**158**
$(O-E)^2$	**16**	**16**	
$X^2=(O-E)^2/E$	**0.20**	**0.20**	**0.40**

The test has one degree of freedom and the chi-square value is 0.40. The two tailed P-value is 0.53.

Table 15 Number of peltate and non peltate individuals in BOB 360

	Peltate	**Non peltate**	**Total**
Observed	**22**	**274**	**296**
Expected	**18.5**	**277.5**	**296**
$(O-E)^2$	**12.25**	**12.25**	
$X^2=(O-E)^2/E$	**0.662**	**0.044**	**0.706**

The test has one degree of freedom and the chi-square value is 0.706. The two tailed P-value is 0.4007.

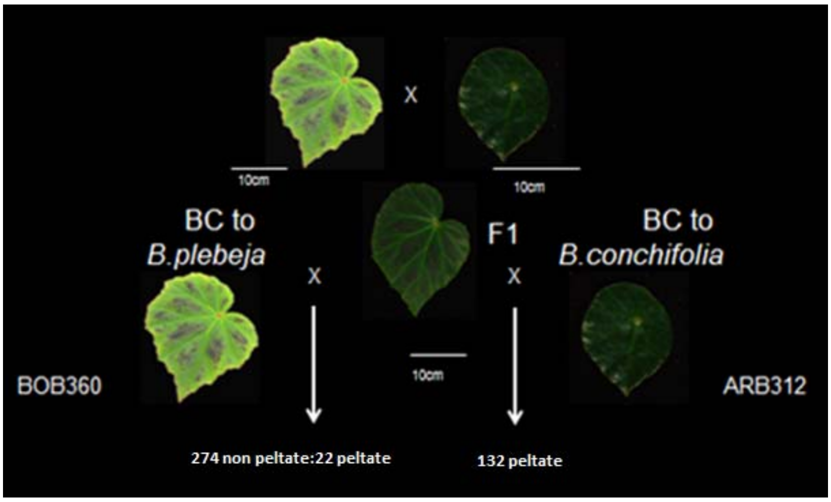


Figure 29 Genetics of peltateness in *B. conchifolia*.
Production of back crosses populations to analyze the peltateness in *B. plebeja* and *B. conchifolia*. (A) Cross between *B. plebeja* and *B. conchifolia* species resulted in non peltate F1 hybrid (CKB137), (B) Cross between CKB137.1 and *B. plebeja* resulted in 22/274 peltate plants, (C) Cross between CKB137.6 and *B. conchifolia* generated 129/132 peltate plants.

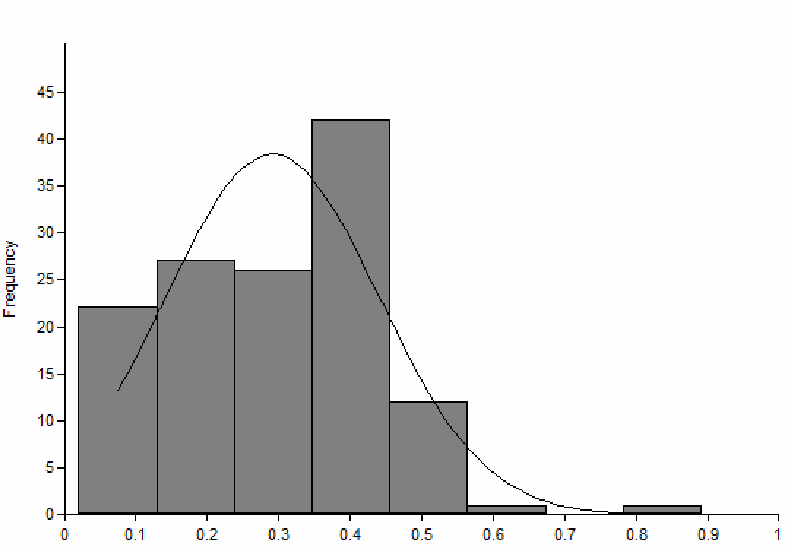


Figure 30 Graph showing the variability of peltateness in *B. conchifolia*.
Peltateness was measured as described in section 3.3. The peltateness shows a continuous variation and may be controlled by more than one loci.

3.6.4. Peltateness in *B. conchifolia* × *B. sericoneura*

F1 hybrids of *B. sericoneura* (non peltate) and *B. conchifolia* (peltate) were peltate (CKB197) (Figure 31). In a backcross to *B. conchifolia* all 72 plants were peltate. This is consistant with the previous findings where backcross of 137.1 with *B. conchifolia* yielded a progeny of 132 plants which were all peltate (section 3.6.3). The backcross with *B. sericoneura* yielded 11/83 peltate plants which suggests the involvement of 3 loci (1:7) for controlling the trait.

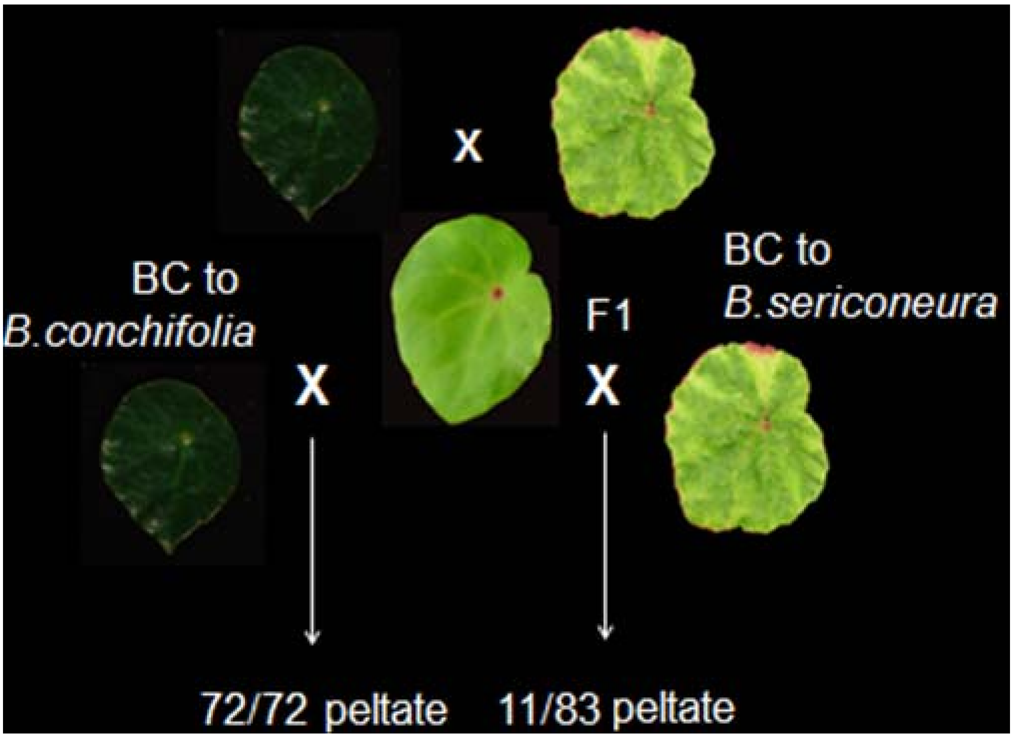


Figure 31 Genetics of peltateness in *B. conchifolia* × *B. sericoneura*.
(A) Cross between *B. sericoneura* and *B. conchifolia* species resulted in non peltate F1 hybrid (CKB197), (B) Cross between CKB 197.1 and *B. sericoneura* resulted ARB210, (C) Cross between CKB197.2 and *B. conchifolia* resulted ARB 552.

3.7. Genetics of compound leaf form

I calculated the dissection index for *B. heracleifolia*, *B. carolineifolia* and B. *theimei* by dividing the leaf perimeter by the square root of leaf area. The leaves of each individual plant were photographed while still attached and backgrounds were changed in Adobe Photoshop to run these leaves though Lamina software (Bylesjo *et al*., 2008) to calculate the square root of leaf area and leaf perimeter. There is no clear trend in variation of dissection index along the rhizome of a plant. The dissection index varies between clones of the same species kept in different greenhouses (Figure 32). Analysis of variance (ANOVA) was carried out in PAST and heritability calculated by dividing between groups variation (between different species) with the sum of between groups variation and within groups variation (between different clones of same species). The heritability of dissection index for *B. heracleifolia* is 60% and for *B. carolineifolia* is 74%. The dissection phenotype of F1s of compound x simple leaved species is usually midway between the phenotypes of the parents. This suggests either a simple genetics with semidominant effects or many loci with additive effects (Figure 33 & 34B). When two compound leaved plants *B. thiemei* and *B. carolineifolia* were crossed, the F1 is compound leaved which suggests the non complementation for compoundness (Figure 34B). This indicates that same loci may be controlling the compoundness in each case.

3.7.1. Compoundness in *B. heracleifolia* × *B. carolineifolia*

The dissection index of F1 hybrid of *B. heracleifolia* and *B. carolineifolia* was midway between the phenotypes of the parents suggesting that compoundness is semi dominant. F1BC1 progeny of *(B. heracleifolia* and *B. carolineifolia*) with *B. heracleifolia* exhibit a wide variety of degrees of lobiness, including many which are less dissected than either parent (Figure 35). This suggests leaf dissection is controlled by multiple loci, and not strictly additive between the two parents.

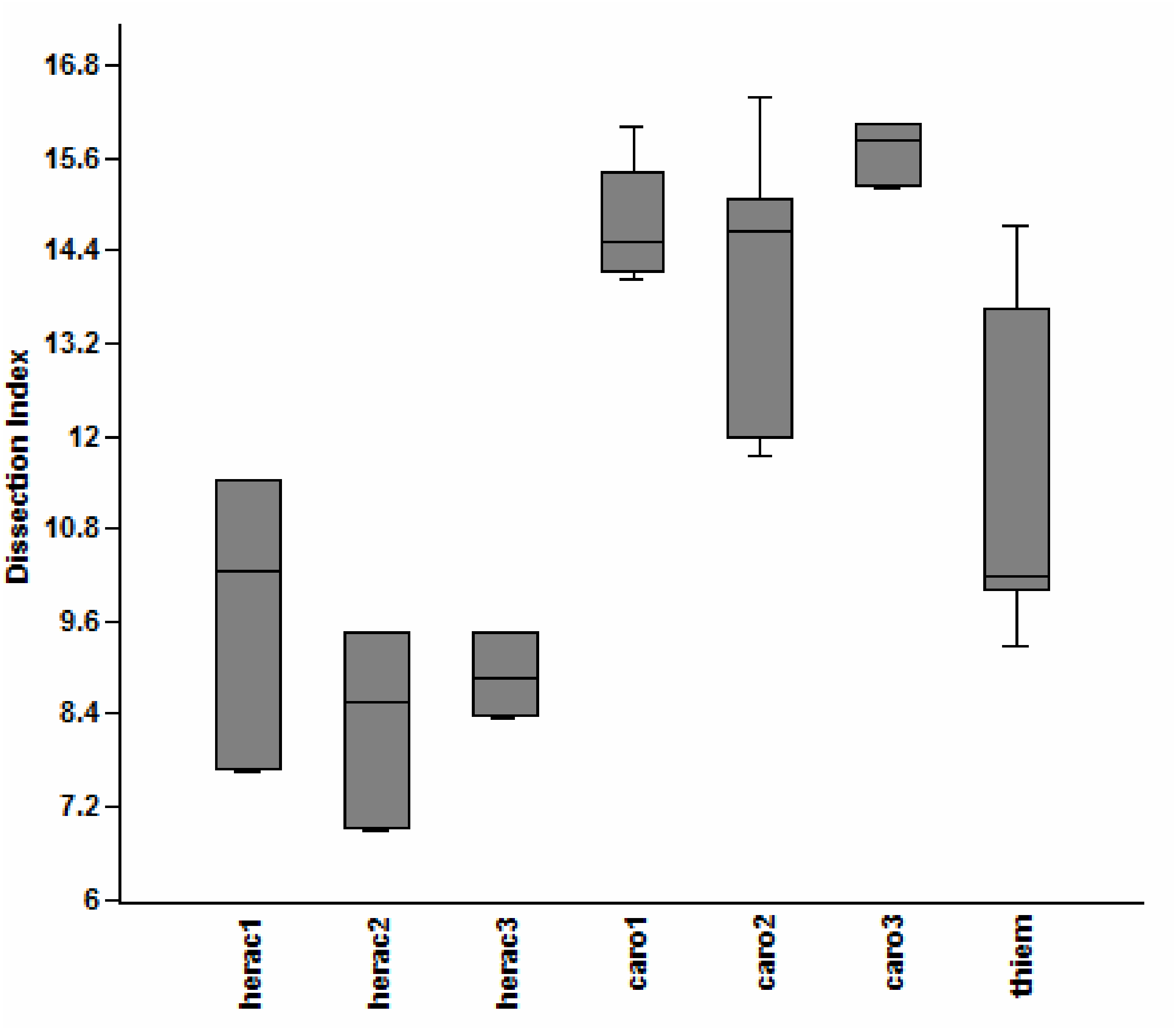


Figure 32 Differences in dissection index of *B. heracleifolia*, *B. carolineifolia* and *B. theimei*.
The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median. The plants of same species are clones kept in different green houses and the difference seen in disection index is due to the environment only, however there is obvious variation in disection index between different species. Abbreviations are as herac- *B. heracleifolia*, caro- *B. carolineifolia*, thiem- *B. theimei*.

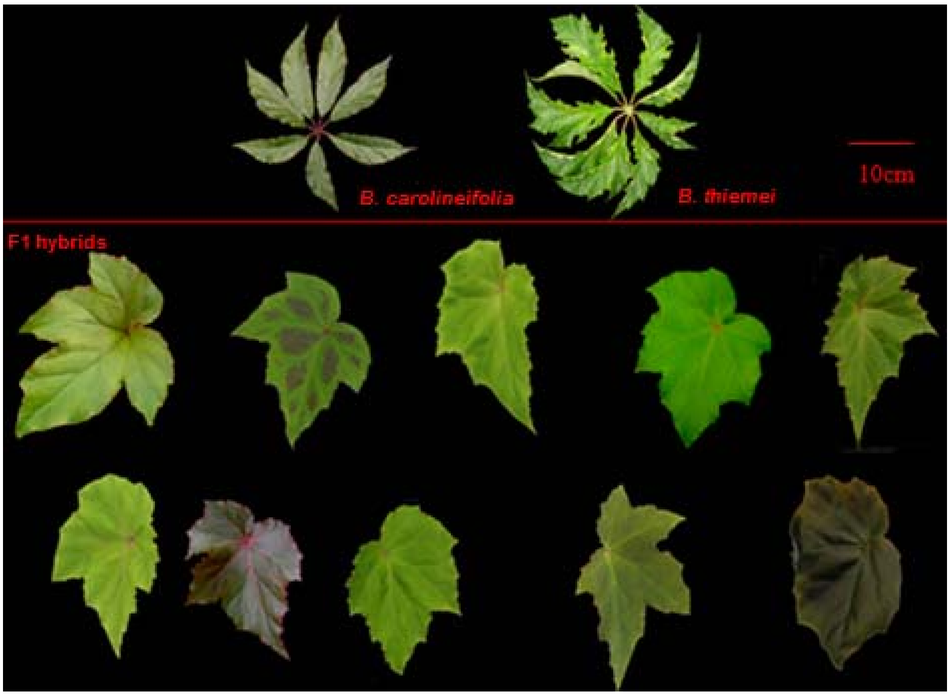


Figure 33 Cross between compound leaved *Begonia* section Gireoudia species.
F1 hybrids of *B. thiemei* and *B. carolineifolia* with simple leaved species showed some degree of dissection index. This suggests the semi dominance of compoundness in this section

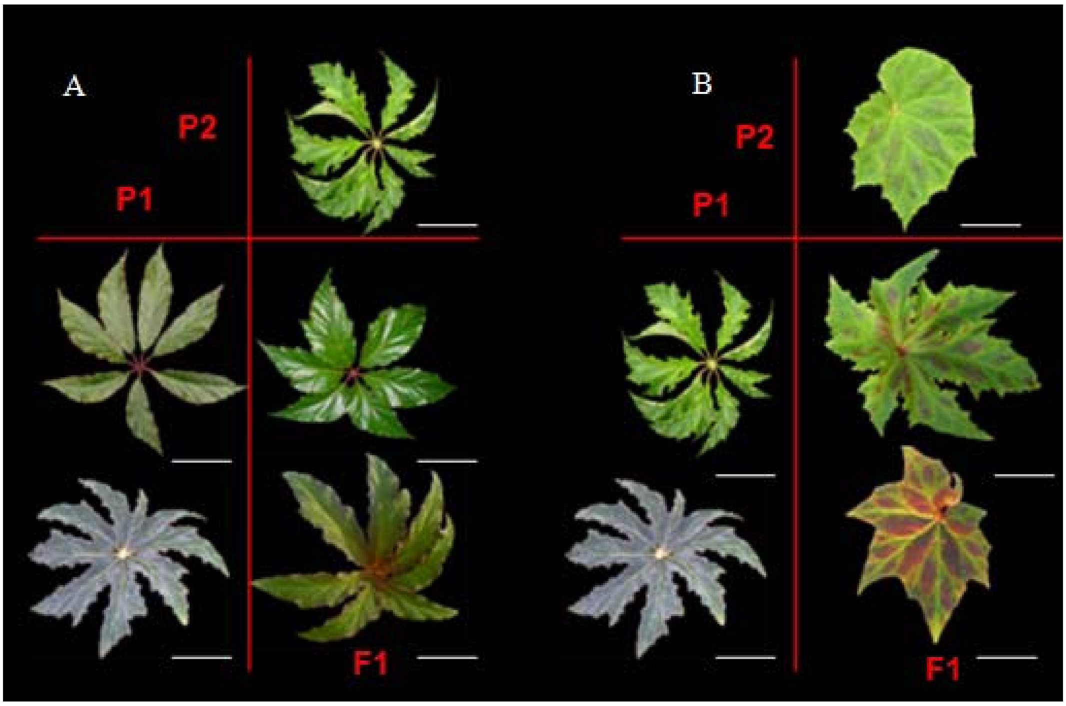


Figure 34 Crosses between compound leaved *Begonia* section Gireoudia species.
***B. thiemei* and *B. carolineifolia* yielded some degree of dissection index in otherwise simple leaved species indicating the semi dominance of compoundness in this section. Scale is 10 cm.**

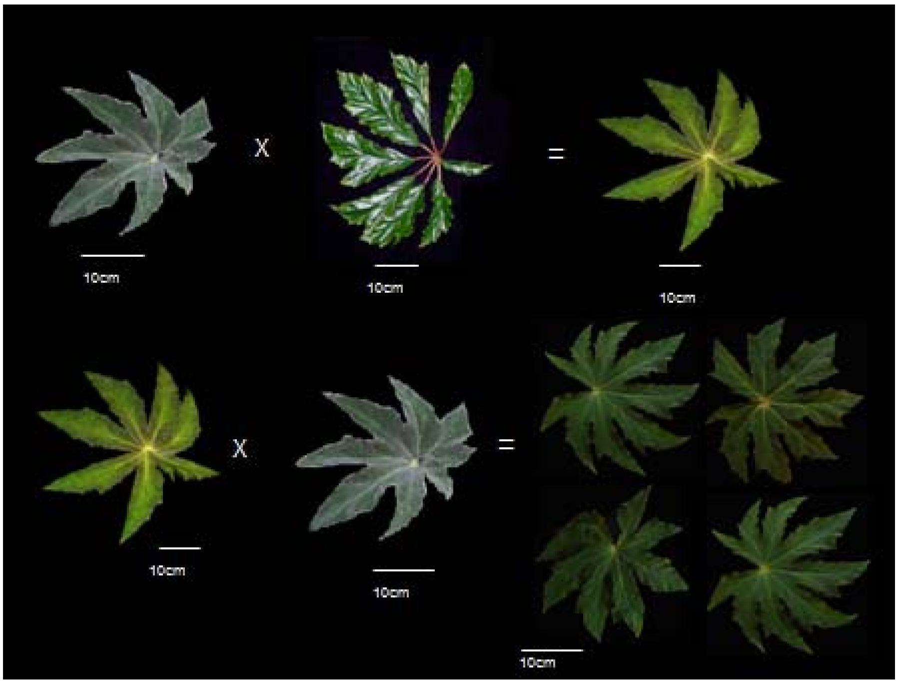


Figure 35 Genetics of compoundness in *B. heracleifolia.*
Cross between *B. heracleifolia* and *B. carolineifolia* resulted in CKB66 and backcross of this F1 hybrid with *B. heracleifolia* yielded ARB 444.

3.8. Discussion

In almost every flora, related species often have leaves that differ markedly in degree of lobing, edge dissection, length/width ratio, symmetry, or combinations of these characters. The leaf form variation exists among leaves within a single individual plant (Ashby 1948; Allsopp 1965), among individuals within populations (Hilu 1983) and between different species (Wyatt and Antonovics 1981; Gurevitch 1988; Andersson 1991; Oyama 1996; Harris *et al.*, 1998). Similarly variation in leaf form exists between leaves of the same plant, within individuals of same species and between different species of *Begonia* section Gireoudia.

Among different species of *Begonia* section Gireoudia the leaf primordia differ in shape and size soon after initiation and peltateness develops with the age of plant. The petioles in all *Begonia* section Gireoudia species are abaxialized which is linked to the development of lamina all around the petiole attachment point. This supports the theory that lamina develops at the boundary between abaxial and adaxial domains and suggests that section Gireoudia is basically peltate, the species varying only in the extent to which the lamina grows out. This is supported by the tendency of *Begonia* leaves to evolve peltateness very frequently. When two different species are crossed to make an F1, heterozygosity in F1 and backcross plants should be high and may encompass many variable genes that have an effect of morphology. In maize the phenotypes of *knox* (*KNOTTED* loss of function) mutants are particularly background dependent (Vollbrecht *et al.*, 2000). Similarly peltateness in *Begonia* seems to be dependent on background and behaved differently in different crosses. This could be due to pleiotropy which means a single gene or allele controls the development of more than one trait (Flatt, 2005). For example the flowering time locus *FLC* controls vernalization response, water use efficiency, circadian leaf movements, and nitrogen content (Swarrup *et al.*, 1999; Sheldon *et al.*, 2000; McKay *et al.*, 2003; Loudet *et al.*, 2003). *HD-ZIPIII* genes have multiple roles like leaf polarity, meristem function and vascular patterning (Prigge *et al.*, 2005). The *AS* pathway regulates the down regulation of *KNOX* genes in the leaf as well as confers disease resistance (Nurmberg *et al.*, 2007). When a certain trait is being selectively maintained, the functional pleiotropic allele can constrain the evolvability of other traits.

In many species compoundness is controlled by few loci (Durst, 1929; Whitaker, 1944; Imrie and Knowely 1970; Shull, 1909; Elmore, 1986; Andersson, 1999; Peter and Prins, 2008; Correns, 1928; Whaley and Whaley, 1942; Clausen and Hiesy, 1958 and Kimura *et al.*, 2008). If a single gene is responsible for the compoundness in *Begonia* we expect two or three classes of dissection index but the fact that there is continuous distribution of leaf dissection in the hybrid offspring suggests compoundness is controlled by many loci in *Begonia*, in constrast to what has been shown using transgenics in *Solanum* and in Brassicacea (Hay and Tsiantis, 2006; Jassen *et al.*, 1998; Bharathan *et al.*, 2002).

3.9. Conclusion

Differences in leaf form become apparent as early as P1 stage and peltateness increases with the age of leaf. Both peltateness and compoundness evolved multiple times

in genus *Begonia* and are polygenic traits in *Begonia* section Gireoudia. Peltateness showed no complementation in interspecific hybrids between peltate species which suggests that this trait has evolved due to the changes at the same locus (peltateness has evolved five times in this section), but modifiers (may be epigenetic factors causing phenotypic variation without changing either the coding sequence of a gene or the upstream promoters region) are present in non peltate species which promote peltateness in backcross populations. Association mapping studies in these back cross populations can reveal if *ARP* or *KNOX* genes are major loci controlling these traits. The genetic analysis suggests that more than one locus affect each of these traits so QTL analysis would be the best method to establish the genetic architecture of these traits in more detail.

Chapter 4

Cloning and molecular characterization of *ARP* and *KNOX* genes in *Begonia* section Gireoudia

4.1. Introduction

Molecular evolutionary analysis provides a method to investigate the history of a particular locus. One way of detecting the signature of selection is by calculating the ratio of non synonymous (change occurs in amino acid sequence) to synonymous (no change in amino acid sequence) substitutions. If there is no selection, non synonymous (dN) and synonymous (dS) substitutions should occur at the same rate ($dN/dS = 1$). If there is negative selection, $dN/dS < 1$ and if there is positive selection, $dN/dS > 1$ (Nielsen, 2005). Negative selection ($dN<dS$) means any change in that sequence may be deleterious and positive selection ($dN>dS$) suggests that there has been selection for mutations. Purifying selection is identical to negative selection in that it describes selection against new mutations. Negative or purifying selection can indicate the regions of functional importance while positive selection is associated with the evolution of a new function (Nielsen *et al*., 2005).

The strength of selection varies between genes and between different domains and residues in a protein. A vital binding site will have experienced stronger purifying selection pressure than a purely structural region. Analysis of the rate of evolution in different regions of a gene can suggest which protein domain is key for its function and which are evolving new functions.

Gene or genome duplication is a major source of evolutionary novelty as it contributes to the evolution of new gene functions (Ohno, 1970; Gu *et al*., 2003; Long *et al*., 1993). The generation of two copies of a gene removes selective pressure from one which can than respond either of three ways: neofunctionalization or through escape from adaptive conflict (EAC) or it can be lost. In neofunctionalization one copy maintains the ancestral function while the second copy is selected to perform a new function. Under escape from adaptive conflict a novel function arises first in the single ancestral gene which reduces the ability of that gene to perform its original function. After duplication either paralog is free to perform either original or novel function or no function (Soltis *et al.*, 2009). Most new genes do not acquire new functions but undergo the process of subfunctionalization in which after duplication the two gene copies specialize to perform complementary functions (Prince and Pickett, 2002; Piatigorsky, 2007).

A large fraction of most plant genomes is composed of duplicated loci due to the frequent occurrence of genomic segmental duplications and polyploidization. For example, up to 90% and 62% of loci are duplicated in the *Arabidopsis thaliana* and rice genomes, respectively, and it is estimated that 70–80% of angiosperm species have undergone polyploidization at some point in their evolutionary history (Simillion *et al.*, 2002; Bowers *et al.*, 2003; Paterson *et al.*, 2004). Therefore gene duplication may have been an important force in the generation of angiosperm diversity.

Genes expression patterns are highly variable within and between species (Oleksiak *et al.*, 2002; Khaitovich *et al.*, 2004; Kliebenstein *et al.*, 2006), and a phenotype can be affected by a slight change in gene expression (Gompel *et al.*, 2005; Kimura *et al.*, 2008; Hay and Tsiantis, 2006). Some examples of variation in regulatory loci conferring adaptive phenotypes are: Class 1 *KNOX* gene promoters affecting diversity in leaf form between *Arabidopsis thaliana* and its closest relative *Cardamine hirsuta* (Hay and Tsiantis, 2006); *PTS* promoter responsible for the complexity of compoundness in *Solanum galapagense* in comparison to its less dissected sister species *Solanum cheesmaniae* (Kimura *et al.*, 2008); the *LdhB* promoter of *Fundulus heteroclitus* which played a role in local adaptation of this fish to its habitat temperature by affecting gene regulation (Schutle *et al.,* 2001; Schutle *et al.,* 1997; Crawford *et al.*, 1999; Schutle *et al.,* 2000), the *hsp70Ba* promoter of *Drosophila melanogaster* which regulates the natural variation in the stress inducible molecular chaperone Hsp70 (Lerman *et al.*, 2003; Michalak *et al.*, 2001); a point mutation in the chalcone synthase promoter was linked with functional cis regulatory variation of F1 individuals to light response in *Arabidopsis thaliana* (de Meaux *et al*., 2003); insertion of a transposable element in the *Cyp6g1* promoter of *Drosophila melanogaster* causes the over transcription of *Cyp6g1* confering insecticide resistance (Daborn *et al*., 2002). Multiple instances of adaptive evolution at cis regulatory loci have been observed in humans (Rockman *et al*., 2003; Rockman *et al*., 2004; Hahn *et al*., 2004).

Numerous segments of conserved non coding sequences between different species have been found through whole genome sequence comparisons (Dermitzakis *et al*., 2004). Identification of cis-regulatory evolution is mostly based on patterns of DNA conservation and it is useful to identify some functionally important elements in promoter sequences (Uchida *et al*., 2007; Koch *et al*., 2001; Cliften *et al*., 2003,).

ARP gene function has been shown to cause peltate leaf formation in a range of species and they regulate the number and placement of leaflets in compound leaved-species (Kim *et al*., 2003a; Kim *et al*., 2003b). *KNOX* genes have been identified as regulating natural variation in leaf form in two independent cases differential expression of class 1 *KNOX* genes is correlated with the variation in leaf form between *Arabidopsis thaliana* and its closest relative *Cardamine hirsuta* (Hay and Tsiantis, 2006); up regulation of *PTS* genes confers the complexity of compoundness in *Solanum galapagense* in comparison to its less dissected sister species *Solanum cheesmaniae* (Kimura *et al*., 2008).

In order to determine the role that *ARP* and *KNOX* genes play in species level variation in leaf form in *Begonia* I have cloned and sequenced *ARP* genes and Class 1 *KNOX* genes from *Begonia* section Gireoudia species. Sequence analysis of the CDS of *BARP* (*ARP* orthologue in *Begonia*) and *KNBs* (*STM* orthologue in *Begonia*) will reveal if there is evidence of selection on these genes and will show whether changes in the CDS of either *BARP* or *KNBs* are correlated with changes in leaf form. The promoters of *BARP* and *KNB* genes have been sequenced to isolate the elements common to *ARP* and *KNOX* promoters in other species.

4.2. Amplification of *ARP* genes from *Begonia* section Gireoudia

At the outset of my research a 300 base pair fragment of *BARP1* (*Begonia ARP* gene) had been amplified from genomic DNA by Catherine Kidner using degenerate primers designed by Jill Harrison (Harrison *et al.*, 2005). Piere Cattenoz subsequently obtained single reads of partial CDS (coding sequences) for *BARP1* from four *Begonia* species (*B. nelumbiifolia*, *B. peltata*, *B. lindlyana* and *B. conchifolia*) (MSc thesis, Lille. 2006).

The first step of my research was to obtain multiple reads of the complete genomic sequence for *BARP* from 14 *Begonia* species. Several rounds of TAIL PCRs were performed on genomic DNA to get the complete CDS, 3′UTR and 5′ region of *BARP1* genes. It was hard to make cDNA from *Begonia* due to the difficulty in RNA isolation. It may be due to the presence of the large amounts of oxalic acid and/or carbohydrates in *Begonias* (Kopperund and Einset, 1995).

PCR products were cloned into pGEM T easy vectors and sequenced. The number of plasmids sequenced for each *BARP* gene and the number of *BARP* genes found in different species is given in table 16. Short stretches of sequences were assembled into contigs in Geneious to produce a full length sequence for each species.

The sequences were blasted at the NCBI non-redundant database and they show strong similarity to *ARPs* from Apple, *Medicago truncatula, Nicotiana tobacum, Glycine max, Arabidopsis thaliana* and many other species. The composition of each *BARP1* gene amplified from *Begonia* is summarised in table 17.

Table 16 Sequencing depth for different species

Species	**Section**	**No. of plasmids sequenced**	**Direct PCR fragments sequenced**	**Complete gene**	**No of *BARP* genes**
B. thiemei	Gireoudia	7	Approximately 5	Yes	1
B. carolineifolia	Gireoudia	16	Approximately 5	Yes	1
B. heracleifolia	Gireoudia	10	Approximately 5	Yes	1
B. nelumbiifolia	Gireoudia	7	Approximately 5	Yes	1
B. conchifolia	Gireoudia	10	Approximately 5	Yes	1
B. peltata	Gireoudia	7	Approximately 5	Yes	1
B. kellemanii	Gireoudia	3	Approximately 5	Yes	1
B. sarcophyla	Gireoudia	1	None	no	1
B. plebeja	Gireoudia	3	Approximately 5	Yes	1
B. sericoneura	Gireoudia	2	Approximately 5	Yes	1
B. multinervia	Gireoudia	1	Approximately 3	No	1
B. cardiocarpa	Gireoudia	3	Approximately 5	Yes	1
B. stigmosa	Gireoudia	2	Approximately 3	no	2
B. lymansmithi	Gireoudia	1	None	no	1
B. luxurians	Scheidweileria	3	Approximately 12	Yes	2

Table 17 *BARP* genes statistics

***BARP* genes statistics**						
Species	**Total (bp)**	**Promoter (bp)**	**Exon1 (bp)**	**Exon2 (bp)**	**Intron (bp)**	**3UTR (bp)**
B. thiemei BARP1	2719	1124	153	1077	365	_
B. carolineifolia BARP1	1993	242	153	1077	364	149
B. heracleifolia BARP1	2225	633	153	1074	357	_
B. nelumbiifolia BARP1	2197	339	153	1077	369	251
B. conchifolia BARP1	2133	437	153	1077	364	94
B. peltata BARP1	2084	443	153	1077	364	45
B. kellemanii BARP1	1993	346	153	1077	364	45
B. sarcophyla BARP1	669	_	_	669	_	_
B. plebeja BARP1	1951	330	150	1077	364	22
B. sericoneura BARP1	2014	336	153	1077	365	75
B. multinervia BARP1	1086	_	_	958	_	128
B. cardiocarpa BARP1	1763	130	153	1077	371	24
B. stigmosa BARP1	1166	_	_	1053	_	113
B. stigmosa BARP2	256	_	_	256		_
B. lymansmithi BARP1	549	_	_	549	_	_
B. luxurians BARP1	1358	_	_	1077	281	_
B. luxurians BARP2	935	_	_	935	_	_

4.3. Duplication of *BARP* genes in *Begonia*

In order to see if there are more than one copy of *BARP* genes in *Begonia* I used Southern hybridization according to the protocol described in chapter 2. The conserved C-terminal domain of *BARP* was used as a probe (Figure 36) because this domain is conserved amongst *ARP* genes but not amongst the large MYB gene family. The genomic DNA of *B. nelumbiifolia* and *B. sericoneura* digested with *Dra*1 enzyme were used as templates for Southern hybridization. The results suggested the presence of two copies of *BARP* genes in *Begonia* (Figure 37).

4.4. Amplification of *BARP2* from *Begonia*

After the confirmation of a *BARP* duplication in *Begonia* through Southern hybridization the next step was to amplify the second copy of *BARP*. I tried two different approaches to amplify the second copy initially from four *Begonia* species (*B. nelumbiifolia*, *B. sericoneura, B. theimei* and *B. luxurians*).

Firstly I used the degenerate primers designed by Jill Harrison (Harrison *et al.*, 2005) on four different species followed by sequencing of four plasmids from each species. Products of the expected sizes (305 bp) were produced and cloned into pGEM T plasmids. I sequenced four plasmids from each species which all coded for the *BARP* gene already characterized (*BARP1*). If both genes are equally likely to amplify with *ARP* degenerate primers (Harrison *et al*., 2005) the chances of finding the second *BARP* gene in 16 plasmids is greater than 99.9%.

Secondly I did tail PCRs using genomic DNAs as templates with primers designed on the conserved Myb domain1 (named Myb domain T1, Myb domain T2 and Myb domain T3) of *Begonia*, *Antirrhinum majus*, *Arabidopsis thaliana*, *Zea mays*, *Nicotiana tabacum*, *Pisum sativum* and *Glycine max*. Three rounds of tail PCR were performed on *B. nelumbiifolia*, *B. sericoneura, B. theimei* and *B. luxurians* as described in chapter 2 which yielded *BARP2* from *B. luxurians* only in four independent PCR products. This second copy of *BARP* from *B. luxurians* has an 18bp deletion in Myb domain 2 of *BARP* genes (Figure 38). A single partial read from *B. stigmosa* also had an 18 bp deletion in the Myb domain 2. The position of *BARP2* with respect to *BARP1* genes is shown in a phylogenetic tree in Figure 39 rooted with *Cucumber ARP* genes.

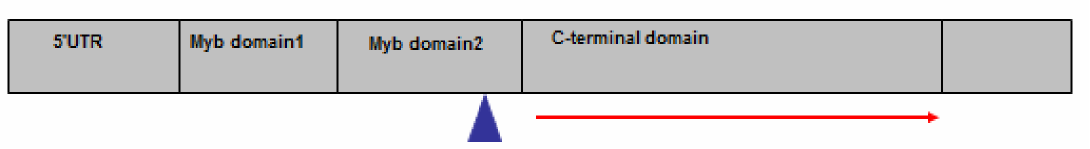


Figure 36 Schematic map of *BARP1* genes showing the location of probe used in southern hybridization.
Blue triangle indicates the position of *Dra*1 and red arrow represents the region used as a probe.

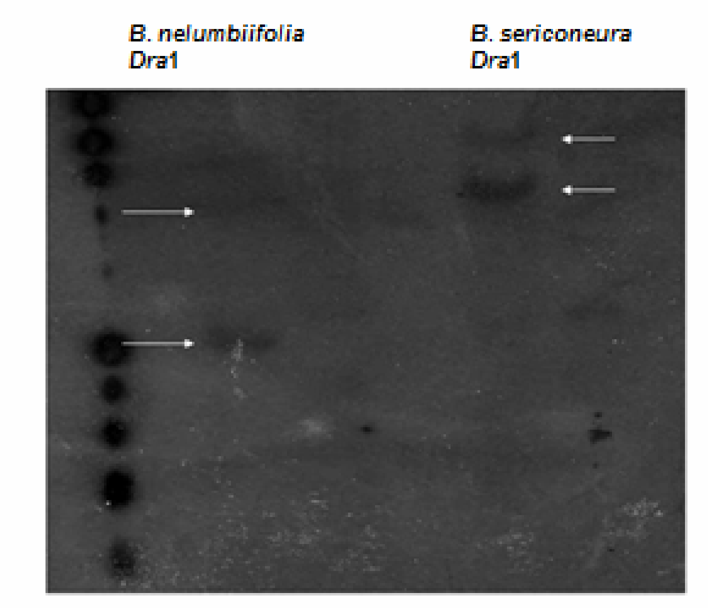
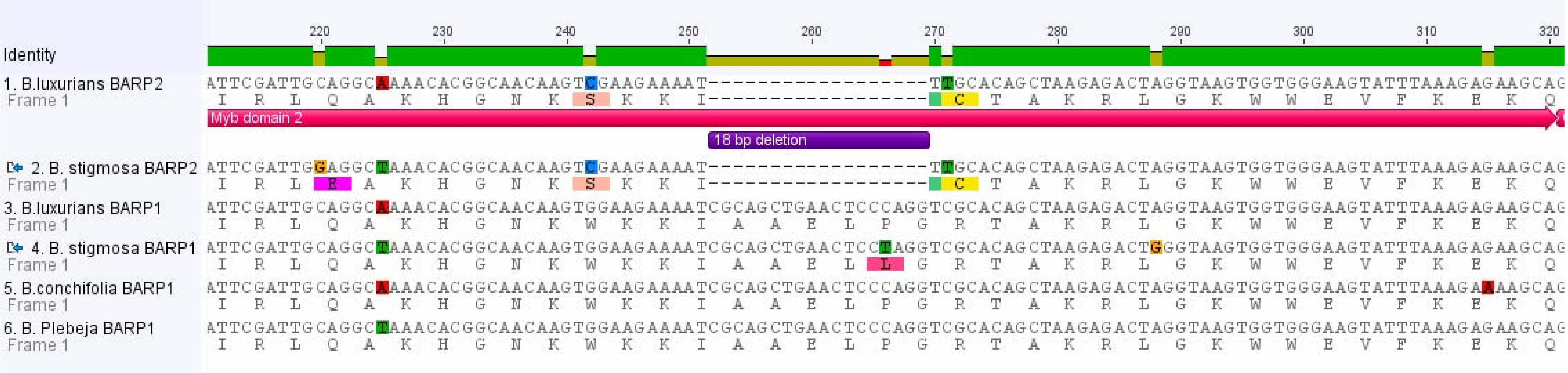


Figure 37 Southern hybridization results.
X-Ray photograph showing two copies of *BARP* genes in *B. nelumbiifolia* and *B. sericoneura.*



```
1. B.luxurians BARP2    ATTCGATTGCAGGCAAAACACGGCAACAAGTCGAAGAAAAT------------------TTGCACAGCTAAGAGACTAGGTAAGTGGTGGGAAGTATTTAAAGAGAAGCAG
   Frame 1               I  R  L  Q  A  K  H  G  N  K  S  K  K  I------------------  C  T  A  K  R  L  G  K  W  W  E  V  F  K  E  K  Q
2. B. stigmosa BARP2    ATTCGATTGGAGGCTAAACACGGCAACAAGTCGAAGAAAAT------------------TTGCACAGCTAAGAGACTAGGTAAGTGGTGGGAAGTATTTAAAGAGAAGCAG
   Frame 1               I  R  L  E  A  K  H  G  N  K  S  K  K  I------------------  C  T  A  K  R  L  G  K  W  W  E  V  F  K  E  K  Q
3. B.luxurians BARP1    ATTCGATTGCAGGCAAAACACGGCAACAAGTGGAAGAAAATCGCAGCTGAACTCCCAGGTCGCACAGCTAAGAGACTAGGTAAGTGGTGGGAAGTATTTAAAGAGAAGCAG
   Frame 1               I  R  L  Q  A  K  H  G  N  K  W  K  K  I  A  A  E  L  P  G  R  T  A  K  R  L  G  K  W  W  E  V  F  K  E  K  Q
4. B. stigmosa BARP1    ATTCGATTGCAGGCTAAACACGGCAACAAGTGGAAGAAAATCGCAGCTGAACTCCTAGGTCGCACAGCTAAGAGACTGGGTAAGTGGTGGGAAGTATTTAAAGAGAAGCAG
   Frame 1               I  R  L  Q  A  K  H  G  N  K  W  K  K  I  A  A  E  L  L  G  R  T  A  K  R  L  G  K  W  W  E  V  F  K  E  K  Q
5. B.conchifolia BARP1  ATTCGATTGCAGGCAAAACACGGCAACAAGTGGAAGAAAATCGCAGCTGAACTCCCAGGTCGCACAGCTAAGAGACTAGGTAAGTGGTGGGAAGTATTTAAAGAAAAGCAG
   Frame 1               I  R  L  Q  A  K  H  G  N  K  W  K  K  I  A  A  E  L  P  G  R  T  A  K  R  L  G  K  W  W  E  V  F  K  E  K  Q
6. B. Plebeja BARP1     ATTCGATTGCAGGCTAAACACGGCAACAAGTGGAAGAAAATCGCAGCTGAACTCCCAGGTCGCACAGCTAAGAGACTAGGTAAGTGGTGGGAAGTATTTAAAGAGAAGCAG
   Frame 1               I  R  L  Q  A  K  H  G  N  K  W  K  K  I  A  A  E  L  P  G  R  T  A  K  R  L  G  K  W  W  E  V  F  K  E  K  Q
```

Figure 38 Alignment of *BARP* genes.
The deletion of 18bp nucleotides (6 amino acids) in Myb domain 2 of *B. stigmosa BARP2* and *B. luxurians BARP2* is highlighted with the purple bar showing the 18bp deletion.

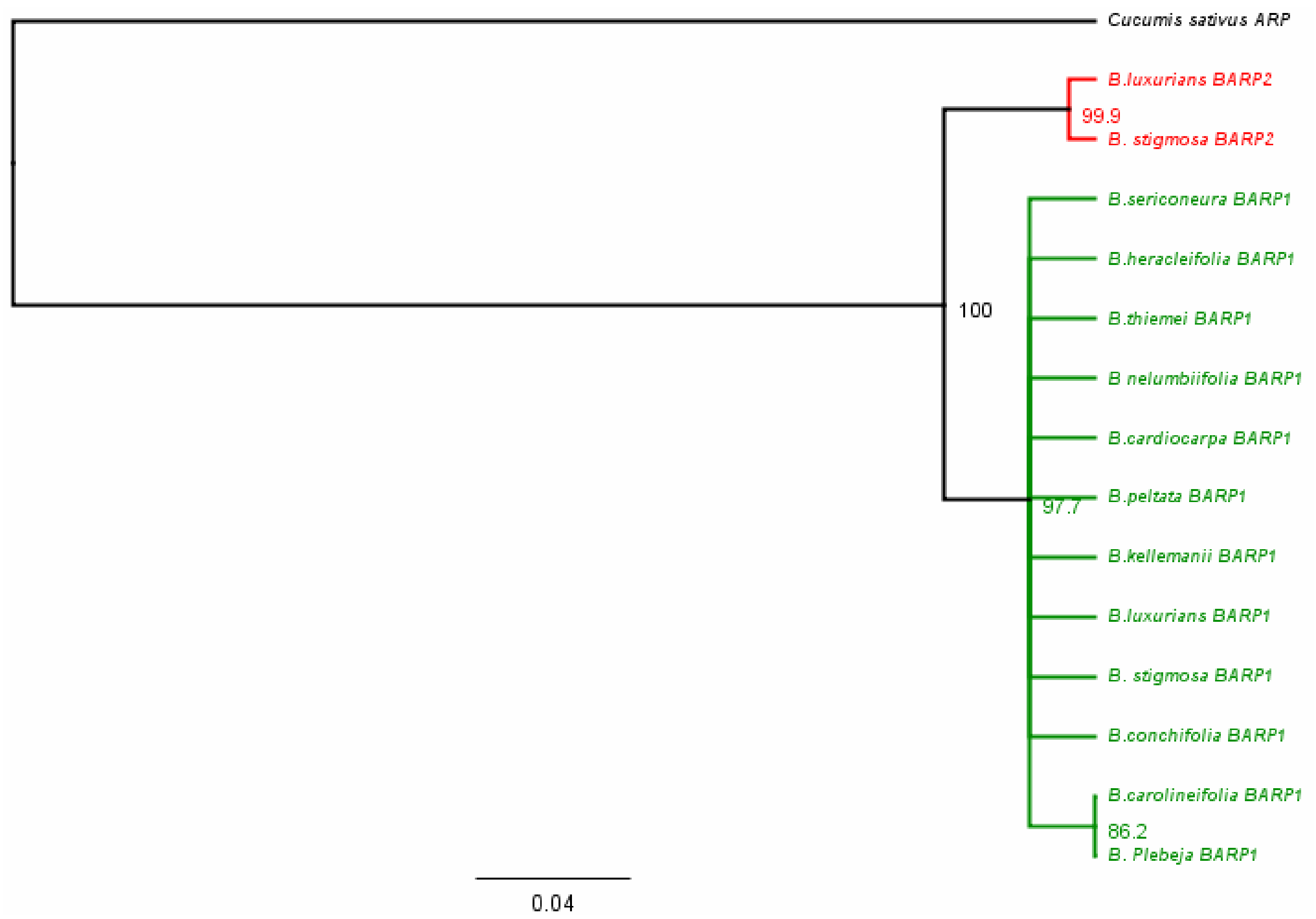


Figure 39 Phylogenetic relationships of *ARP* genes in *Begonia*.
Jukes-Cantor UPGMA tree with 1000 bootstrap replicates based on *ARP* CDS using *Cucumis sativus ARP* as an out group.

4.5. Annotation of *BARP1* genes and features of *BARP1* intron

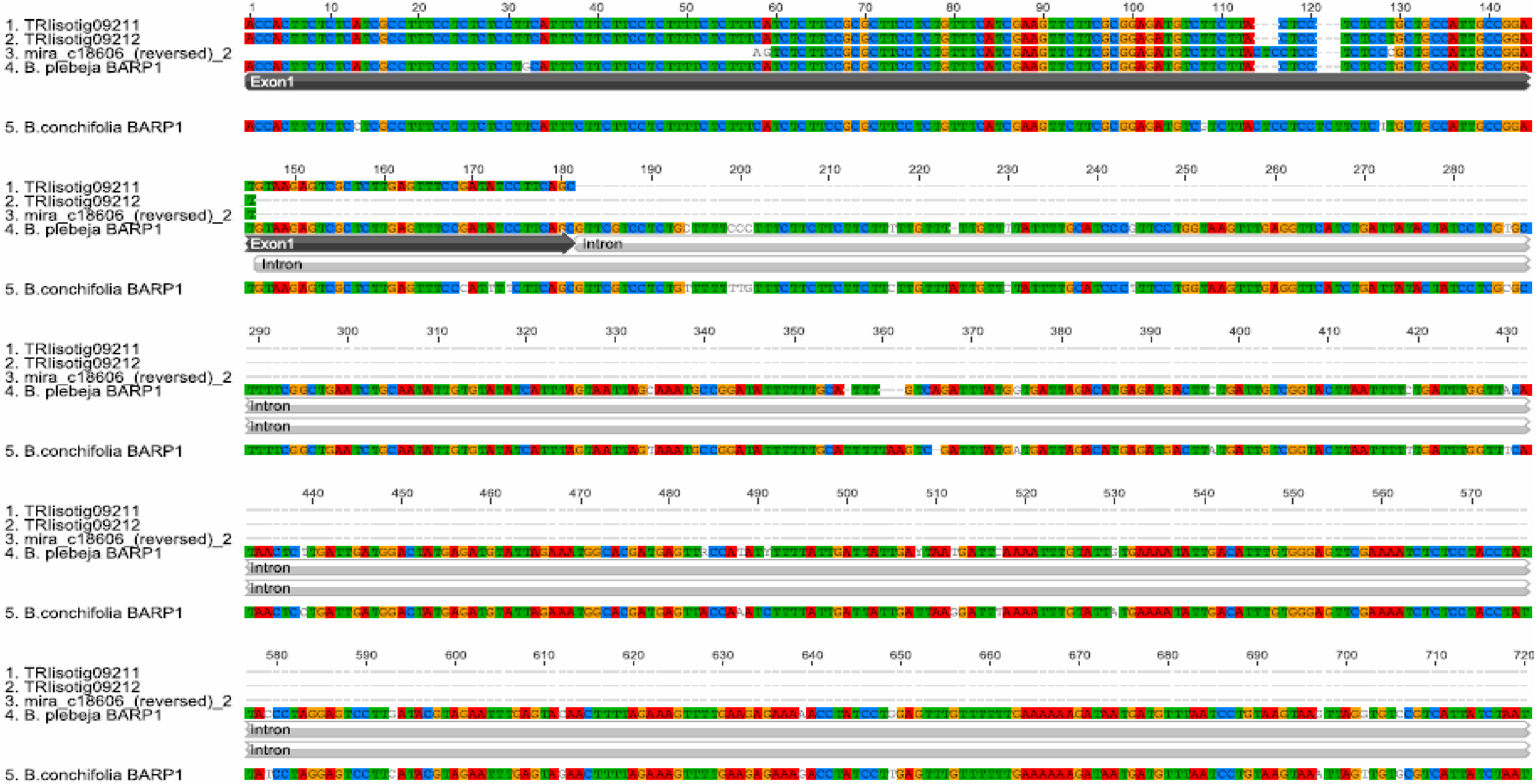
The RNA was isolated from *plebeja, B. conchifolia* and *B. venusta* and submitted for 454 sequencing to the Edinburgh University GenePool services. The *BARP1* genes were annotated for exons and intron on the basis of partial cDNA reads of transcriptome sequences from *B. plebeja, B. conchifolia* and *B. venusta* (Figure 40). More than one transcript matched the *B. conchifolia BARP1* but all except one resembled non-*ARP* Myb like genes. Thus transcriptomes sequences of *B. plebeja, B. conchifolia* and *B. venusta* have only one *BARP* gene.

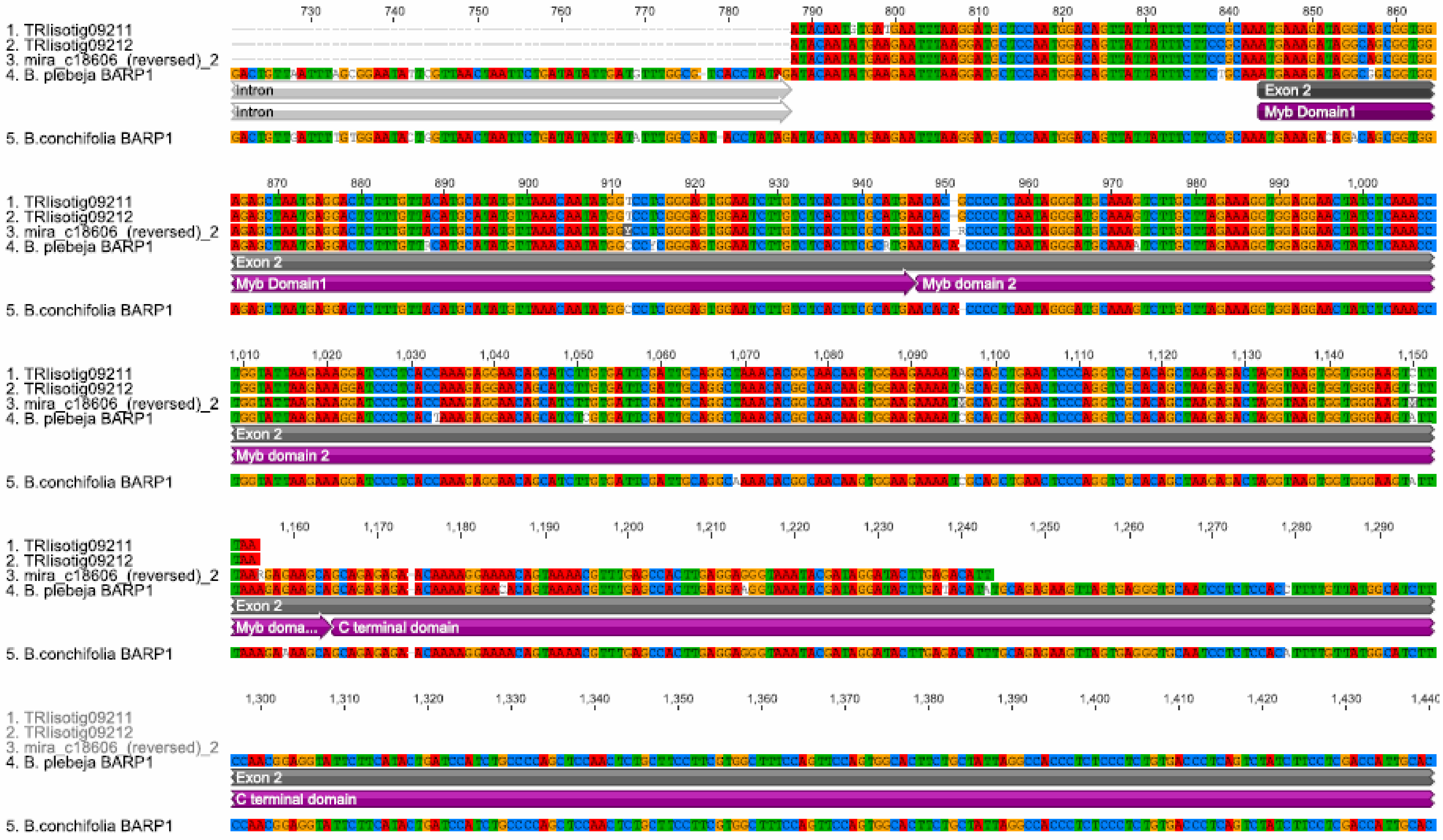
Introns are non-coding DNA regions which are transcribed to pre-mRNA and subsequently removed by a process called splicing during the processing to mature RNA. The sequences at the intron/exon junctions called splice sites are required for the removal of the introns from a pre-mRNA. A GT dinucleotide is commonly found at the 5′ end of the intron which is usually less conserved. The 3′ end of the introns has three conserved features: the branch point, followed by an AT rich tract, followed by a terminal AG at the extreme 3′ end (Lynch and Richardson, 2001; Black, 2003). These conserved features of introns exist in the *BARP* 5′ intron (Figure 40).

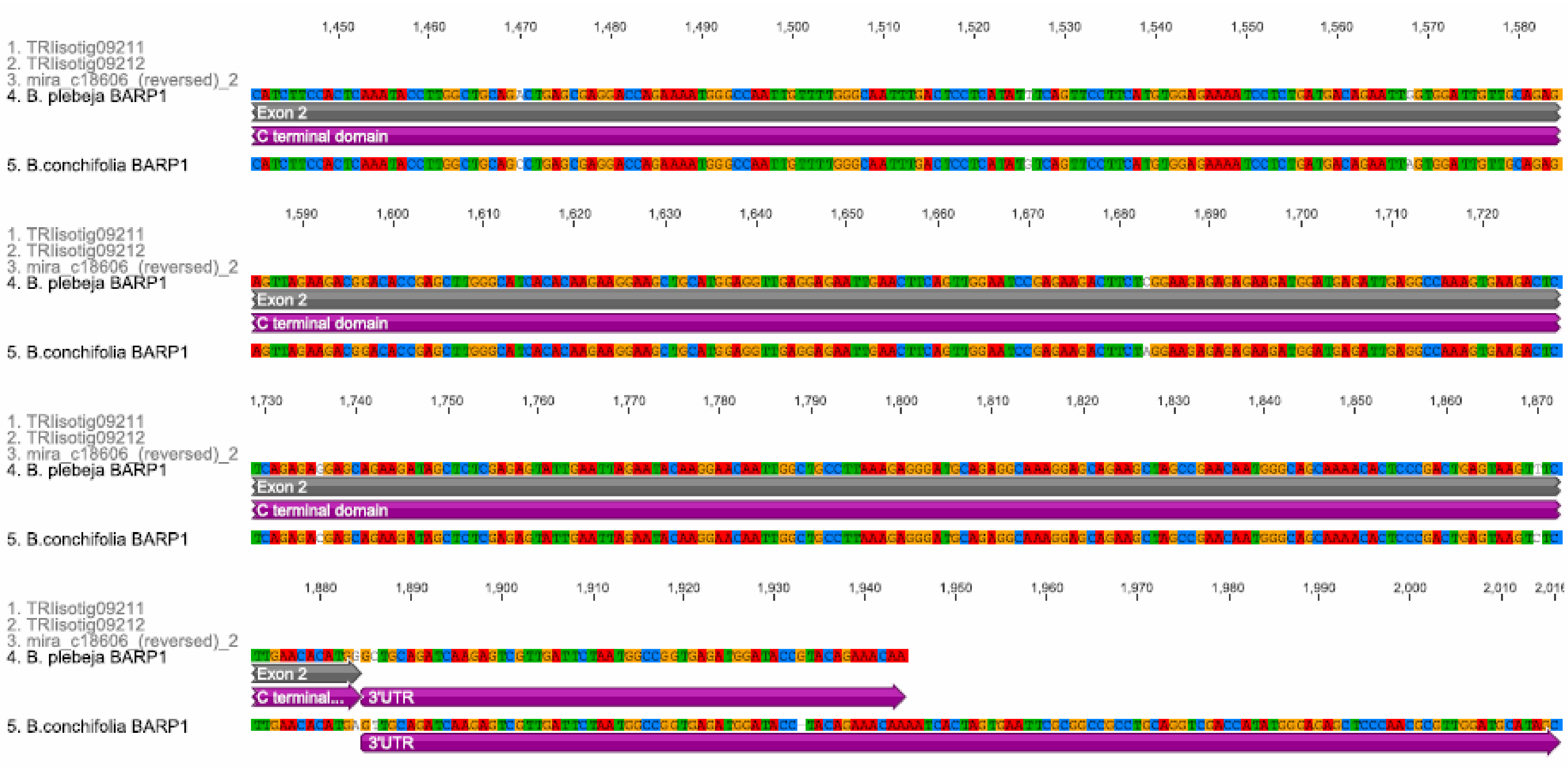
Transcriptome sequences show the evidence for alternate splicing sites in *BARP1* genes of *B. plebeja, B. conchifolia* and *B. venusta* (Figure 41). Alternative splicing can affect the mRNA and protein products of a gene in many different ways. For example alternative splicing can yield protein isoforms that differ in their amino acid sequences and hence chemical and biological activity such as change in protein ligand binding, enzymatic activity or protein localization (Grabowski and Blake, 2001). Alternative splicing has been reported to be a major source of protein diversity in metazoan organisms. Approximately 60% of human gene products undergo alternative splicing (Modrek and Lee, 2002). Alternative splicing is important in many cellular and developmental processes such as sex determination, apoptosis, axon guidance, cell excitation and contraction, and many others. Errors in splicing regulation underlie a number of different disease states (Eckardt, 2002; Drea *et al.*, 2006).

Microsatellites are detected in the *BARP1* intron which varies from 2-9 repeats of TTC (Figure 42). Microsatellites are repeating sequences of <10bp of DNA. If properly expressed, inherited length changes in microsatellites can act as ‘digital’ genetic data, which allows gradual changes in physical properties, reducing the risk of drastic mutations that might be lethal for the organism (Sosinski *et al.*, 1997). Replication slippage caused by mismatches between DNA strands during replication is the most common cause of length changes in microsatellites (Tautz, 1994). Microsatellites within introns can influence phenotypes, a GAA triplet expansion in the first intron of the *X25* gene appears to cause Friedreich Ataxia by interfering with transcription (Bidichandani, 1998). Microsatellites in the first intron of the *Asparagine synthetase* gene are linked to acute lymphoblastic leukemia (Akagi, 2009). Reductions in the repeat lengths of the EGFR gene are linked with osteosarcomas (Kersting 2008).

Figure 40 Annotations for *BARP1* genes.
Different domains of *BARP* genes are annotated on the basis of transcriptomes reads of *B. conchifolia, B. plebeja* and *B. venustra.* TRisotig09211,TRisotig09212 and mira_c18606 reversed_2 are transcriptomes reads and *B. plebeja BARP1* is a genomic DNA sequence.

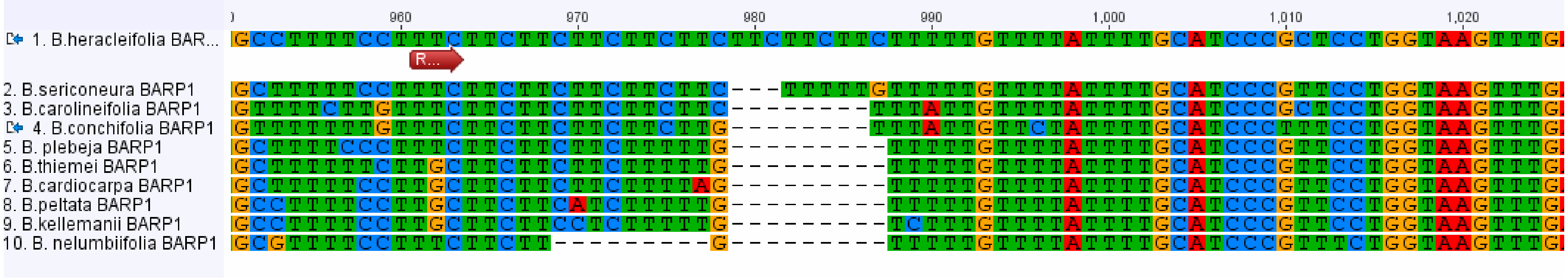



730
740
750
760
770
780
790
800
810
820
830
840
850
860
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Intron
Intron
Exon 2
Myb Domain1
5. B.conchifolia BARP1
870
880
890
900
910
920
930
940
950
960
970
980
990
1,000
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
Myb Domain1
Myb domain 2
5. B.conchifolia BARP1
1,010
1,020
1,030
1,040
1,050
1,060
1,070
1,080
1,090
1,100
1,110
1,120
1,130
1,140
1,150
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
Myb domain 2
5. B.conchifolia BARP1
1,160
1,170
1,180
1,190
1,200
1,210
1,220
1,230
1,240
1,250
1,260
1,270
1,280
1,290
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
Myb doma...
C terminal domain
5. B.conchifolia BARP1
1,300
1,310
1,320
1,330
1,340
1,350
1,360
1,370
1,380
1,390
1,400
1,410
1,420
1,430
1,440
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
C terminal domain
5. B.conchifolia BARP1


1,450 1,460 1,470 1,480 1,490 1,500 1,510 1,520 1,530 1,540 1,550 1,560 1,570 1,580
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
C terminal domain
5. B.conchifolia BARP1
1,590 1,600 1,610 1,620 1,630 1,640 1,650 1,660 1,670 1,680 1,690 1,700 1,710 1,720
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
C terminal domain
5. B.conchifolia BARP1
1,730 1,740 1,750 1,760 1,770 1,780 1,790 1,800 1,810 1,820 1,830 1,840 1,850 1,860 1,870
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
C terminal domain
5. B.conchifolia BARP1
1,880 1,890 1,900 1,910 1,920 1,930 1,940 1,950 1,960 1,970 1,980 1,990 2,000 2,010 2,01
1. TRlisotig09211
2. TRlisotig09212
3. mira_c18606_(reversed)_2
4. B. plebeja BARP1
Exon 2
C terminal...
3'UTR
5. B.conchifolia BARP1
3'UTR

```
B.heracleifolia BARP1           TTCTTCACGGAGATGTCTTCTTA---CTCCTCTCCTGCTGCTAATGCTGGATGTAAGTGTCGCTCTTGAGTTTCCGATTTTCTTCAGCGTTCGTCCTCTGCCTNTTCCTTTCTTCTTCTTCTTCT
B.thiemei BARP1                 TTCTTCGCGGAGATGTCTTCTTACTCCTCCTCTCCTGCTGCCATTGCCGGATGTAAGAGTCGCTCTTGAGTTTCCGATTTTCTTCACCGTTCGTCCTCTGCTTTTTTCTTGCTTCTTCTTCTTTT
B. plebeja BARP1                TTCTTCGCGGAGATGTCTTCTTA---CTCCTCTCCTGCTGCCATTGCCGGATGTAAGAGTCGCTCTTGAGTTTCCGATATCCTTCAGCGTTCGTCCTCTGCTTTTCCCTTTCTTCTTCTTCTTTT

TRlisotig09211                  TTCTTCGCGGAGATGTCTTCTTA---CTCCTCTCCTGCTGCCATTGCCGGATGTAAGAGTCGCTCTTGAGTTTCCGATATCCTTCAGC-------------------------------------
PLEisotig06154_(reversed)       TTCTTCGCGGAGATGTCTTCTTA---CTCCTCTCCTGCTGCCATTGCCGGATGTAAGAGTCGCTCTTGAGTTTCCGATATCCTTCAGC-------------------------------------
TRlisotig09212                  TTCTTCGCGGAGATGTCTTCTTA---CTCCTCTCCTGCTGCCATTGCCGGAT-------------------------------------------------------------------------
mira_c18606_(reversed)_2        TTCTTCGCGGAGATGTCTTCTTACTCCTCCTCTCCCGCTGCCATTGCCGGAT-------------------------------------------------------------------------
```

Figure 41 Alternative splicing in *BARP1* genes.
The alligments of genomic DNA sequences of *B. heracleifolia BARP1, B. thiemei BARP1*, B. *plebeja BARP1* and transcriptomes reads TRisotig09211, PLEisotig06154, TRisotig09212 and mira_c18606 reversed_2 of *B. conchifolia, B. plebeja* and *B. venustra* were performed in Geneious align which suggest the occurenec of alternative splicing.



```
                              )        960       970       980       990       1,000     1,010     1,020
1. B.heracleifolia BAR...     GCCTTTTCCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTTTTGTTTTATTTTGCATCCCGCTCCTGGTAAGTTTG
2. B.sericoneura BARP1        GCTTTTTCCTTTCTTCTTCTTCTTCTTC---TTTTTGTTTTTGTTTTATTTTGCATCCCGTTCCTGGTAAGTTTG
3. B.carolineifolia BARP1     GTTTTCTTGTTTCTTCTTCTTCTTCTTC--------TTTATTGTTTTATTTTGCATCCCGCTCCTGGTAAGTTTG
4. B.conchifolia BARP1        GTTTTTTTGTTTCTTCTTCTTCTTCTTG--------TTTATTGTTCTATTTTGCATCCCTTTCCTGGTAAGTTTG
5. B. plebeja BARP1           GCTTTTCCCTTTCTTCTTCTTCTTTTTG---------TTTTTGTTTTATTTTGCATCCCGTTCCTGGTAAGTTTG
6. B.thiemei BARP1            GCTTTTTTCTTGCTTCTTCTTCTTTTTG---------TTTTTGTTTTATTTTGCATCCCGTTCCTGGTAAGTTTG
7. B.cardiocarpa BARP1        GCTTTTTCCTTGCTTCTTCTTCTTTTAG---------TTTTTGTTTTATTTTGCATCCCGTTCCTGGTAAGTTTG
8. B.peltata BARP1            GCCTTTTCCTTGCTTCTTCATCTTTTTG---------TTTTTGTTTTATTTTGCATCCCGTTCCTGGTAAGTTTG
9. B.kellemanii BARP1         GCCTTTTCCTTGCTTCTTCCTCTTTTTG---------TCTTTGTTTTATTTTGCATCCCGTTCCTGGTAAGTTTG
10. B. nelumbiifolia BARP1    GCGTTTTCCTTTCTTCTT---------G---------TTTTTGTTTTATTTTGCATCCCGTTTCTGGTAAGTTTG
```

Figure 42 Microsatellites in the 5 □ intron of *Begonia* section Gireoudia.
The repeated sequences are TTC which is shown with red arrow named R. There are 9 repeats of TTC in *B. heracleifolia*, 6 in *B. sericoneura*, 6 in *B. carolineifolia*, 5 in *B. conchifolia*, 4 in *B. plebeja*, 3 in *B. thiemei*, 3 in *B. cardiocarpa*, 2 in *B. peltata*, 2 in *B. kellemanii* and 1 in *B. nelumbiifolia*.

4.6. Homology of *Begonia* genes to other *ARP* genes

To determine the orthology of *BARP* genes with other *ARP* genes amino acid sequences were retrieved from Gene Bank and compared using Geneious align. The other species *ARP* genes sequences were collected from NCBI and Phytozome. A neighbour joining tree was built using *Selaginella ARP* as an out group (Figure 43). As expected, *Begonia ARP* genes have highest similarity with *Cucumber ARP* genes (Figure 43).

4.7. *BARP1* CDS and leaf form in *Begonia*

ARP genes from *Begonia* showed a high degree of similarity to each other in coding regions as well as in non coding parts of the gene indicating that they are orthologs (see appendix B for *BARP1* genes allignments). This is supported by the fact that I was able to amplify *BARP* genes from most *Begonia* species using primers designed against a single species.

The sequence alignment for BARP1 proteins from different *Begonia* species showed high homology (pair wise identity is 97.3%) and no amino acid residues specific to simple, peltate or compound leaves (Figure 44) were found. Phylogenetic trees of these BARP protein sequences (Figure 45-46) showed that the sequences were not grouped based on final leaf morphology which suggests that evolution in coding regions of the *BARP1* does not parallel the evolution of leaf form in *Begonia* section Gireoudia.

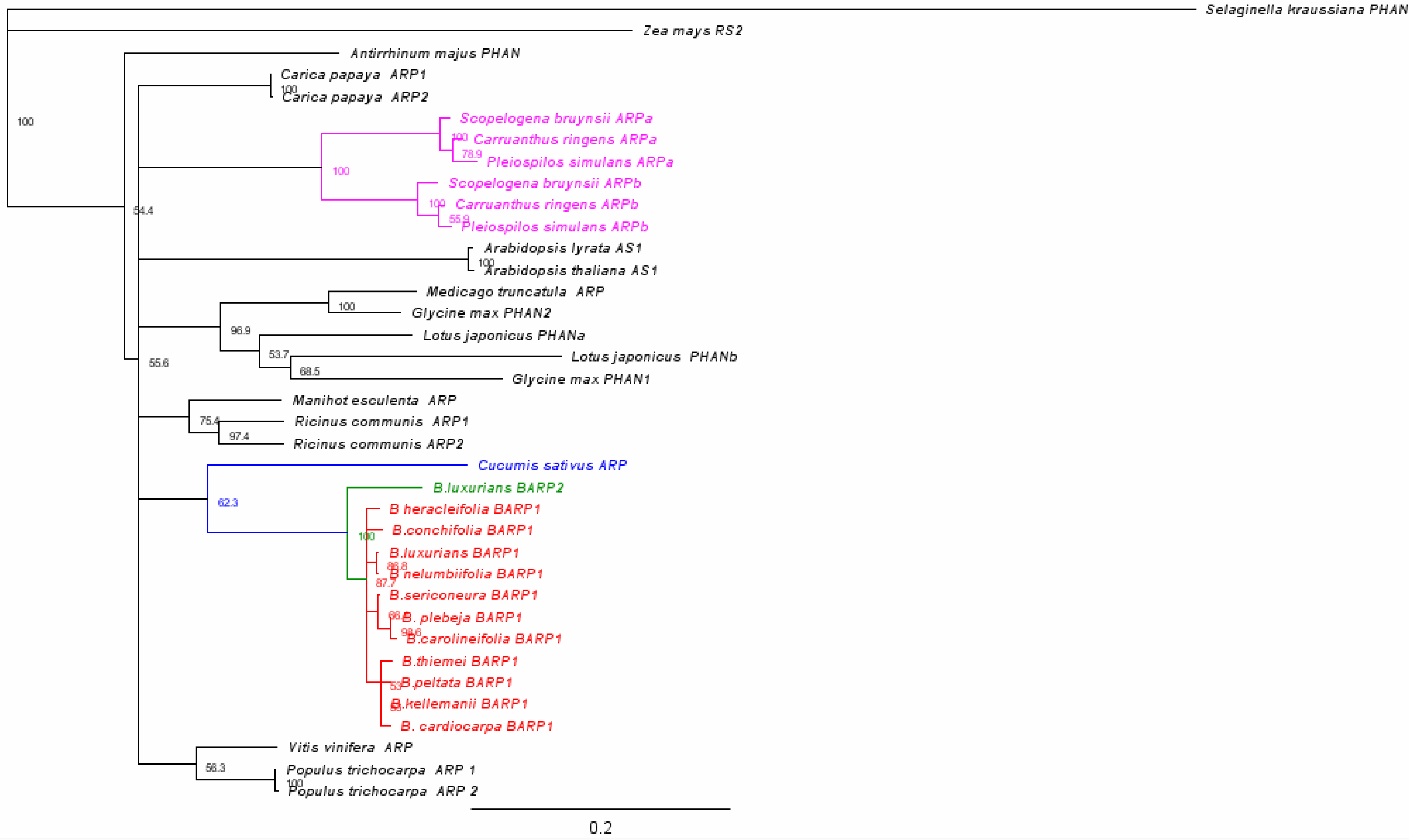


Figure 43 Phylogenetic relationships of *ARP* genes.
Jukes-Cantor Neighbour Joining tree of ARP proteins with 1000 bootstrap replicates based on proteins using *Seleginella kraussiana* PHAN as an out group. The species are not grouped on the basis of leaf morphology instead it is consistant with ITS phylogenetic tree where *BARP* genes are closely related to *Cucumis sativus ARP* genes.

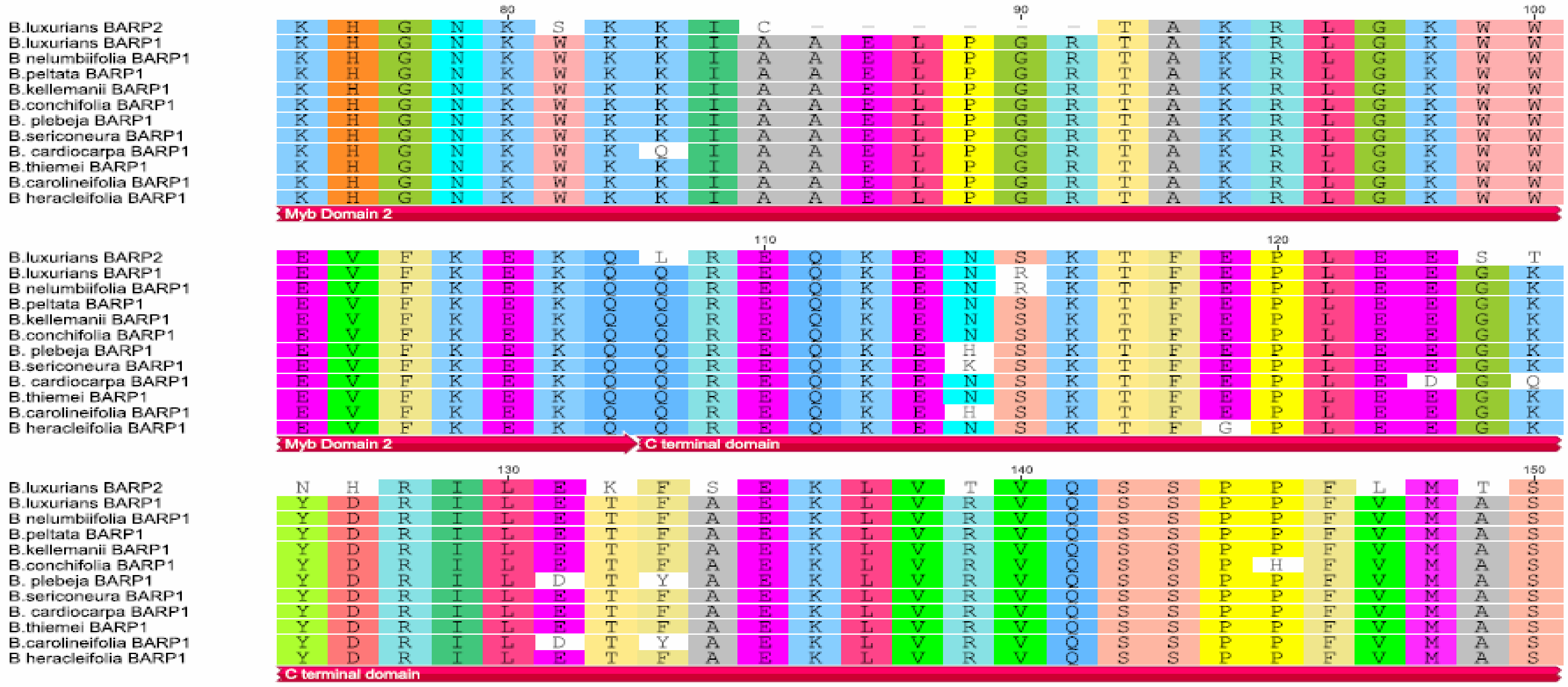
Figure 44 BARP protein alignments.

The sequence alignment for BARP1 proteins from different *Begonia* species showed high homology and no amino acid residues specific to simple, peltate or compound leaves were found.

	1									10										20					
B.luxurians BARP2	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B.luxurians BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B nelumbiifolia BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B.peltata BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B.kellemanii BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B.conchifolia BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B. plebeja BARP1	M	K	D	R	R	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B.sericoneura BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B. cardiocarpa BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B.thiemei BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B.carolineifolia BARP1	M	K	D	R	R	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R
B heracleifolia BARP1	M	K	D	R	Q	R	W	R	A	N	E	D	S	L	L	H	A	Y	V	K	Q	Y	G	P	R

					30										40										50
B.luxurians BARP2	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B.luxurians BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B nelumbiifolia BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B.peltata BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B.kellemanii BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B.conchifolia BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B. plebeja BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B.sericoneura BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B. cardiocarpa BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B.thiemei BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	K
B.carolineifolia BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R
B heracleifolia BARP1	E	W	N	L	V	S	L	R	M	N	T	P	L	N	R	D	A	K	S	C	L	E	R	W	R

										60										70					
B.luxurians BARP2	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B.luxurians BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B nelumbiifolia BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B.peltata BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B.kellemanii BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B.conchifolia BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B. plebeja BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B.sericoneura BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B. cardiocarpa BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B.thiemei BARP1	N	Y	L	K	P	G	V	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B.carolineifolia BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A
B heracleifolia BARP1	N	Y	L	K	P	G	I	K	K	G	S	L	T	K	E	E	Q	H	L	V	I	R	L	Q	A



					80										90										100
B.luxurians BARP2	K	H	G	N	K	S	K	K	I	C	–	–	–	–	–	–	T	A	K	R	L	G	K	W	W
B.luxurians BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B nelumbiifolia BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B.peltata BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B.kellemanii BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B.conchifolia BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B. plebeja BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B.sericoneura BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B. cardiocarpa BARP1	K	H	G	N	K	W	K	Q	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B.thiemei BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B.carolineifolia BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W
B heracleifolia BARP1	K	H	G	N	K	W	K	K	I	A	A	E	L	P	G	R	T	A	K	R	L	G	K	W	W

Myb Domain 2

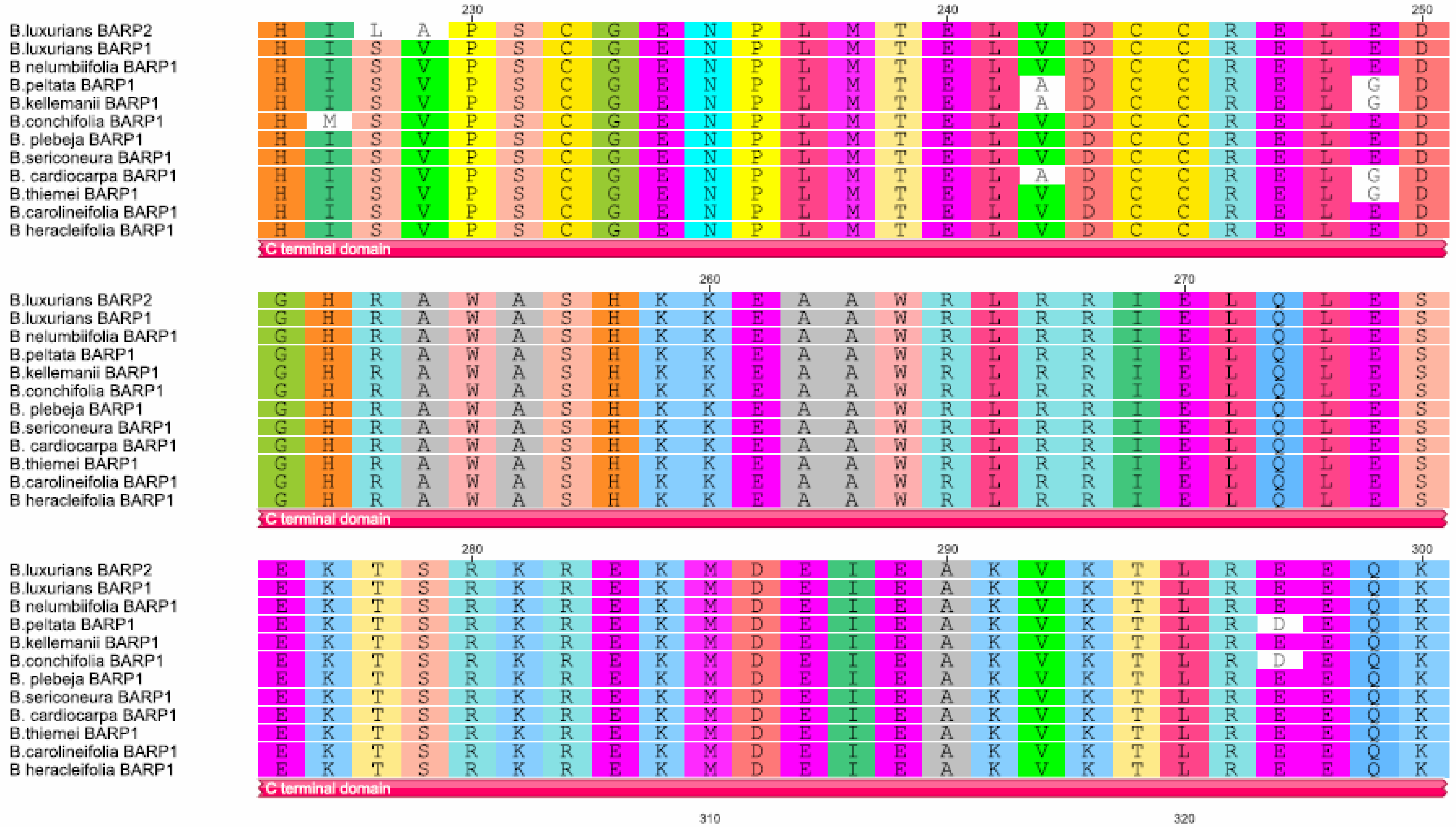
										110										120					
B.luxurians BARP2	E	V	F	K	E	K	Q	L	R	E	Q	K	E	N	S	K	T	F	E	P	L	E	E	S	T
B.luxurians BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	R	K	T	F	E	P	L	E	E	G	K
B nelumbiifolia BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	R	K	T	F	E	P	L	E	E	G	K
B.peltata BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	S	K	T	F	E	P	L	E	E	G	K
B.kellemanii BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	S	K	T	F	E	P	L	E	E	G	K
B.conchifolia BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	S	K	T	F	E	P	L	E	E	G	K
B. plebeja BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	H	S	K	T	F	E	P	L	E	E	G	K
B.sericoneura BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	K	S	K	T	F	E	P	L	E	E	G	K
B. cardiocarpa BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	S	K	T	F	E	P	L	E	D	G	Q
B.thiemei BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	S	K	T	F	E	P	L	E	E	G	K
B.carolineifolia BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	H	S	K	T	F	E	P	L	E	E	G	K
B heracleifolia BARP1	E	V	F	K	E	K	Q	Q	R	E	Q	K	E	N	S	K	T	F	G	P	L	E	E	G	K

Myb Domain 2 C terminal domain

					130										140										150
B.luxurians BARP2	N	H	R	I	L	E	K	F	S	E	K	L	V	T	V	Q	S	S	P	P	F	L	M	T	S
B.luxurians BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B nelumbiifolia BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B.peltata BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B.kellemanii BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B.conchifolia BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	H	F	V	M	A	S
B. plebeja BARP1	Y	D	R	I	L	D	T	Y	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B.sericoneura BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B. cardiocarpa BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B.thiemei BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B.carolineifolia BARP1	Y	D	R	I	L	D	T	Y	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S
B heracleifolia BARP1	Y	D	R	I	L	E	T	F	A	E	K	L	V	R	V	Q	S	S	P	P	F	V	M	A	S

C terminal domain

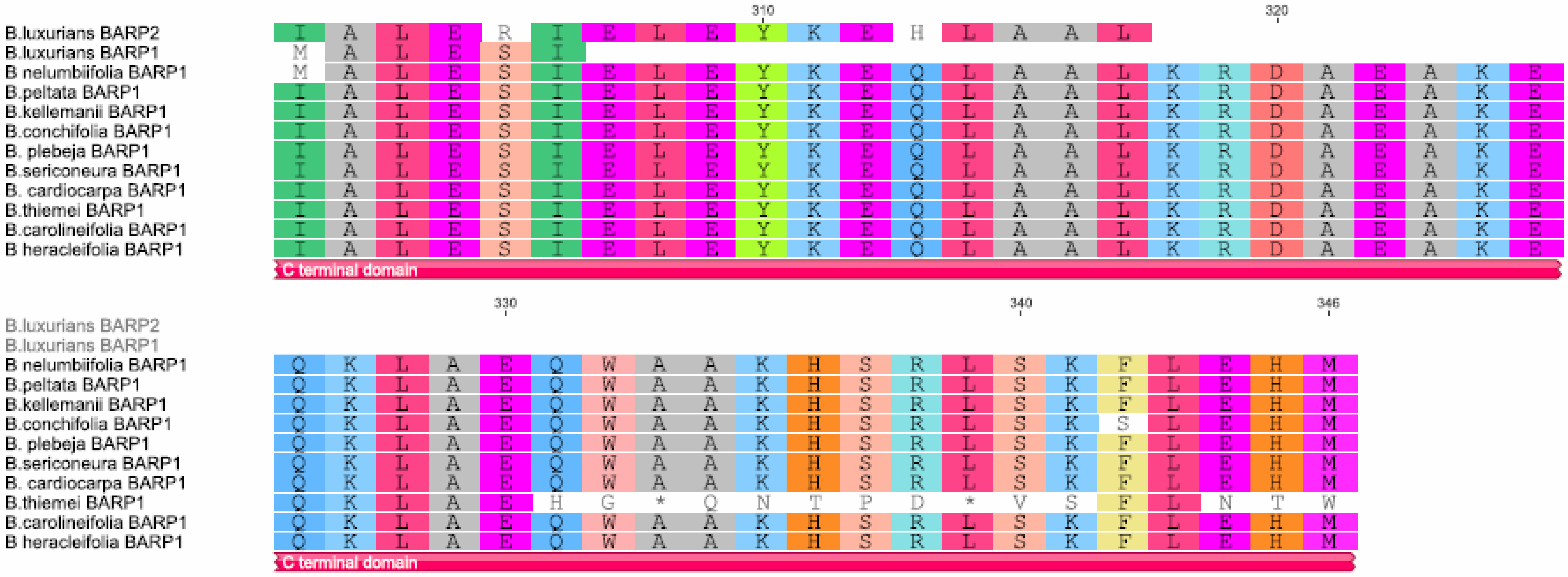
```
                                    160       170
B.luxurians BARP2          SNGGILHTDPSTPAPPLLPSWLSSS
B.luxurians BARP1          SNGGILHTDPSAPAPTLLPSWLSSS
B nelumbiifolia BARP1      SNGGILHTDPSAPAPTLLPSWLSSS
B.peltata BARP1            SNGGILHTDPSAPAPTLLPSWLSSS
B.kellemanii BARP1         SNGGILHTDPSAPAPTLLPSWLSSS
B.conchifolia BARP1        SNGGILHTDPSAPAPTLLPSWLSSS
B. plebeja BARP1           SNGGILHTDPSAPAPTLLPSWLSSS
B.sericoneura BARP1        SNGGILHTDPSAPAPTLLPSWLSSS
B. cardiocarpa BARP1       SNGGILHTDPSAPAPTLLPSWLSSS
B.thiemei BARP1            SNGGILQTDPSAPAPTLLPSWLSSS
B.carolineifolia BARP1     SNGGILHTDPSAPAPTLLPSWLSSS
B heracleifolia BARP1      SNGGILHTDPSAPAPTLLPSWLSGS
                           C terminal domain

                               180       190       200
B.luxurians BARP2          SGTSAIRPPSPSVTLSLSSSTIAPS
B.luxurians BARP1          SGTSAIRPPSPSVTLSLSSSTIAPS
B nelumbiifolia BARP1      SGTSAIRPPSPSVTLSLSSSTIAPS
B.peltata BARP1            GGTSAIRPPSPSVTLSLSSSTIAPS
B.kellemanii BARP1         SGTSAIRPPSPSVTLSLSSSTIAPS
B.conchifolia BARP1        SGTSAIRPPSPSVTLSLSSSTIAPS
B. plebeja BARP1           SGTSAIRPPSPSVTLSLSSSTIAPS
B.sericoneura BARP1        SGSSAIRPPSPSVTLSLSSSTIAPS
B. cardiocarpa BARP1       SGTSAIRPPSPSVTLSLSSSTIAPS
B.thiemei BARP1            SGTSAIRPPSPSVTLSLSSSTIAPS
B.carolineifolia BARP1     SGTSAIRPPSPSVTLSLSSSTIAPS
B heracleifolia BARP1      SGTSAIRPPSPSVTLSLSSSTIAPS
                           C terminal domain

                                    210       220
B.luxurians BARP2          STQIPWLQPERGPENGPIVLGNLTP
B.luxurians BARP1          STQIPWLQPERGPENGPIVLGNLTP
B nelumbiifolia BARP1      STQIPWLQPERGPENGPIVLGNLTP
B.peltata BARP1            STQIPWLQPERGPENGPIVLGNLTP
B.kellemanii BARP1         STQIPWLQPERGPENGPIVLGNLTP
B.conchifolia BARP1        STQIPWLQPERGPENGPIVLGNLTP
B. plebeja BARP1           STQIPWLQTERGPENGPIVLGNLTP
B.sericoneura BARP1        STQIPWLQPERGPENGPIVLGNLTP
B. cardiocarpa BARP1       STQIPWLQPERGPENGPIVLGNLTP
B.thiemei BARP1            STQIPWLQPERGPENGPIVLGNLTP
B.carolineifolia BARP1     STQIPWLQTERGPENGPIVFGNLTP
B heracleifolia BARP1      STQIPWLQPERGPENGPIVLGNLT-
                           C terminal domain
```



B.luxurians BARP2	H I L A P S C G E N P L M T E L V D C C R E L E D
B.luxurians BARP1	H I S V P S C G E N P L M T E L V D C C R E L E D
B nelumbiifolia BARP1	H I S V P S C G E N P L M T E L V D C C R E L E D
B.peltata BARP1	H I S V P S C G E N P L M T E L A D C C R E L G D
B.kellemanii BARP1	H I S V P S C G E N P L M T E L A D C C R E L G D
B.conchifolia BARP1	H M S V P S C G E N P L M T E L V D C C R E L E D
B. plebeja BARP1	H I S V P S C G E N P L M T E L V D C C R E L E D
B.sericoneura BARP1	H I S V P S C G E N P L M T E L V D C C R E L E D
B. cardiocarpa BARP1	H I S V P S C G E N P L M T E L A D C C R E L G D
B.thiemei BARP1	H I S V P S C G E N P L M T E L V D C C R E L G D
B.carolineifolia BARP1	H I S V P S C G E N P L M T E L V D C C R E L E D
B heracleifolia BARP1	H I S V P S C G E N P L M T E L V D C C R E L E D

B.luxurians BARP2	G H R A W A S H K K E A A W R L R R I E L Q L E S
B.luxurians BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B nelumbiifolia BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B.peltata BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B.kellemanii BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B.conchifolia BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B. plebeja BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B.sericoneura BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B. cardiocarpa BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B.thiemei BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B.carolineifolia BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S
B heracleifolia BARP1	G H R A W A S H K K E A A W R L R R I E L Q L E S

B.luxurians BARP2	E K T S R K R E K M D E I E A K V K T L R E E Q K
B.luxurians BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B nelumbiifolia BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B.peltata BARP1	E K T S R K R E K M D E I E A K V K T L R D E Q K
B.kellemanii BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B.conchifolia BARP1	E K T S R K R E K M D E I E A K V K T L R D E Q K
B. plebeja BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B.sericoneura BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B. cardiocarpa BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B.thiemei BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B.carolineifolia BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K
B heracleifolia BARP1	E K T S R K R E K M D E I E A K V K T L R E E Q K



```
B.luxurians BARP2          I A L E R I E L E Y K E H L A A L
B.luxurians BARP1          M A L E S I
B nelumbiifolia BARP1      M A L E S I E L E Y K E Q L A A L K R D A E A K E
B.peltata BARP1            I A L E S I E L E Y K E Q L A A L K R D A E A K E
B.kellemanii BARP1         I A L E S I E L E Y K E Q L A A L K R D A E A K E
B.conchifolia BARP1        I A L E S I E L E Y K E Q L A A L K R D A E A K E
B. plebeja BARP1           I A L E S I E L E Y K E Q L A A L K R D A E A K E
B.sericoneura BARP1        I A L E S I E L E Y K E Q L A A L K R D A E A K E
B. cardiocarpa BARP1       I A L E S I E L E Y K E Q L A A L K R D A E A K E
B.thiemei BARP1            I A L E S I E L E Y K E Q L A A L K R D A E A K E
B.carolineifolia BARP1     I A L E S I E L E Y K E Q L A A L K R D A E A K E
B heracleifolia BARP1      I A L E S I E L E Y K E Q L A A L K R D A E A K E

B.luxurians BARP2
B.luxurians BARP1
B nelumbiifolia BARP1      Q K L A E Q W A A K H S R L S K F L E H M
B.peltata BARP1            Q K L A E Q W A A K H S R L S K F L E H M
B.kellemanii BARP1         Q K L A E Q W A A K H S R L S K F L E H M
B.conchifolia BARP1        Q K L A E Q W A A K H S R L S K S L E H M
B. plebeja BARP1           Q K L A E Q W A A K H S R L S K F L E H M
B.sericoneura BARP1        Q K L A E Q W A A K H S R L S K F L E H M
B. cardiocarpa BARP1       Q K L A E Q W A A K H S R L S K F L E H M
B.thiemei BARP1            Q K L A E H G * Q N T P D * V S F L N T W
B.carolineifolia BARP1     Q K L A E Q W A A K H S R L S K F L E H M
B heracleifolia BARP1      Q K L A E Q W A A K H S R L S K F L E H M
```

Figure 45 Phylogenetic trees of BARP1 proteins.
The alignment showed that the sequences were not grouped based on final leaf morphology which suggests that evolution in coding regions of the *BARP1* does not parallel the evolution of leaf form in *Begonia* section Gireoudia. Peltate species are highlighted with red colour, the compound leaved species with blue and the simple leaved species with black colour.

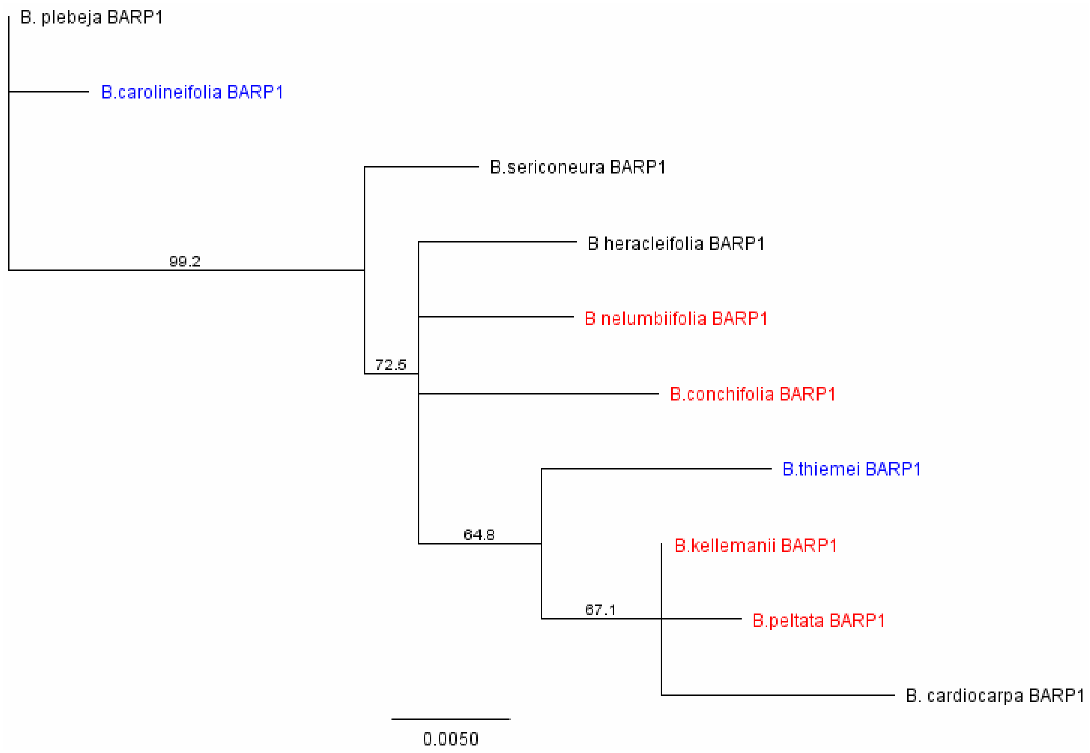
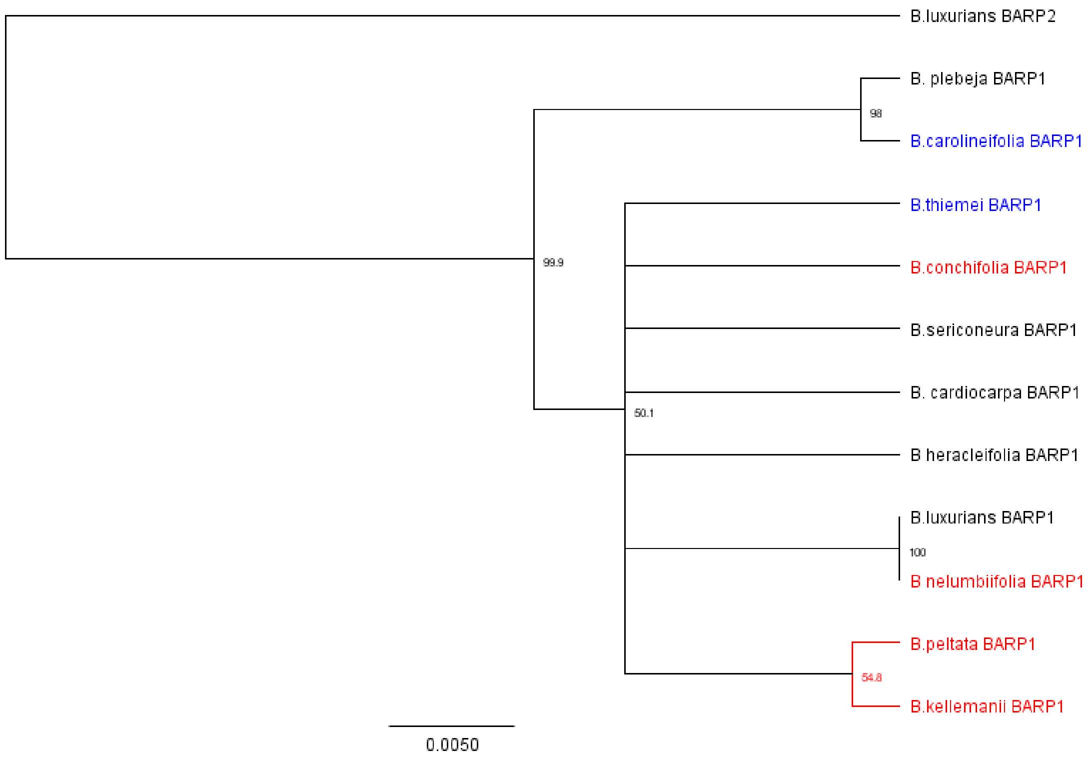


Figure 46 Phylogenetic trees of BARP proteins using *B. luxurians* BARP2 as an out group.

The tree showed that the sequences are not grouped based on final leaf morphology which suggests that evolution in coding regions of the *BARP1* does not parallel the evolution of leaf form in *Begonia* section Gireoudia*.* The "red nodes" represent the peltate leaved species, "blue nodes" represent the compound leaved species and "black nodes" represent the simple leaved species in *Begonia* section Gireoudia.



4.8. Rates of evolution for *ARP* genes

In order to check the selective pressure on different domains of *ARP* genes I calculated the synonymous (dS) and non synonymous (dN) substitutions of *ARP* genes for *Arabidopsis thaliana, Selaginella kraussiana*, Tobacco, *Glycine max*, *Vitis*, *Sorghum, Medicago*, *Pisum sativum*, *Zea mays*, *Malus domestica*, *Cardamine hirsuta* and ten *Begonia* species using the codon based Z-Test of selection in MEGA (Molecular Evolutionary Genetics Analysis) (Ttable 18). Myb domain1 has more synonymous substitutions (dS) than non synonymous (dN) substitutions suggesting the occurrence of purifying selection (Z- Test P value for purifying selection is 0.00). Myb domain 2 has more non synonymous (dN) substitutions than synonymous substitutions (dS) indicating the presence of positive selection (dN>dS) (Z- Test P value for positive selection is 0.00) whilst the C-terminal domain is under neutral selection (dN=dS) (Z- Test P value for purifying selection is 0.362). When I tested the selective pressure on *Begonia BARP1* only, again Myb domain1 has more synonymous substitutions (dS) than non synonymous (dN) substitutions suggesting the prevalence of purifying selection. Myb domain2 has more non synonymous (dN) substitutions than synonymous substitutions (dS) indicating the presence of positive selection (dN>dS) while C-terminal domain is under neutral selection (dN=dS) (Table 18).

So Myb domain 2 is under selection pressure in Angiosperm *ARPs* and two copies of *BARP* genes have differences in this domain. This suggests that the 18 bp deletion in the Myb domain 2 of *BARP1* may be significant from an adaptive point of view.

Table 18 Tests for selection on *ARP* genes.

		Myb Domain1	Myb Domain2	C-terminal Domain
dS	*BARP1*	0.063	0.004	0.01
	ARP	0.09442	0.0945	1.3196
dN	*BARP1*	0.005	0.016	0.01
	ARP	0.0938	0.2434	1.43
dS/dN	*BARP1*	1.26	0.25	1.15
	ARP	10.06	0.389	1.00
Test for selection (Codon based Z test for Positive selection (HA dN > dS))				
P-Value	*BARP1*	1.00	0.049	1.00
	ARP	1.00	0.00	1.00

The number of synonymous substitutions per synonymous site and the number of non-synonymous substitutions per non-synonymous site from averaging over all sequence pairs are shown. This was obtained by a bootstrap procedure (1000 replicates). Analyses were conducted using the Nei-Gojobori model. The probability of rejecting the null hypothesis of strict-neutrality (dN = dS) in favor of the alternative hypothesis (dN > dS) is shown. Values of P less than 0.05 are considered significant at the 5% level. dS and dN are the numbers of synonymous and nonsynonymous substitutions per site, respectively. The analysis involved 10 nucleotide sequences for *BARP1* genes and 21 for *ARP* genes. All ambiguous positions were removed for each sequence pair. Evolutionary analyses were conducted in MEGA5.

4.9. Structural conservation of protein sequences for *ARP* genes

3D molecular modelling of the *BARP1*, *BARP2* and *AS1* proteins was done to determine the level of structural conservation at the protein level. The CDS of *BARP1*, *BARP2* and *AS1* were translated in Geneious. The proteins were submitted into SWISS-MODEL WORKSPACE (Arnold *et al*., 2006; Kiefer *et al*., 2009) and PHYRE 2 (**P**rotein **H**omology/analog**Y** **R**ecognition **E**ngine 2) (Kelley and Sternberg, 2009) one by one for the prediction of protein structures under automated mode. Both generated a similar 3D structure for *BARP1*, *BARP2* and *AS1*. The proteins were superimposed on each other in Deep view (Arnold *et al*., 2006; Kiefer *et al*., 2009) to deduce the structural alignments and to find the similarities and differences in their structures.

The predicted 3D overlays of *BARP1* and *BARP2*, *BARP1* and *AS1*, and *BARP2* and *AS1* superimposed onto each other revealed that Myb domain1 is highly conserved between *BARP1*, *BARP2* and *AS1* (Figure 47). The predicted 3D overlays for *BARP1* and *AS1* superimposed onto each other revealed very high levels of conservation for Myb domain 2 as well (Figure 47 B). The only noticeable structural dissimilarity observed between the *BARP1* and *BARP2*, and *BARP2* and *AS1* is in Myb domain2 (Figure 47 A & C).

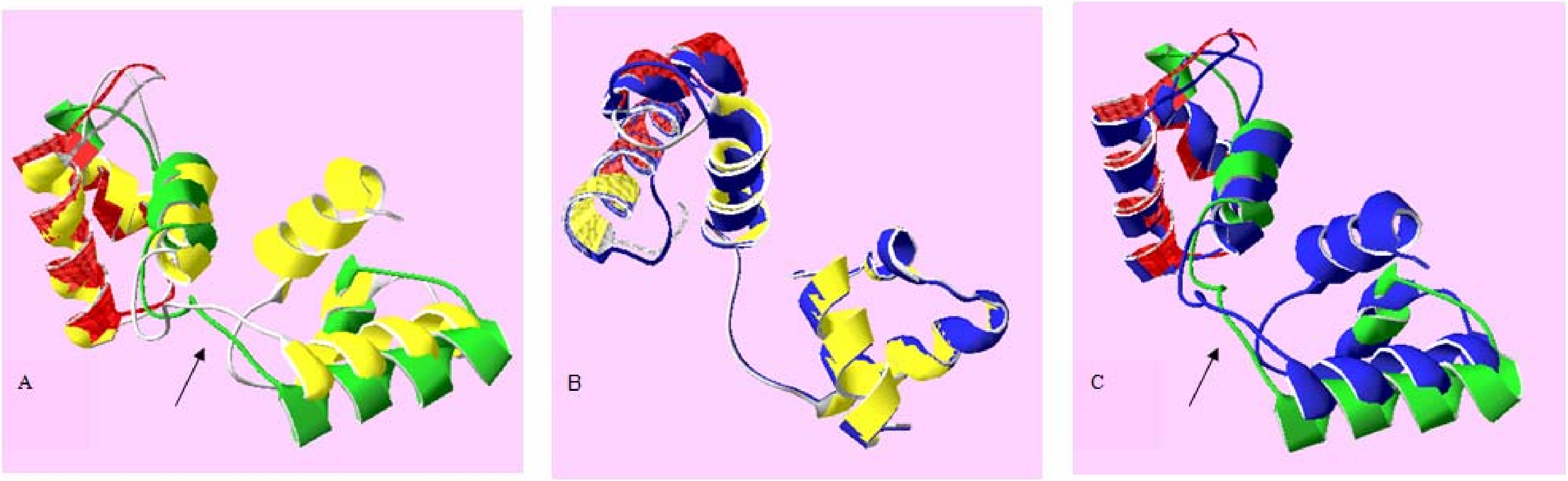


Figure 47 3D predictive modeling of *ARPs* (A) *BARP1* and *BARP2* from *B. luxurians* (B) *BARP1* and *AS1* and (C) *BARP2* and *AS1* superimposed on to each other.

The 3D structure of *BARP1* (A & B) is represented in yellow, *BARP2* (A & C) in green and *AS1* (B & C) in blue. Myb Domain1 is represented with red color in A, B & C. Black arrows indicated the major structural dissimilarity between the *BARP1* and *BARP2*; *BARP2* and *AS1*.

4.10. Amplification of *KNOX* genes from *Begonia*

KNOX genes were first sequenced from *Begonia* by Will Goodall-Copestake in an attempt to find single copy nuclear loci for phylogenetics (Goodall-Copestake 2005, PhD thesis). Unfortunately *STM*-like *KNOX* genes proved to be duplicated in all species tested in *Begonia* and in the closest relative, *Hildebrandia sanwichensis*. Using the degenerate primers designed by Will, Jack Cavers (undergraduate student 2005) amplified two *STM*-like *KNOX* genes (*KNB1* and *KNB2*) from *B. luxurians*. He showed that the two genes are expressed at different levels (Figure 48). Another MSc student (2007) Fatima Dahmani and Clare Rickerby (undergraduate student, 2007) amplified partial CDS of *KNOX* genes from compound leaved species of *Begonia*; *B. carolineifolia* and *B. theimei* and reported the presence of two *STM*-like genes from *B. carolineifolia* and one locus from *B. theimei*.

I started amplifying complete CDS of *KNOX* genes from fourteen species of *Begonia* section Gireoudia. As with *BARP*, I used TAIL PCR to extend sequences 5′ and 3′. Intron 2 is very long (over 1.5kb in some species) and had proved difficult to sequence through. I designed new primers in exon2 and exon3 to sequence the complete genes.

4.11. Amplification of *KNOX* promoters

Differential expression of *KNOX* genes across different species has been found to be linked with leaf morphology (Hay & Tsiantis, 2006; Kimura *et al*., 2008). The difference in the expression pattern of *KNOX* genes between these species is driven by the variation in the promoter region of *KNOX* genes. Two conserved non coding sequence (CNS) in the *STM* promoter were found and named as the RB box and the K-box. The RB box is present in Eudicot species only except in *Arabidopsis* and *Cardamine* whereas the K-box is found in all Angiosperms surveyed. The K-box was shown to control the repression of *KNOX* genes in many simple leaved species (Uchida *et al*., 2007).

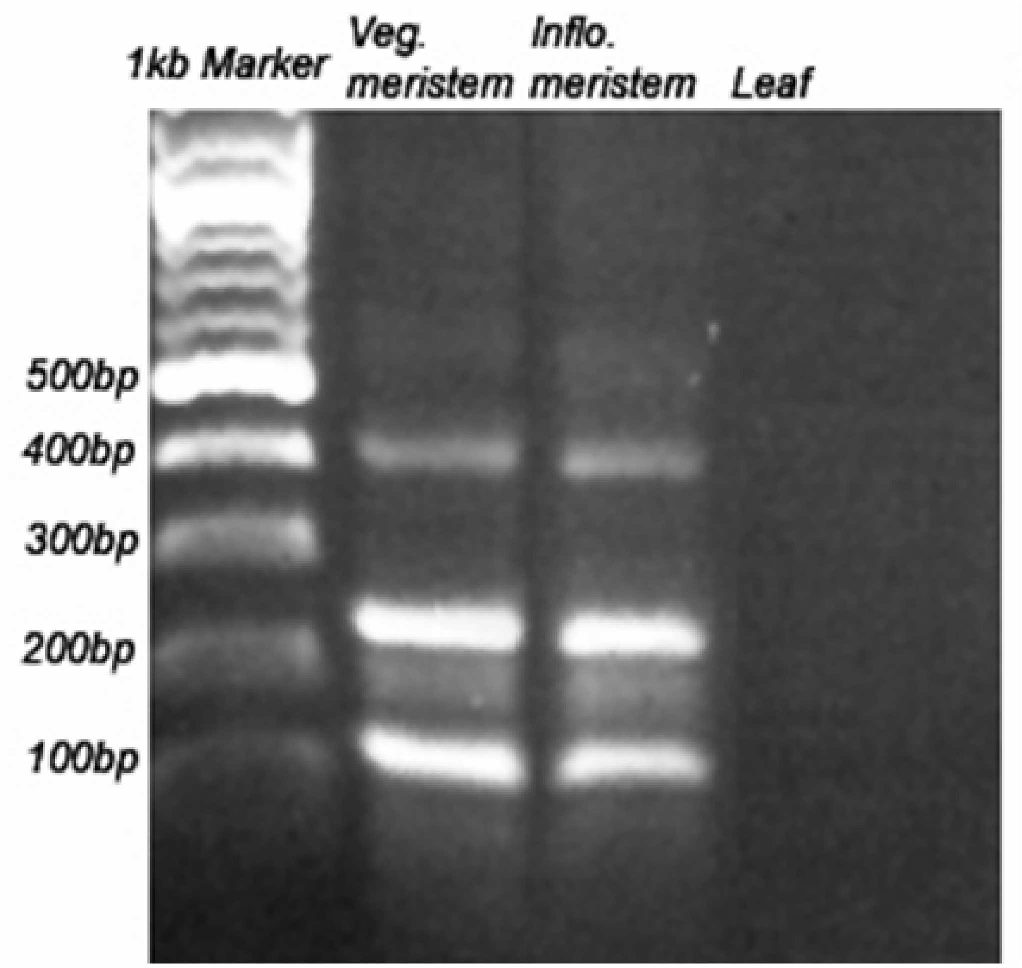


Figure 48 RT PCR on *B. luxurians* cDNA for *KNBs*.
Both *KNB1* and *KNB2* are expressed in the vegetative meristem and the inflorescence meristem. Neither *KNB1* nor *KNB2* showed expression in leaves. Expected sizes for *KNB1* digested with *Rsa*1 are 412bp and 195 bp. Expected sizes for *KNB2* digested with *Rsa*1 are 70bp, 106bp, 132bp and 213 bp. *KNB2* is expressed at higher levels than *KNB1* (C. Kidner).

I amplified the *KNB* promoter region from *Begonia KNB1* and *KNB2* genes of section Gireoudia species in order to determine if *Begonia KNOX* genes contain a K-Box and to look for similarities in the promoter regions of compound as compared to simple leaved species. The length of *KNB1* and/or *KNB2* genes and presence of K-BOX amplified from *Begonia* species is presented in table 19.

4.12. Amplification of K-box from *KNB1*

I amplified the K-box of *KNB1* from *Begonia* species using degenerate primers designed on the core RB-box (forward primer) and core K-box (reverse primer) shown in black boxes in Figure 49. The region between the K-box and exon1 was amplified with primers designed on K-box (forward primer) and exon1 (28R) shown in red boxes in Figure 49. The sequence gap between two PCR products of above described primer sets was filled using primers (-440F & 28R and -321F and 116R) as shown in Figure 49. Multiple sequences reads for the same regions were amplified with primer sets of -321F & 1087R, -440F & *KNOX*HOX2 and -321F and *KNOX*Beg R.

4.13. Amplification of K-box from *KNB2*

I amplified the *KNB2* K-BOX from *Begonia* section Gireoudia species using TAIL PCRs as described in 2.2.4. I did 3 rounds of TAIL PCR 3 times using forward primers *KNB*T1, T2, T3; *KNB*T4, T5, T6 and *KNB*herT1, T2 & T3 and reverse primers AD1-AD5 to amplify 5′UTR and some parts of promoters (Figure 50).

I cloned the TAIL3 PCR products of approximately 700-800bp into pGEM TEasy vector to sequence with SP6 and T7 primers. The sequences I got back were only 130-150bp long and did not overlap so I designed primers at the ends of the sequences to amplify the middle portion. This was the problem with every TAIL PCR products cloned into pGEM TEasy vector and it took longer to sequence the K-BOX from *KNB2* through TAIL PCRs (Figure 51).

Table 19 Class 1 *KNOX* genes in *Begonia*

Species	*KNB1*			*KNB2*		
	Total (bp)	Complete gene?	K-Box	Total (bp)	Complete?	K-Box
B. carolineifolia	2210	Yes	Yes	2220	Yes	Yes
B. heracleifolia	2324	Yes	Yes	771	No	Yes
B. thiemei	2203	Yes	Yes	952	No	_
B. nelumbiifolia	1937	Yes	Yes	817	No	_
B. conchifolia	445	No	_	1612	No	_
B. cardiocarpa	1448	No		_	_	_
B. peltata	939	No	Yes	_	_	_
B. sericoneura	1496	No	Yes	_	_	_
B. stigmosa	_	_	_	892	No	_
B. plebeja	400	_	_	482	No	_

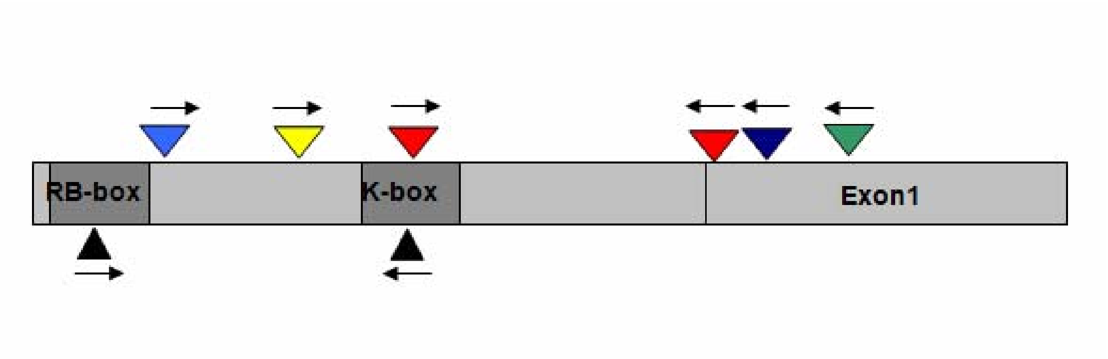


Figure 49 Schematic map of *KNB1* gene showing the location of primers.
-440F (blue triangle) & 28R (dark blue triangle), -321F (yellow triangle), 116R (green triangle), RBF, K-boxR (black triangle), K-boxF, 28R (red triangle) primers.

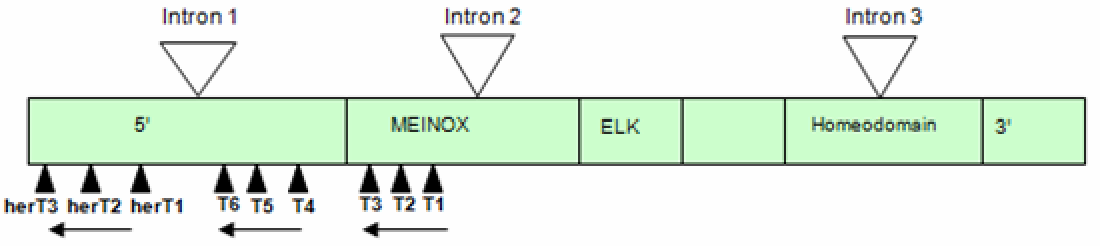


Figure 50 Schematic map of *KNB2* gene showing location of TAIL primers.
Primers are shown with black triangles.

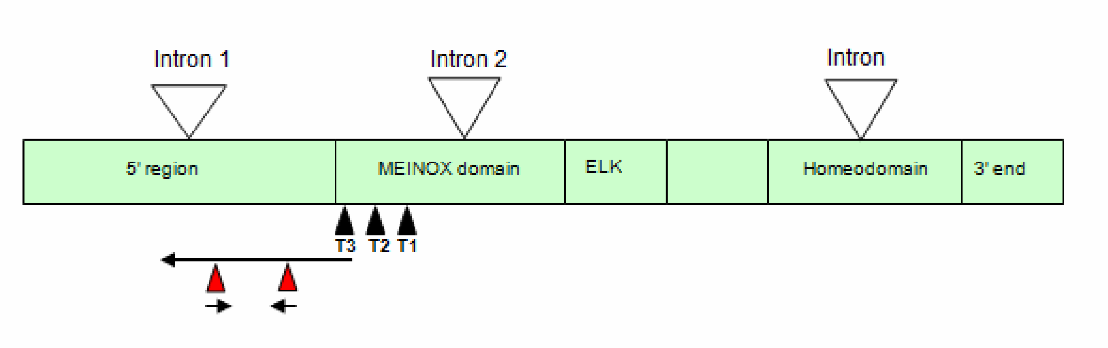


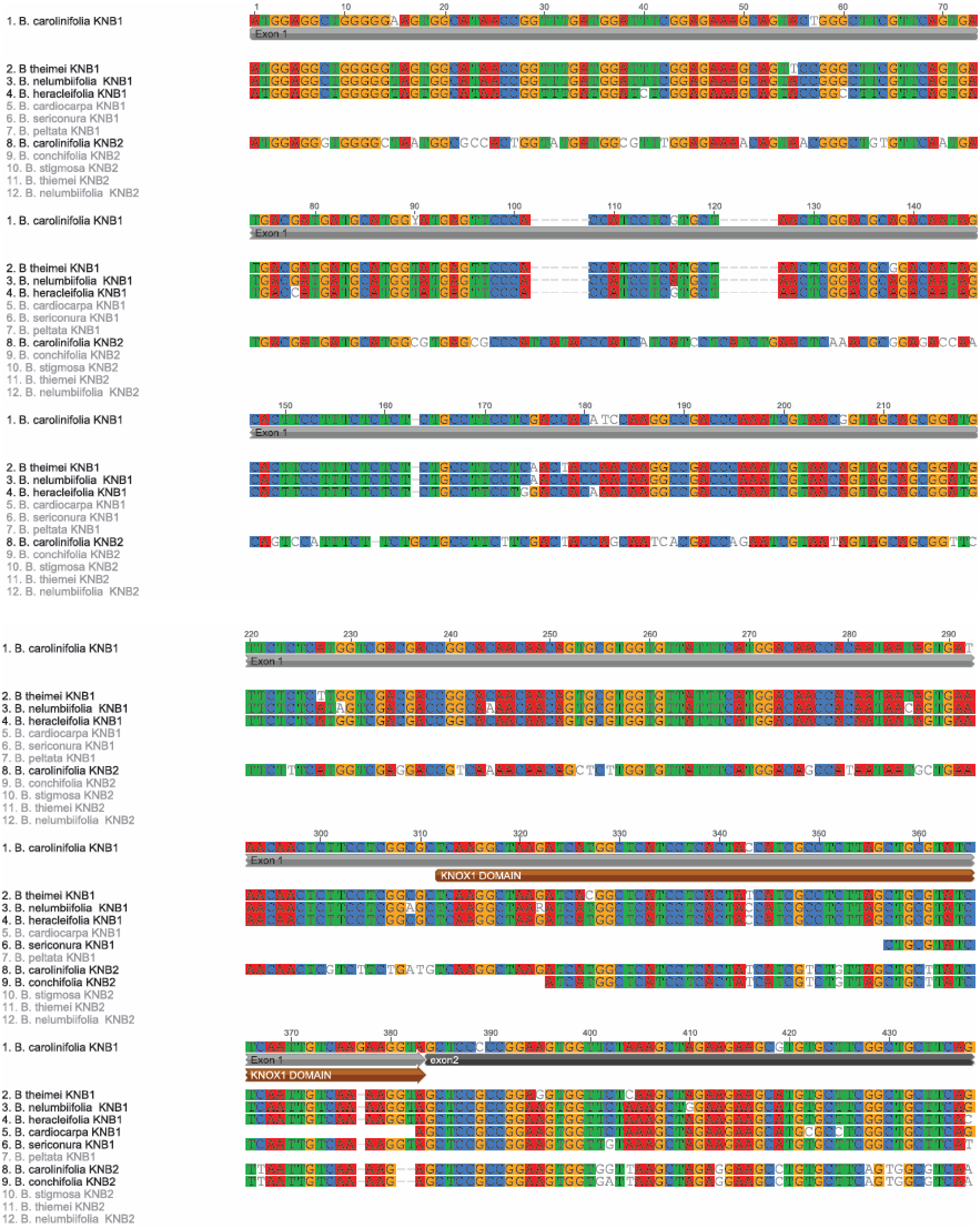
Figure 51 Physical map of *KNB2* gene showing the location of sequencing primers.
The primers used to amplify fragments are indicated by black triangles names T1, T2 and T3. The PCR fragment is represented by long arrow. The sequencing primers are indicated by red triangles.

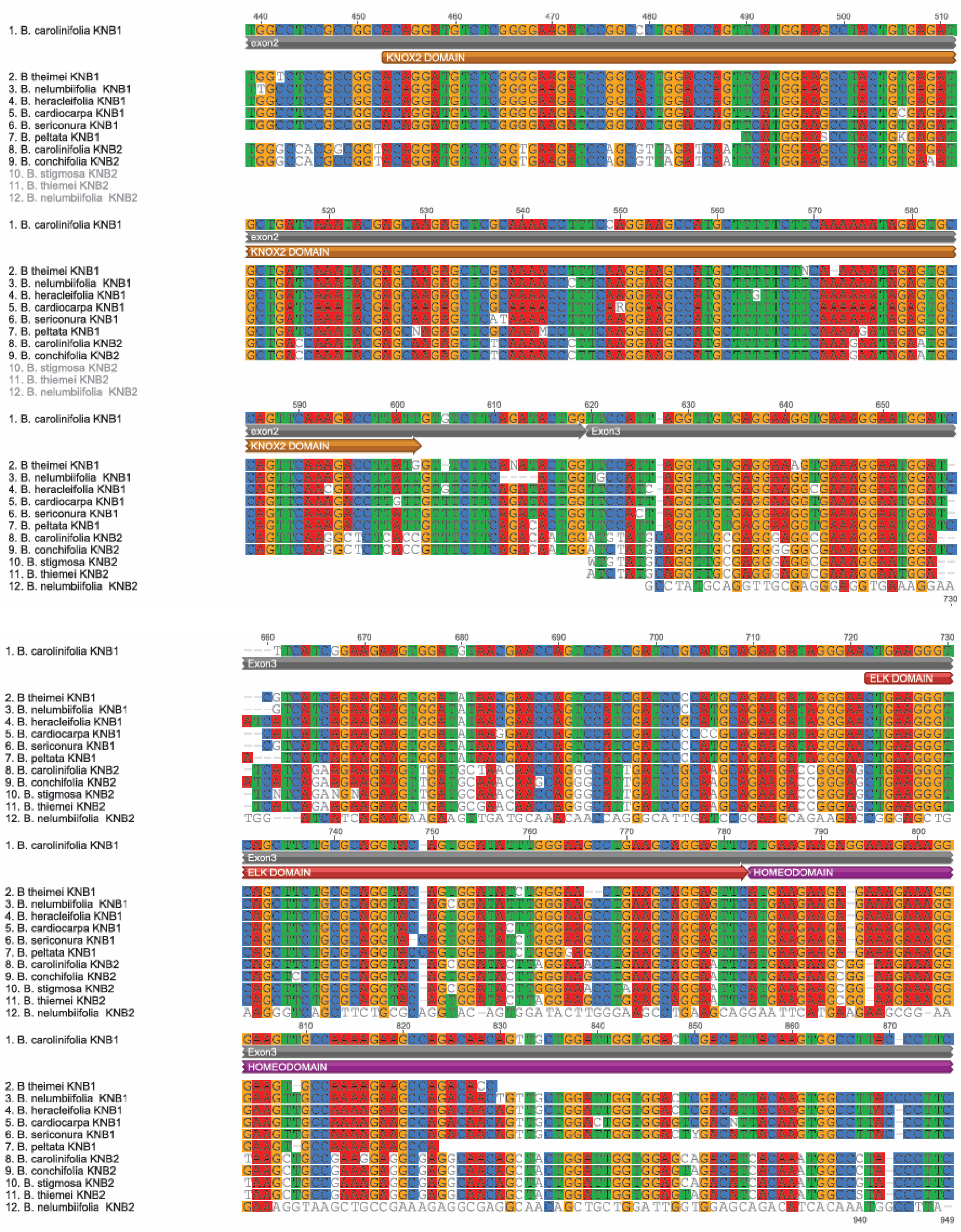
4.14. Sequence analysis for *Begonia* Class1 *KNOX* genes (*KNBs*)

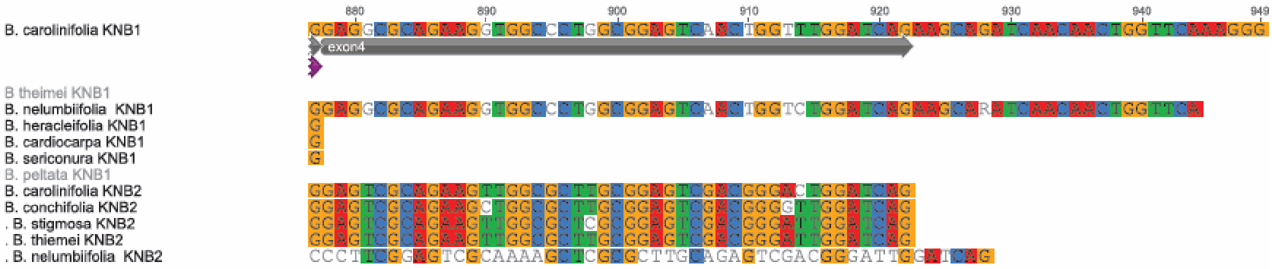
KNOX protein consists of a MEINOX domain, an ELK domain and a Homeo domain. *KNOX* genes are divided into two classes in plants partly by the presence of an intron in the ELK domain of class 2 *KNOX* genes but not class 1 *KNOX* genes (Kerstetter *et al*., 1994). *KNOX* genes from *Begonia* (*KNB*) showed all of the known conserved domains and belong to class 1 *KNOX* genes as they did not have an intron in the ELK domain. The *KNOX* like genes from *Begonia* showed higest homology with *STM* like genes when NCBI Blast programme was used.

There were two copies of *SHOOTMERISTEMLESS* like genes in *Begonia* named as *KNB1* and *KNB2*. Both *KNBs* showed a high degree of similarity to each other in coding regions as well as in noncoding parts of the gene indicating that they are paralogs. CDS of *KNB1* and *KNB2* genes are 62.5% identical in *B. carolineifolia*. *B. carolineifolia KNB1* matched with *Prunus persica STM2* genes (89%) and *Weddellina squamulosa WsSTM* (90%) on blast hit. And *B. carolineifolia KNB2* matched with *Populas trichocarpa* (97%), *Prunus persica STM2* genes (90%) and *Weddellina squamulosa WsSTM* (95%) and large number of other *STM* like genes. All *KNB1* CDS from *B. heracleifolia*, *B. thiemei*, *B. nelumbiifolia, B. peltata, B. cardiocarpa, B. sericoneura* and *B. carolineifolia* have 72.5% identical sites with 87.2% pair wise identity with each other. *KNB2* CDS from *B. heracleifolia*, *B. thiemei*, *B. nelumbiifolia, B. conchifolia, B. stigmosa* and *B. carolineifolia* has 94.9% identical sites with 94.6% pair wise identity. Table 19 shows the number and length of Class 1 *KNOX* genes amplified from *Begonia* section Gireoudia species. The alignments of CDS of *KNB1* and *KNB2* genes are presented in Figure 52.

Figure 52 *KNB1* and *KNB2* are aligned in Geneious align. Different domains are annotated using NCBI domain finder tool and named accordingly.




1. B. carolinifolia KNB1
2. B theimei KNB1
3. B. nelumbiifolia KNB1
4. B. heracleifolia KNB1
5. B. cardiocarpa KNB1
6. B. sericonura KNB1
7. B. peltata KNB1
8. B. carolinifolia KNB2
9. B. conchifolia KNB2
10. B. stigmosa KNB2
11. B. thiemei KNB2
12. B. nelumbiifolia KNB2
exon2
Exon3
KNOX2 DOMAIN
ELK DOMAIN
HOMEODOMAIN
440
450
460
470
480
490
500
510
520
530
540
550
560
570
580
590
600
610
620
630
640
650
660
670
680
690
700
710
720
730
740
750
760
770
780
790
800
810
820
830
840
850
860
870
940
949



B. carolinifolia KNB1 GGAGGCGCAGAAGGTGGCCCTGGCGGAGTCAACTGGTTTGGATCAGAAGCAGATCAACAACTGGTTCAAAGGG

B theimei KNB1

B. nelumbiifolia KNB1 GGAGGCGCAGAAGGTGGCCCTGGCGGAGTCAACTGGTCTGGATCAGAAGCARATCAACAACTGGTTCA

B. heracleifolia KNB1 G

B. cardiocarpa KNB1 G

B. sericonura KNB1 G

B. peltata KNB1

B. carolinifolia KNB2 GGAGTCGCAGAAGTTGGCGCTTGCGGAGTCGACGGGACTGGATCAG

B. conchifolia KNB2 GGAGTCGCAGAAGCTGGCGCTTGCGGAGTCGACGGGGTTGGATCAG

. B. stigmosa KNB2 GGAGTCGCAGAAGTTGGCGCTCGCGGAGTCGACGGGATTGGATCAG

. B. thiemei KNB2 GGAGTCGCAGAAGTTGGCGCTTGCGGAGTCGACGGGATTGGATCAG

. B. nelumbiifolia KNB2 CCCTTCGGAGTCGCAAAAGCTCGCGCTTGCAGAGTCGACGGGATTGGATCAG

4.15. Homology of *Begonia* genes with other *KNOX* genes

To test the homology of *Begonia KNOX* genes with other species' *KNOX* genes, amino acid sequences were aligned in Geneious. The other species' *KNOX* genes sequences were collected from GenBank and Phytozome. *Begonia KNOX* genes have highest homology to *SHOOTMERISTEMLESS* (*STM*) genes. A Neighbour Joining tree based on amino acid sequences was built in Geneious and rooted with *Acetabularia acetabulum KNOX* genes (Figure 53). A Neighbour Joining tree based on nucleotide sequences of *STM*-like genes is also built in Geneious and rooted with the class 2 *KNOX* gene, *KNAT4* of *Arabidopsis* (Figure 54). *Begonia KNB* genes formed a clade indicating that the *KNB* duplication is *Begonia* specific. The Homeodomain is highly conserved between different *KNOX* genes. A tree based on Homeodomain protein sequences is built in Geneious using *Acetabularia acetabulum KNOX* genes as an out group (Figure 55).

4.16. Similarity of *KNBs* K-box with other species *STM* K-box

The K-BOX is present close to the translation initiation codon of *STM* genes and the K-BOX controls *STM* expression (Uchida *et al.*, 2007). A GUS gene was fused with *STM* promoter with or without K-BOX to drive reporter gene expression in *Arabidopsis* and tobacco. These experiments showed that the K-BOX regulates the post initiation repression of *STM* genes in leaf primordia. Neither the *STM* expression in SAM nor its down regulation at the leaf initiation sites is regulated by K-BOX (Uchida *et al.*, 2007). The K-box is located 148 bp upstream of the translation initiation codon in *KNB1* and 150 bp upstream of the translation initiation codon in *KNB2*. The length of the core K-box is 23 bp in *Begonia* as it is in most other *STM* like promoters (Figure 56-57).

Figure 53 Phylogenetic relationship of *KNOX* genes.
Jukes-Cantor Neighbour Joining tree based on amino acid sequences is built in Geneious and rooted with *Acetabularia acetabulum KNOX* genes.

Figure 54 Phylogenetic relationship of *STM* like genes.

Jukes-Cantor Neighbour Joining tree based on nucleotide sequences of class 1 *KNOX* genes is also built in Geneious and rooted with the class 2 *KNOX* gene, *KNAT4* of *Arabidopsis*. *Begonia KNB* genes form a clade indicating that the *KNB* duplication is *Begonia* specific.

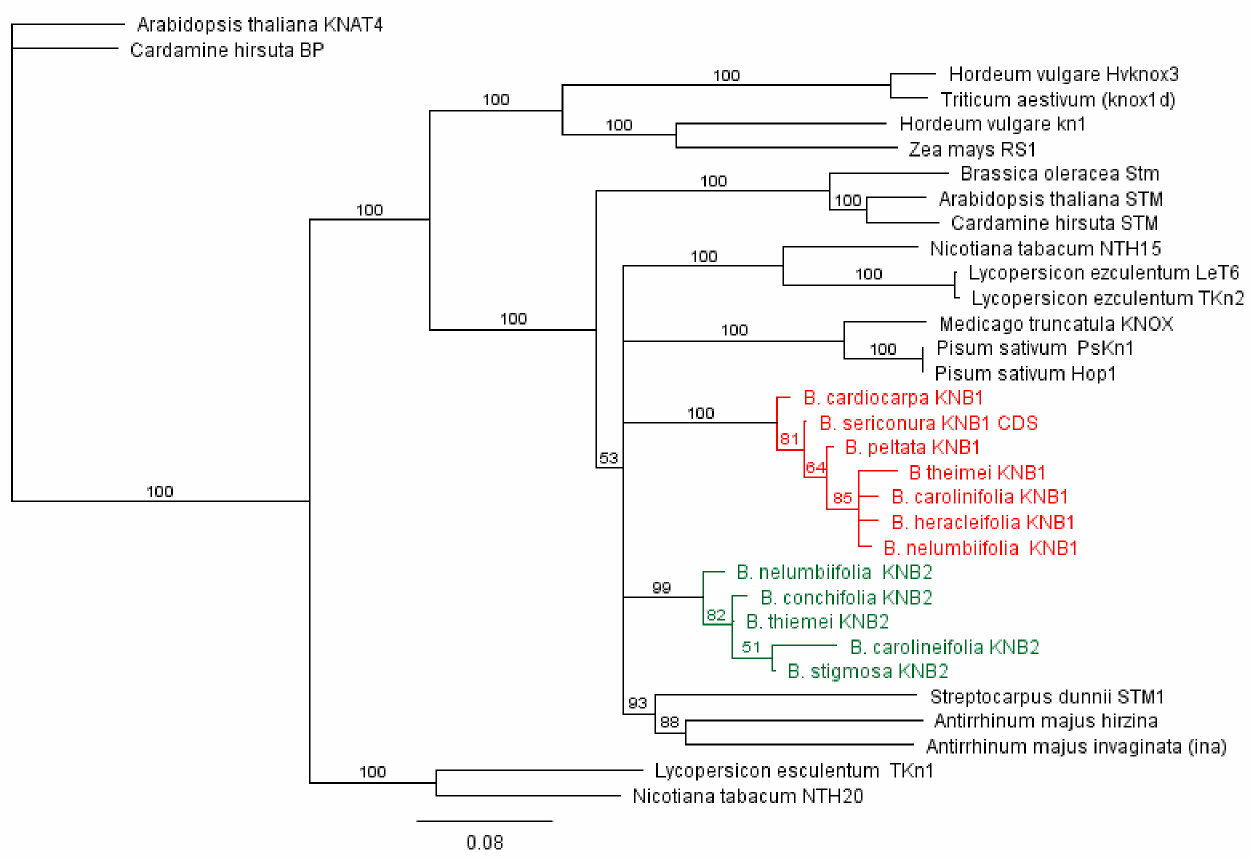


Figure 55 Phylogenetic relationship of Homeodomain of *KNOX* genes.
The Homeodomain is highly conserved between different *KNOX* genes. A tree based on Homeodomain protein sequences is built in Geneious using *Acetabularia acetabulum* KNOX proteins as an out group.

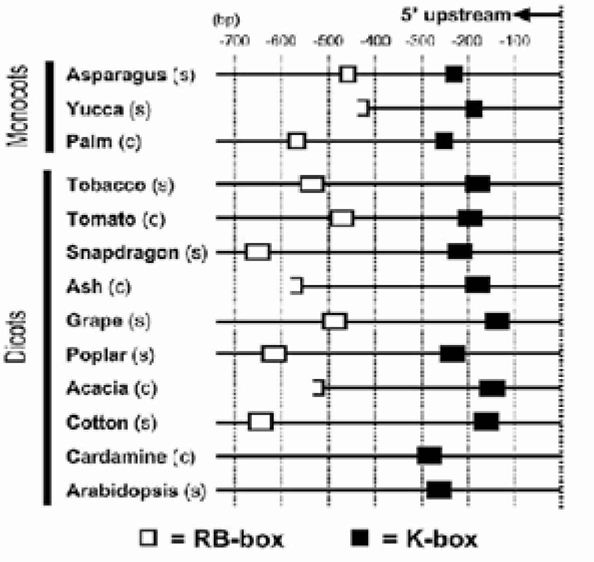
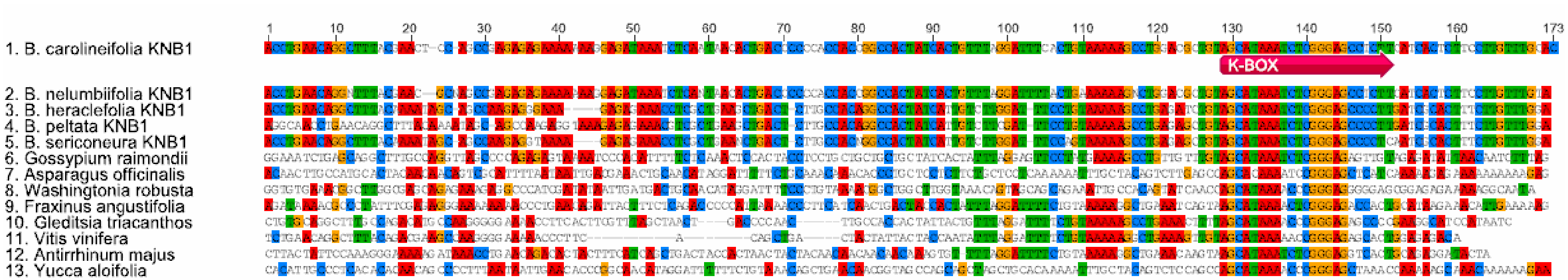


Figure 56 The schematic map of *STM* like genes alignment.
The core K-box (23 bp) is located in the region between 105 and 271 nucleotides upstr;eam from the translation initiation codon (Uchida *et al*., 2007).

Figure 57 The K-BOX alignment from *STM* gene's promoters.
The K-box is located 148 bp upstream of the translation initiation codon in *KNB1* and 150 bp upstream of the translation initiation codon in *KNB2*. The length of the core K-box is 23 bp in *Begonia* as it is in most other *STM* like promoters.



4.17. Rates of evolution for class 1 *KNOX* genes

I determined the selective pressure for different domain of class 1 *KNOX* genes from 67 different species and for *KNB1* and *KNB2* genes separately. I also used the codon based Z-Test of selection in MEGA (Molecular Evolutionary Genetics Analysis) (table 20 and 21). I calculated the ratio of synonymous substitution to non synonymous substitution for different domains of 67 different class 1 *KNOX* genes including *KNB1* and *KNB2.*

All three domains MEINOX, the ELK and HOMEODOMAIN of class1 *KNOX* genes appear to be under purifying or neutral selection. The ELK and HOMEODOMAIN of *KNB1* genes are under positive selection. All domains except the ELK domain of *KNB2* are under purifying or neutral selection. The ELK domain is under positive selection.

Table 20 Tests for selection on Class1 *KNOX* genes for whole CDS

	Class 1 *KNOX* genes
dS	0.805
dN	0.229
dS/dN	3.51
Test for selection (Codon based Z test for Positive selection (HA dN > dS))	
P-Value	1.00

The number of synonymous substitutions per synonymous site and the number of nonsynonymous substitutions per nonsynonymous site from averaging over all sequence pairs are shown. This was obtained by a bootstrap procedure (1000 replicates). Analyses were conducted using the Nei-Gojobori model. The probability of rejecting the null hypothesis of strict-neutrality (dN = dS) in favor of the alternative hypothesis (dN > dS) is shown. Values of P less than 0.05 are considered significant at the 5% level. The analysis involved 67 nucleotide sequences. All ambiguous positions were removed for each sequence pair. There were a total of 132 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Table 21 Tests of selection on *KNOX* genes

		MEIKNOX Domain	**ELK Domain**	**HOMEODOMAIN**
dS	Class1 *KNOX*	1.189	0.581	0.117
	KNB1 & *KNB2*	0.337	0.374	0.592
	KNB1	0.051	0.00	0.00
	KNB2	0.021	0.146	0.134
dN	Class1 *KNOX*	0.180	0.206	0.208
	KNB1 & *KNB2*	0.048	0.099	0.023
	KNB1	0.017	0.048	0.601
	KNB2	0.00	0.012	0
dS/dN	Class1 *KNOX*	6.60	2.82	0.562
	KNB1 & *KNB2*	7.02	3.77	25.7
	KNB1	3.00	0	0
	KNB2	Not valid	12.16	Not valid
Test for selection (Codon based Z test for Positive selection (HA dN > dS))				
P-Value	Class1 *KNOX*	1.00	1.00	0.087
	KNB1 & *KNB2*	1.00	1.00	1.00
	KNB1	1.00	0.028	0.068
	KNB2	1.00	1.00	1.00

The number of synonymous substitutions per synonymous site and the number of nonsynonymous substitutions per nonsynonymous site from averaging over all sequence pairs are shown. This was obtained by a bootstrap procedure (1000 replicates). Analyses were conducted using the Nei-Gojobori model. The probability of rejecting the null hypothesis of strict-neutrality (dN = dS) in favor of the alternative hypothesis (dN > dS) is shown. Values of P less than 0.05 are considered significant at the 5% level. Evolutionary analyses were conducted in MEGA5.The analysis involved 67 nucleotide sequences for calss1 *KNOX* genes,6-11 for *KNB1* and *KNB2*, 6 for *KNB1* and 2-5 for *KNB2*. All ambiguous positions were removed for each sequence pair.

4.18. Discussion

4.18.1. Evolution of *BARP* genes and leaf form in *Begonia* section Gireoudia

ARP like genes were isolated from *Begonia*. Southern blotting and sequence analysis supported the presence of two copies of *BARP* (*Begonia ARP*) genes. *ARP* genes are transcription factors and consist of three domains each of which is responsible for a specific aspect of gene function i.e. Myb domain 1 and Myb domain 2 are involved in nucleic acid binding and the C-terminal domain is linked with homodimerization. The second copy of *BARP* genes has 18 bp deletions in Myb domain 2 and this domain is under adaptive pressure (positive selection) in Angiosperm and also in *Begonia BARP1* genes. This suggests that duplication of *BARP* genes may have resulted in sub functionalization or neofunctionalization in *Begonia*.

At the protein level the 3D structure of *BARP1* overlaps with that of the *Asymmetric leaves 1* (*AS1*) protein of *Arabidopsis* while *BARP2* showed disagreements with the *AS1* protein structure in Myb domain 2. *BARP2* may bind to different sequence to *BARP1* or it may not bind at all and act to affect the function of *BARP1*. Expression and functional analysis is required to determine the role of the second *BARP* genes.

BARP genes show high similarity between *Begonia* species and neither the CDS nor 5′ variation of the gene shows any link with leaf form variation in this section. The *BARP* gene always formed a monophyletic group with other species' *ARP* genes and closely resembles the *Cucumis sativus ARP* genes. Phylogenetic analysis suggested that duplication in *BARP* genes is recent and has happened after the divergence of *Begonia*ceae from Cucurbitaceae. It would be interesting to isolate *ARP* genes in the only other genus of the *Begonia*ceae; *Hillebrandia,* to confirm whether the duplication is specific to genus *Begonia* species. A limitation of doing this was lack of a *Hillebrandia* specimen. Seed is now germinating at RBGE, so this could be done next year.

4.18.2. Evolution of *KNOX* genes and compoundness in *Begonia* section Gireoudia

STM-like genes were isolated from *Begonia* section Gireoudia. By homology these genes appeared to fall into two classes named as *KNB1* (*KNOX* in *Begonia* 1) and *KNB2* (*KNOX* in *Begonia* 2) and both are orthologues of *STM*. Sequences of *KNB1* and *KNB2* showed high homology with each other as compared to any other species *STM* like genes indicating that duplication of this gene has happened in the *Begonia* lineage. The test of selective pressure has indicated that all domains of *KNB1* and *KNB2* are under purifying selection except the ELK domain of *KNB2* genes. The MEIKNOX domain of class1 *KNOX* genes is responsible for protein-protein interaction with BELL-like transcription factors and the Homeodomain binds DNA. The MEIKNOX domain of *LeT6* (*STM* orthologue of Tomato) competes with the MEIKNOX domain of *PTS* (another *KNOX*- like gene) to interact with BELL like proteins (Kimura *et al*., 2008) to modulate leaf form, showing the importance of the MEIKNOX domain for carrying out the usual class 1 *KNOX* genes functions. Any mutation in this domain will be deleterious and should be selected against. This is consistant with the domain being under purifying selection. The ELK domain controls nuclear localization. It has been shown that KNOX proteins are capable of moving between cells via plasmodesmata (Kim *et al*., 2003) and this movement of KNOX proteins is important for controlling meristem development. The fact that the ELK domain is under positive selection indicates that *KNOX* genes may be regulating phenotype diversity by protein movement.

In the phylogenetic tree *KNB* sequences are not grouped on the basis of leaf morphology in *Begonia* and no amino acid polymorphisms are unique to compound leaved species. This suggests that the variation in leaf form in section Gireoudia is not due to changes in the coding sequences.

The promoter regions of *KNB1* and *KNB2* were amplified from *Begonia* species representing different leaf forms. The K-box regulates the expression of *STM* like genes in various simple leaved species (Uchida *et al*., 2009). I looked to see if variation at the K-Box or surrounding regions was associated with leaf form variation in section Gireoudia. The promoter regions of *KNB1* and *KNB2* are highly conserved and no common motifs specific to simple, peltate and compound leaved species were found. I amplified up to 750 bp upstream of the *KNB* CDS and more promoter regions could be amplified in order to search for common motifs extensively but time is a big limitation for doing this.

Chapter 5

Functional characterization of *BARP1* and *BARP2* genes in *Arabidopsis thaliana*

5.1. Introduction

There are numerous examples showing that hopeful monsters might have contributed to evolution by mutations in key regulatory genes. The model for such genes is the homeobox genes which play a key role in the specification of the animal body plan in both development and evolution. For example the *Pax*-6 *Hox* gene encoding a paired-box and homeobox, from mice (which triggers eye formation) can induce the formation of fly eyes all over the body, even on the wings in the fruitfly *Drosophila* (Brakefield and French, 1995; Quiring *et al*., 1994; Halder *et al*., 1995a; Walther and Gruss., 1991). Remarkably, the *Pax*-6 helps to organize compound eyes in flies and camera eyes in both squid and vertebrates (Bonini *et al*., 1997). Mutations in *Pax*-6 gene, *Small eye* (*Sey*) result in a reduction of external eye size, a characteristic iris hypoplasia, and at later stages, corneal opacification and cataracts in mouse (Hogan *et al*. 1986). Aniridia in man has also been shown to be caused by heterozygous mutations of *Pax*-6 and is characterized by a varying degree of iris hypoplasia, corneal opacification, cataracts, and glaucoma (Jordan *et al*., 1992; Glasser *et al*., 1992). The *Hox* gene, *tinman*, induces heart formation in both insects and vertebrates (Bodmer., 1993; Harvey., 1996; Komuro and Izumo, 1993; Lints *et al.,* 1993; Evans *et al*., 1995). *Distal-less* controls the development of fly legs, fish fins and the tube feet of sea urchins (Cohen *et al*., 1989; Akimenko *et al*., 1994; Holland *et al*., 1996).

Interspecific transgenic experiments involve the moving of a single gene (whose function needs to be tested) from a donor species to replace the function of the endogenous gene in a recipient species. If there is divergence in gene regulation or protein function then this orthologue would yield phenotypes not normally seen in the recipient species. For example species-specific *Yellow* expression patterns were retained in *D. melanogaster* transformants carrying the *D. subobscura* and *D. virilis yellow* genes, indicating that sequence evolution within the *yellow* gene underlies the divergence of *Yellow* expression (Wittkopp *et al.* 2002). The courtship songs are generated by *Drosophila* males by modulating the interpulse intervals which facilitate mating. These interpulse intervals are species specific and the clock gene *period* was tested for its possible involvement in this mechanism. The germline transformation experiments between two *Drosophila* species showed that four base pairs substitution in *period* gene of one species was responsible for species specific modulation of the interpulse intervals (Wheeler *et al.* 1991).

However when a protein from one species can complement a mutant or produce a similar phenotype in a second organism, even when the two species have been separated for long evolutionary periods, this suggests conservation in the function of proteins. One of the most excellent examples is the ability of the mouse *Small-eye* (*Pax-6*) gene, which controls eye formation, to induce ectopic eye formation in *Drosophila*, indicating that the function of this protein has been conserved for the 500 million years since the divergence of arthropods and vertebrates (Halder *et al*., 1995). In another example, *Drosophila Hox* mutant (*lab*) phenotypes can be rescued by the chicken *gHoxb-1* gene which plays a role in head involution (Lutz *et al*., 1996). When the promoter of the *gooseberry* (*gsb*) gene was tagged with the coding region of the *paired* (*prd*) gene, it rescued the *gooseberry* (*gsb*) mutant phenotypes which showed the conservation of proteins despite the fact that the coding sequences of both genes are highly diverged and both genes play different roles in *Drosophila* development (Li and Noll, 1994). In plants the *cdc2* homolog of maize can rescue the *cdc28* mutant of *Saccharomyces*, indicating that the function of the proteins has been conserved over long evolutionary periods as these diverged approximately 600 million years ago (Colasanti *et al*., 1991).

The direct transformation of *Arabidopsis thaliana* with genes from other species has provided evidence to support a role for *KNOX* genes in the evolution of leaf morphology (Hay and Tsiantis, 2006). Two class I *KNOX* homologues

(*CRKNOX1* and *CRKNOX2*) from *Ceratopteris richardii* were expressed in the SAM and in the incipient leaf primordia, as was observed for another fern, *Anogramma charophylla* (Bharathan *et al*,. 2002; Sano *et al*,. 2005). In order to reveal the function of *CRKNOX1* and *CRKNOX2*, over expression experiments of these genes were conducted in *Arabidopsis*. The resulting phenotypes resembled the *KNAT1* (class1 *KNOX* genes) over expression phenotype, suggesting that these proteins can function similarly to their angiosperm counterparts in meristem development and leaf architecture (Sano *et al*. 2005).

Expression of *C. hirsuta AS1* under the control of the broadly expressed CaMV 35S promoter complemented the *A. thaliana as1* mutant phenotype and repressed expression of the *KNOX* gene in *AS1* leaves, indicating that the function of the two proteins is conserved. The function of either *LjPHANa* or *LjPHANb* (*Lotus japonicus ARP* genes orthologs) was the same as other *ARP* genes orthologs at the protein level as the transgenic plants typically displayed elongated leaves with narrower, longer blades and longer petioles as compared with the leaves of wild type Landsberg *erecta* (Luo *et al*., 2005). These transgenic phenotypes mimicked those of a previous study, in which the overexpression of *RS2* (maize *ARP* genes ortholog) and *AS1* in *Arabidopsis* caused narrower leaves with longer petioles (Theodoris *et al*., 2003). *SkARP1* (*Selaginella kraussiana ARP* genes ortholog) was also functionally equivalent to eudicot *ARP* genes as it complemented *as1* leaf phenotypes and repressed the *KNOX* genes in *Arabidopsis* (Harrison *et al*, 2005).

Interspecies transformation studies have both strengths and limitations when used to study the genetic basis of species differences. On the positive side, studying genes from different species can reveal conserved mechanisms that are hard to detect in traditional mutant or reverse-genetic studies (Harrison *et al*, 2005; Luo *et al*., 2005; Hay and Tsiantis, 2006; Theodoris *et al*., 2003; Yoon and Baum, 2004). Conversely transgenic experiments that do not yield perfect conversion of the recipient species into the phenotype of the donor species makes the interpretation more complex as it may provide neither rigorous rejection nor support of a prior hypothesis (Sliwinski *et al*., 2007). One explanation for such partial complementation can be that mostly proteins function in a network of genes where they perform a specific function through interactions with lots of other downstream genes and cofactors. A change in these downstream targets can alters the binding capability of the upstream protein leaving it only a part of its original function or the protein can acquire a totally different role. In such a scenario the original role of the protein is either fulfilled by some other mechanism or through gene duplication. Such a protein will complement the original function in a second species where there has been no change in downstream targets.

LFY plays a role in the evolution of plant architecture as introduction of the *L. crassa LFY* gene into an *A. thaliana lfy*-6 background rescued the production of petals and stamens (Yoon and Baum, 2004). In Angiosperms *LEAFY* is responsible for flower induction whereas it regulates different aspects of the life cycle in a basal land plant, the moss *Physcomitrella patens* (Maizel *et al*. 2005). The DNA binding domain of *LEAFY* is strongly conserved indicating the possibility of conservation of the molecular function as well. However when moss genes *PpLFY1* and *PpLFY2* cDNAs were linked to the *Arabidopsis LFY* promoter and introduced into a strong *lfy* mutant they were inactive. This indicates that *LFY* function has diverged during evolution. However Angiosperm LFY genes fully complement the *lfy* mutant, whereas gymnosperm genes provide only partial rescue. Among homologs from the most basal groups, fern genes have some rescue ability, although less than the gymnosperm genes. This gradient of complementation reflects the phylogenetic distance from angiosperms and suggests that a continuum of discrete and nonneutral changes, rather than a sudden modification, is responsible for changes in function (Maizel *et al*. 2005).

The major source of new genes and functions is gene duplication. Until recently, it was generally assumed that duplicate genes were free to evolve new functions (‘neofunctionalization’) because the original function was maintained by the other copy (Taylor and Raes, 2004). However, several recent case studies and comparisons of genome content have suggested that most new genes do not have novel functions (Prince and Pickett, 2002). Instead, paralogous gene pairs are often ‘subfunctionalized’ with two or more functions being partially or completely subdivided between the two genes after gene duplication.

Work in the previous chapter described the isolation of two *ARP* like genes from *Begonia luxurians*. To check whether expression of these two *Begonia ARP* genes (*BARP1* and *BARP2*) could complement the *as1* phenotype and deduce the phenotypes of *BARP1* and *BARP2* overexpression on leaves, I generated *Arabidopsis* plants expressing *AS1*, modified *AS1* (*BARP2* like *AS1*) and *BARP1* from 35S promoter of cauliflower mosaic virus (35S CaMV). 35S CaMV is a strong constitutive promoter which causes a high level of expression in most parts of transformed plants irrespective of developmental and environmental factors. The rationale for theses experiments was that if *BARP1* and *BARP2* are functionally equivalent to other *ARP* genes than these should rescue *Arabidopsis as1* mutant phenotypes and over expression should generate plants with narrower leaves and longer petioles in a wild type background (Theodoris *et al*., 2003; Luo *et al*., 2005). However if the duplicated *BARP* genes has undergone subfunctionalization, neither of the *BARP* genes copy should fully complement *as1* mutants alone. In case of neofunctionalisation one of the copies should rescue the mutant phenotype and the other copy, which could have acquired new functions, should not complement.

BARP1 and *BARP2* are different in Myb domain 2; *BARP2* has 6 amino acid deletions in this domain. Theodoris (Theodoris *et al*., 2003) performed deletion experiments to reveal the importance of Myb domain for *RS2* gene function. They introduced constructs with *RS2* coding sequences having the Myb domain deleted (*RS2* △Myb) in *as1* mutant and wild type *Arabidopsis* plants. These constructs partially rescued the *as1* mutant phenotypes and over expressed lines generated unexpected dominant phenotypes with narrower leaves than over expressed *AS1* or *RS2* lines. The *RS2* △Myb over expressed plants showed leaf lobes as seen in *as1* mutants though less conspicuous and thses plants also had a novel phenotype where petiole/leaf-blade boundary was more diffused (Theodoris *et al*., 2003). On the bsis of these experiments authors concluded that interaction between Myb domain and other factors are important for *RS2* function in leaf shape regulation. If 6 amino acid deletion of Myb domain2 is important for *ARP* gene function, I expect to see the similar phenotypes with modified *AS1* in as1 mutant and wild type back ground as were observed by Theodoris which may indicate that *BARP2* is only partially active.

5.2. Experimental Design

5.2.1. Cloning of *AS1* and *BARP1* genes

ASYMMETRIC LEAVES1 (*AS1*) cloned in pBluescript vector (named pPOD#12) was kindly provided by Professor Andrew Hudson. I confirmed the presence of the insert with double digestion (*Afl*II *and Hind*III) as shown in Figure 58. The *AS1* gene was recloned into pGEM T easy vector system (named pSU68) as the pBluescript vector provides fewer restriction enzyme choices for later manipulation of the vector. The insert was confirmed with colony PCR followed by sequencing. Complete *AS1* and *BARP1* CDS were amplified with primers having adapters for directional pENTR D-TOPO cloning as described in 2.4 from *Arabidopsis* and *B. luxurians* respectively. PCR products of *AS1* and *BARP1* were column purified and cloned into pENTR D-TOPO vector (Figure 59 and Figure 60) and plasmids were named pSU64 and pSU65 respectively (Figure 61 A & B).

Restriction analysis was done to select the desired clones (Figure 61C & D) for sequencing with M13 forward and reverse primers. DNA sequences were aligned with already sequenced *AS1* and *BARP1* and 100% identity was found.

LR recombination reactions were carried out with Gateway® LR Clonase™ II Enzyme mix using 150 ng of destination vector (pB7WG2.0) and 150 ng of entry clone (pSU65.16/pSU64.3) as described in chapter 2. The resultant vectors were named pSU69 and pSU70 carrying *AS1* and *BARP1* genes respectively. The inserts were confirmed by sequencing using gene specific primers. The flow chart of the LR reaction is shown in Figure 62.

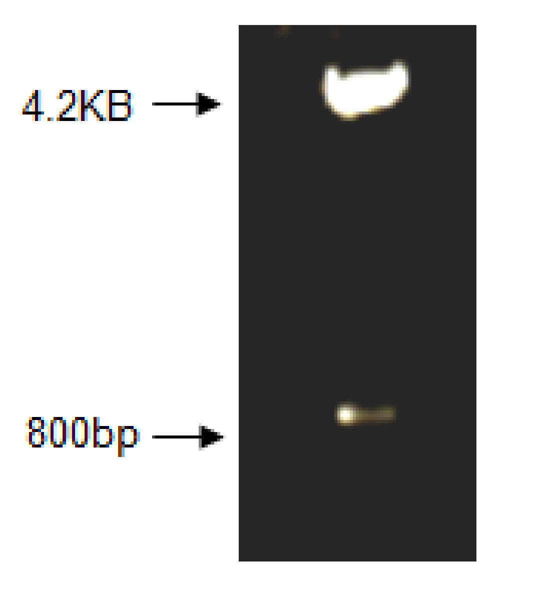


Figure 58 pPOD#12 digested with *Afl*II and *Hind*III. The expected band sizes of 800bp and 4.2KB were observed.

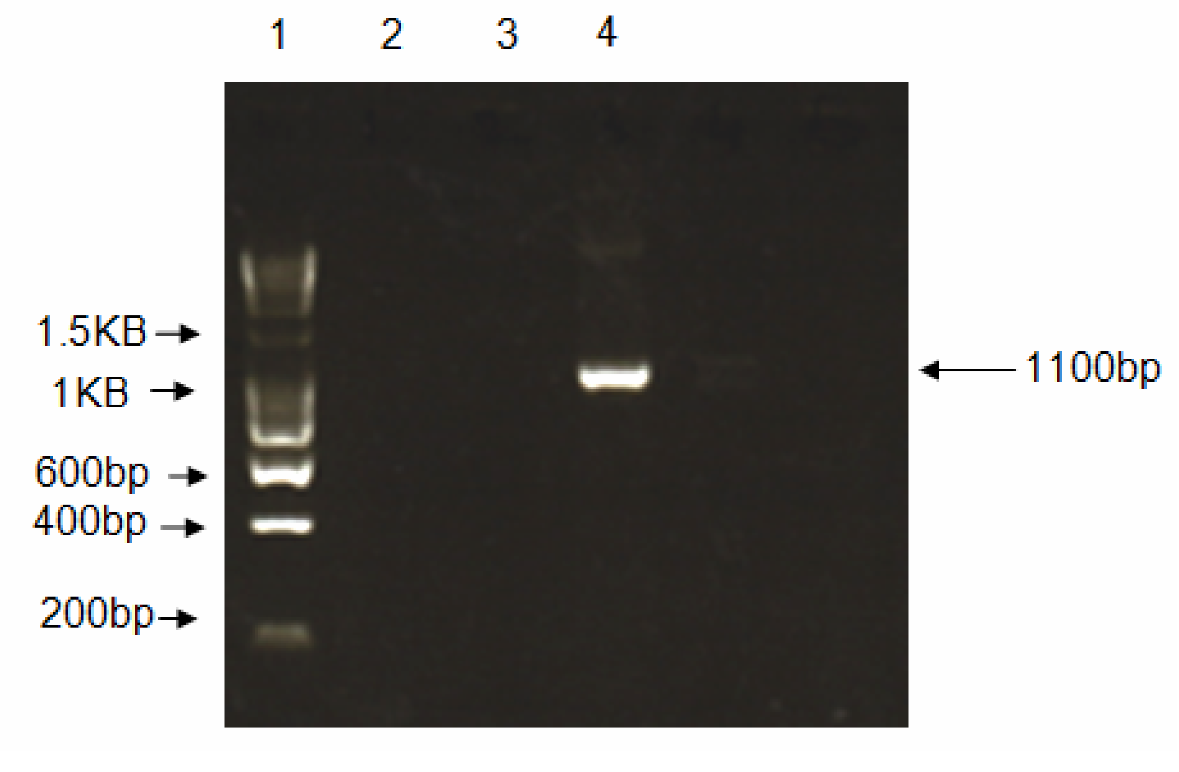


Figure 59 PCR amplification of *AS1* using high fidelity *Taq* polymerase (A) Lane 1 Hyper ladder 1, lane 4 PCR amplified *AS1*

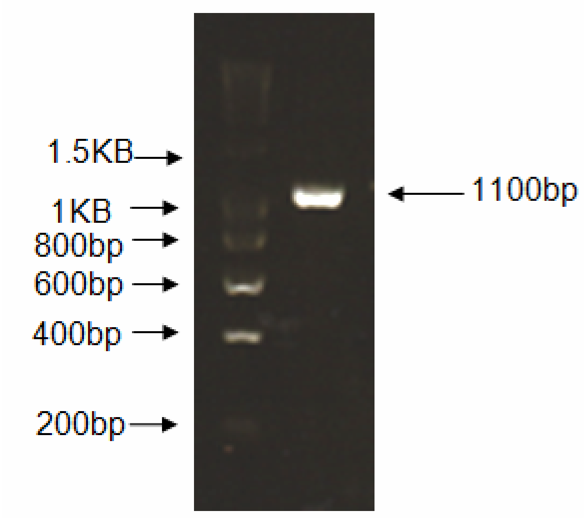


Figure 60 PCR amplification of *BARP1* using high fidelity *Taq* polymerase; Lane 1 Hyper ladder 1, Lane 2 PCR amplified *BARP1*.

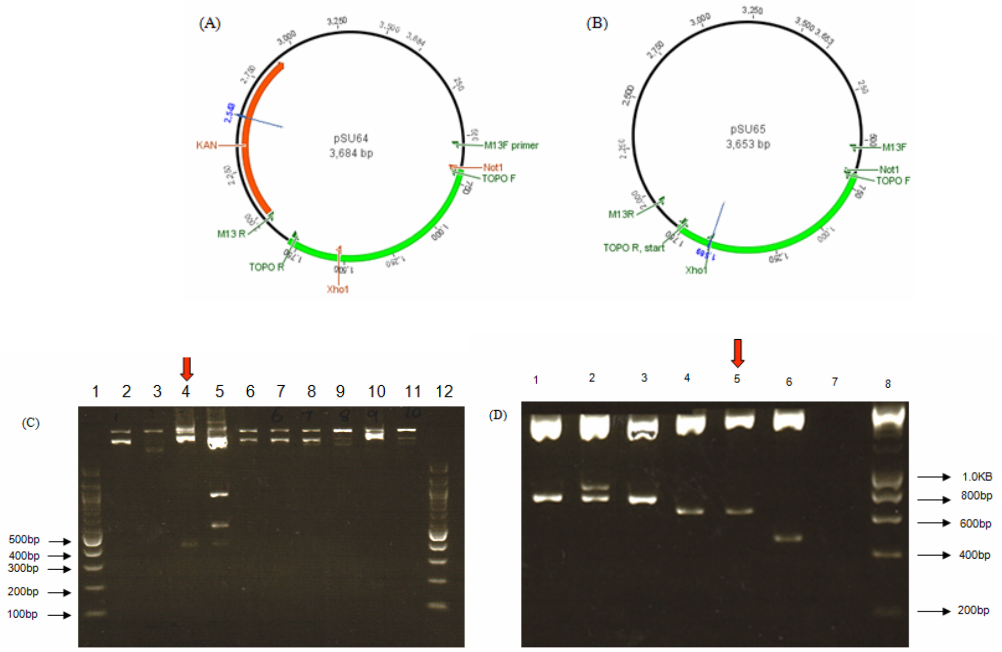


Figure 61 (A) Physical map of pSU64 (pENTR D-TOPO with *AS1*); (B) Physical map of pSU65 (pENTR D-TOPO with *BARP1*);

(C) Lane 1 100bp ladder, lane 2-11 digested products of pSU64.1-pSU64.11 with *Xho*1 and *Not*1, pSU64.3 is a positive clone indicated by red arrow, lane 12 100bp ladder; (D) lane1-6 digested products of pSU65.11-pSU65.16 with *Xho*1 and *Not*1, Lane 5 pSU65.16 is a positive clone indicated by red arrow.

5.2.2. Construction of modified *AS1*

As described in 4.4 the difference between *BARP1* and *BARP2* is an 18 bp deletion (six amino acids) in the Myb domain2 of *BARP2.* It was hard to get the complete exon1 of *BARP2* from *Begonia* because the *BARP1* and *BARP2* are highly similar to each other in Myb domain1 and C-terminal domains. In order to characterize the function of *BARP2* from *Begonia* I modified *AS1* from *Arabidopsis* by deleting the same 18 bp from the Myb domain2 of *AS1*. A 250 bp fragment of *AS1* gene with 18 bp deletion was synthesized by Gene Synthesizer Company DNA 2.0 (Figure 63). The plasmid was rescued from the filter. The yield of plasmid was very low and plasmid was retransformed into TOP 10 cells and named pSU66. Presence of the insert was confirmed through restriction analysis followed by sequencing. I tried to isolate the modified *AS1* gene fragment from pSU66 through double digestion (*Afl*II and *Hind*III) but yield was very low. So I carried out the Infusion PCR reaction. The modified *AS1* fragment was amplified from pSU66 with Infusion F and R primers (Figure 64). pSU68 (pGEM T easy vector with *AS1*) was digested with *Afl*II and *Xho*I enzymes (Figure 65 A, B). The Infusion reaction was carried out as described in chapter 2 and the resulting plasmid was named pSU80 (Figure 66 A). Colony PCR followed by sequencing was used to select a positive clone (Figure 66 B).

Figure 62 LR reaction facilitates recombination of an *att*L substrate (entry clone) with an *att*R substrate (destination vector) to create an *att*B-containing expression clone (see diagram below). This reaction is catalyzed by LR Clonase™ II enzyme mix.

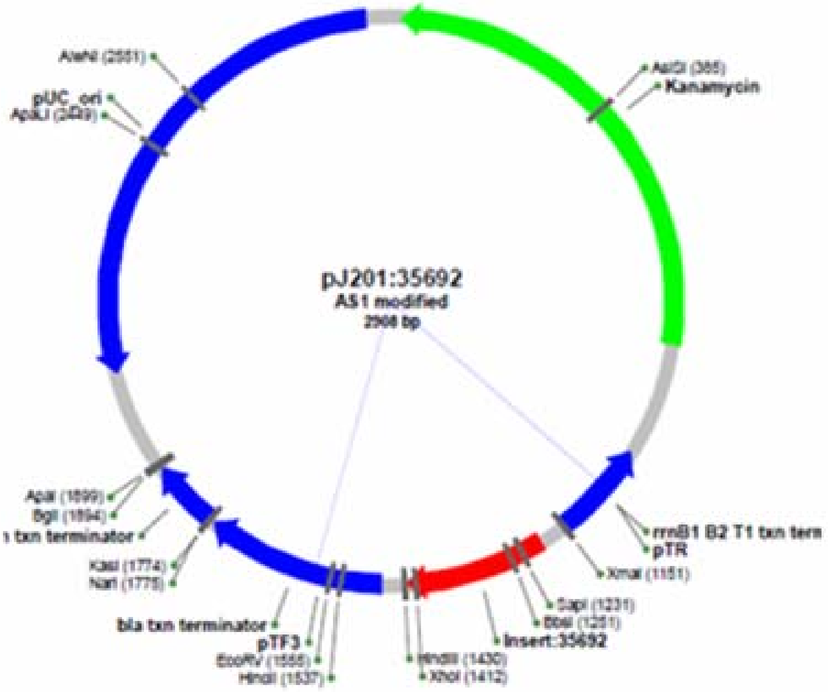


Figure 63 Physical map of plasmid pJ201 sent by DNA 2.0.
The modified *AS1* gene fragment is flanked by *Afl*II and *Xho*I sites as shown in red

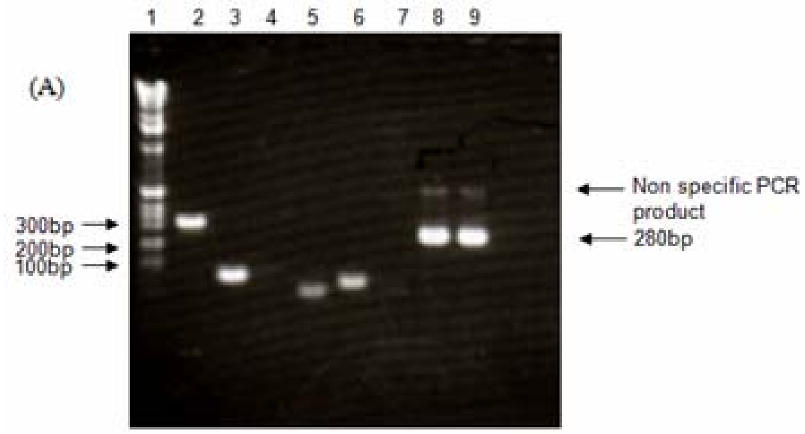


Figure 64 Infusion PCR on pSU66, Lane 1 100 bp ladder, lane 8 & 9 infusion PCR products of modified *AS1* with 15 bp complementary nucleotides of pSU68 on both sides to be used in infusion reaction.

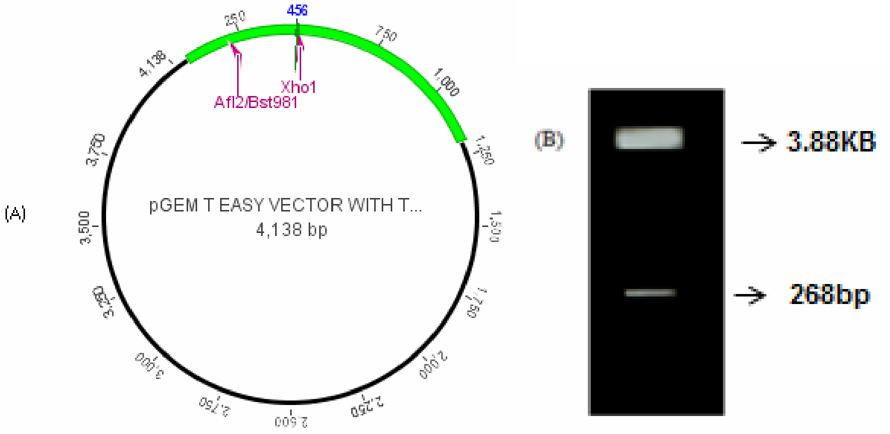


Figure 65 (A) Physical map of pSU68 carrying normal *AS1* gene. (B) Digestion of pSU68.5 with *Afl*II and *Xho*I yielded 268bp and 3.88KB fragments. 3.88KB fragment was used for infusion PCR.

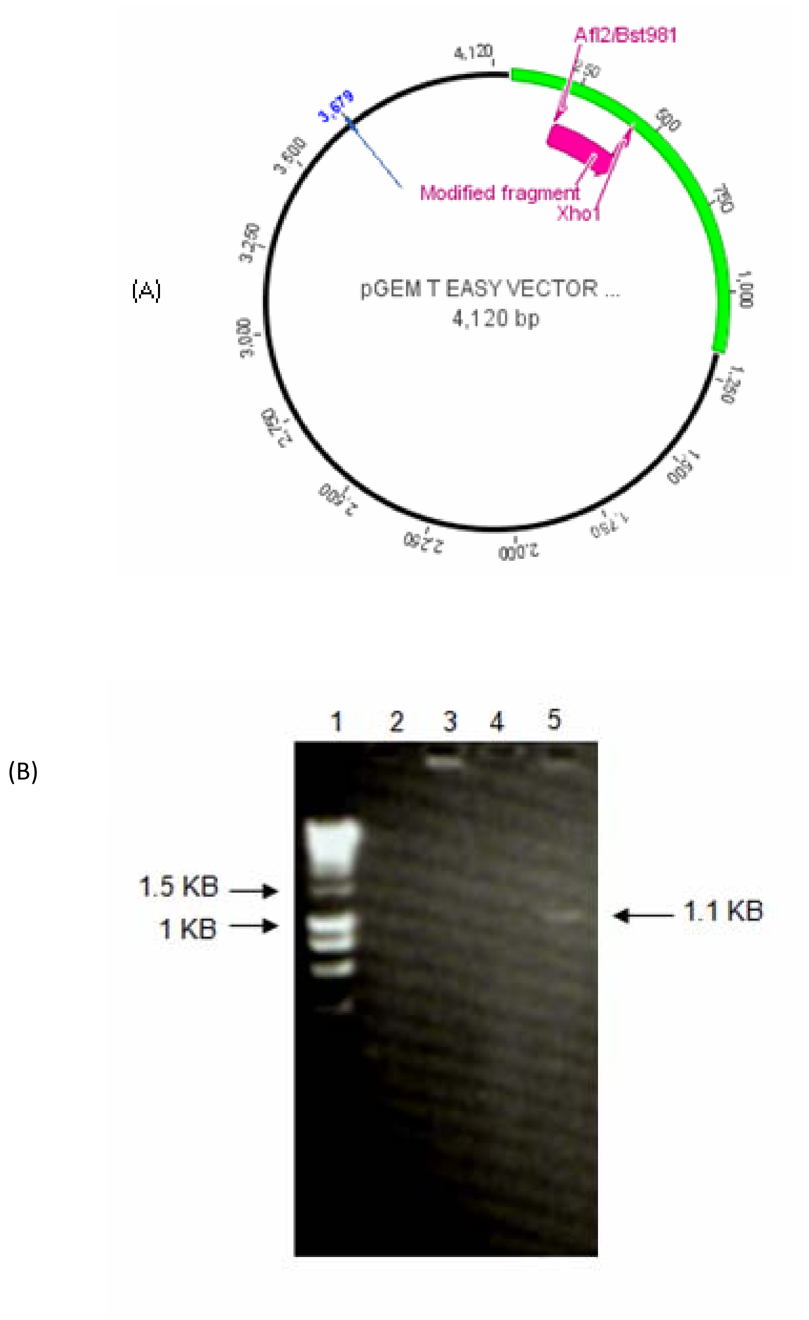


Figure 66 Physical map of pSU80 carrying modified *AS1* gene. (B) Colony PCR for pSU80, Lane 1 100 bp ladder, Lane2-5 pSU80.1-pSU80.4, Lane 5, pSU80.4 is a positive clone.

Infusion PCR was carried out to transfer modified *AS1* from pGEM Teasy vector to pENTR D-TOPO vector. The Infusion804F and Infusion 804R primers were used to amplify modified *AS1* from pGEM Teasy vector (Figure 67), and pSU64.3 was digested with *Xho*I and *Not*I enzymes (Figure 68) to perform the infusion reaction. The infusion reaction was carried out as described in chapter 2 and positive clone was selected with colony PCR (Figure 69) for sequencing with M13F and R primers.

LR recombination reactions were carried out with Gateway® LR Clonase ™ II Enzyme mix using 150 ng of destination vector (pB7WG2.0) and 150 ng of entry clone (pSU82.1) as described in chapter 2. The resultant vector was named pSU95 carrying modified *AS1* gene. The insert was confirmed through sequencing using gene specific primers.

pSU69, pSU70 and pSU95 were transformed into *Agrobacterium* strain GV3101 as described in chapter 2 and resultant vectors were named as pSU96, pSU97 and pSU98 respectively. The transformation of *Arabidopsis* was done according to the floral dipping method with pSU96, pSU97 and pSU98 as described in chapter 2. *Arabidopsis* transformants were selected on ½ MS plates with 50µg/ml kanamycin as described in chapter 2. Kanamycin resistant plants with true leaves and extended root system were than transferred into 9 cm pots and phenotypes of plants were evaluated.

Figure 67 Infusion PCR fragment of modified *AS1* from pSU80, Lane 1 Hyper ladder 1, lane 2 Infusion PCR product.

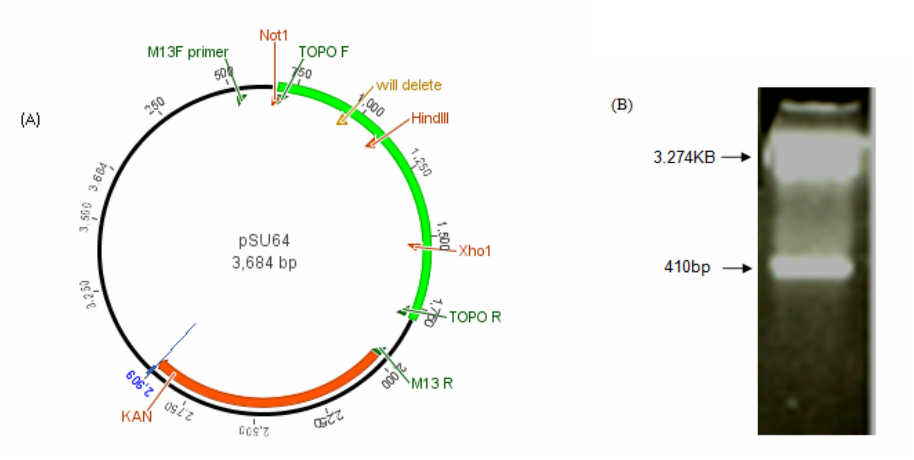


Figure 68 (A)Physical map of pSU64 pENTR D TOPO vector with *AS1* gene. (B) It was digested with *Not*I and *Xho*I for using in infusion reaction.

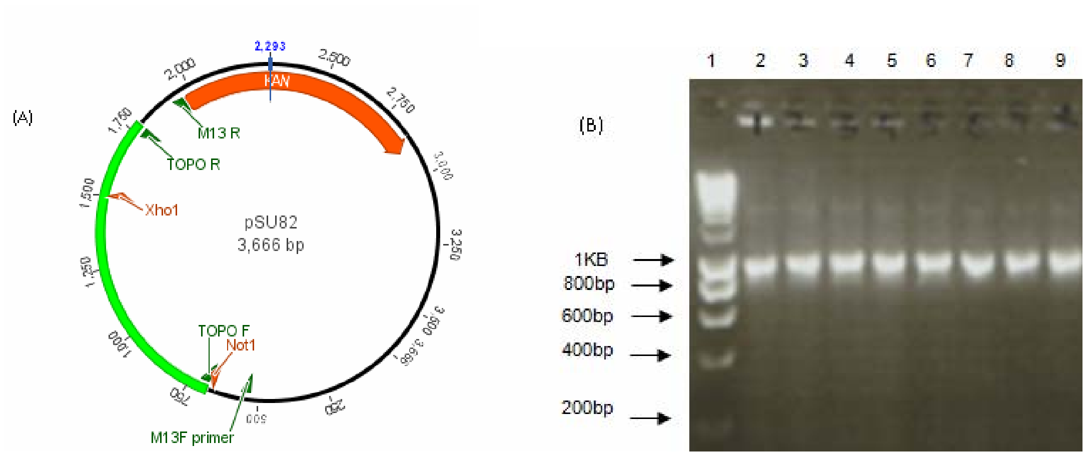


Figure 69 Physical map of pSU82; pENTR D TOPO vector with modified *AS1* gene (B) Colony PCR for pSU82, Lane 1 Hyper ladder1, Lane 2-9 pSU82.1-pSU82.8.

5.3. Results

5.3.1. 35S *AS1*/*BARP1*/modified *AS1* phenotypes in *as1* mutant background

The development of leaves is disrupted in *asymmetric leaves1* (*as1*) mutants (Figure 70). The mutant leaves have increased width to length ratio (shorter petiole and wider leaves). The *as1* lamina has prominent lobes. Later leaves have more lobes as compared to the early leaves (Byrne *et al*., 2000). The *as1* mutant plants transformed with 35S *AS1*, *BARP1* and modified *AS1* were selected on Kanamycin plates. PCR genotyping or northern blotting was not done due to time constraints. Expression of *AS1* and *BARP1* in an *as1* mutants were able to complement the *as1* leaf development phenotypes (compare Figure 71-72 with 70) suggesting that *BARP1* and *AS1* are functionally interchangeable and that this pathway is functionally conserved. Compared with wild-type and *as1* controls, the phenotype of transformed plants varied from *as1*- like to wild type like plants. In 35S *SkARP1 Arabidopsis* plants the degree of phenotypic rescue was directly proportional to *SkARP1* transcript levels (Harrison *et al*, 2005). The difference in *BARP1* and modified *AS1* gene transcripts level may be the cause of variation in rescued phenotypes. This can be tested by performing Quantitative PCR or northern blotting which were not done due to time constraints. The ratios for rescued to non rescued plants were 2:1 and 1:1 for *AS1* and *BARP1* transgenes respectively as described in table 22. *AS1* has rescued *as1* mutants at higher frequency as compare to *BARP1* which may be due to the fact that *AS1* is expressing in its native species. The modified *AS1* rescued *as1* mutants phenotypes partially (compare Figure 73 with 70 and 74). The plants were more similar to wild type spatulate (having broad rounded end) form than rounded *as1* like plants. Furthermore the transformed plants have broader leaves than wild type plants and fewer lobes than *as1* mutant plants. This is consistant with the previous finding where the similar phenotypes were observed with 35S *RS2* with deleted Myb domain in *as1* background (Theodoris *et al*., 2003).

This may suggest that 6 amino acid deletions have made Myb domain of modified *AS1* partially inactive.

Table 22 Scoring of transgenic plants (T1) for leaf form.
***as1* means plants looking like *as1* mutants, o/e is abbreviation for over expressed lines, * indicates plants showing phenotypes similar to over expressing *REV1-D* and ! represents plants lacking apical dominance.**

	Col-0				***as1* Col-0**		
	as1	**Wild**	**o/e**	**Novel phenotypes**	***as1***	**wild**	**Novel phenotypes**
pSU96 **(35S *AS1*)**	**0**	**15**	**46**	**2!**	**15**	**26**	**1!**
pSU97 **(35S *BARP1*)**	**0**	**14**	**25**	**4***	**54**	**55**	**4!**
pSU98 **(35S modified *AS1*)**	**58**	**2**	**0**	**0**	**15**	**32**	**0**

5.3.2. Over expression of *AS1*, *BARP1* and *BARP2* like *AS1* in *Arabidopsis*

Over expression of the genes should increase gene expression in leaves and should produce dominant leaf phenotypes. Whereas loss of *as1* resulted in plants with shorter petioles and wider leaves (increased width to length ratio), 35S expression of either *AS1* or *BARP1* in wild-type plants resulted in the opposite effect, producing elongated leaves with narrower blades and longer petioles (decreased width to length ratio) (compare Figure 75-77 with 74). This is consistant with the previous findings where over-expression of *AS1*, *RS2*, *LjPHANa* and *LjPHANb* under 35S promoter generated the similar phenotypes in wild type background (Theodoris *et al*., 2003; Luo *et al*., 2005).

However when the modified *AS1* was over expressed in wild type plants, it produced *as1* like plants suggesting that modified *AS1* is a negative dominant locus (Figure 78). These phenotypes deviate from previous findings where no *as1* like phenotypes were observed in *RS2* △Myb (*RS2* with myb domain deleted) (Theodoris *et al*., 2003). This may be due to the loss of Myb domain activity to a different extant as compared to *RS2* △Myb constructs where the whole Myb domain was deleted whereas modified *AS1* has only 6 amino acid deletions in Myb domain 2.

I grew the pooled T1 plants seeds and obtained 104 plants. T2 lines generated thirty two wild type plants and seventy two plants showed phenotypes resembling to 35S *RS2* △Myb like plants (Figure 79 and 80) (Theodoris *et al*., 2003). 35S *RS2* △ Myb like plants may be the extreme phenotypes of *as1* like mutant plants and may be dependent on gowth conditions. Both T1 and T2 lines were generated at different times and difference observed in phenotypes may be due to environment as *phan* mutants were temperature sensitive and expressed more at higher temperatures (Waites *et al*., 1995).

5.3.3. Novel Phenotypes

As described in table 22, some novel phenotypes were observed among over expressing transgenic plants with 35S *AS1*, *BARP1* and modified *AS1* in wild type background. Most of these plants have an upward curvature of leaves, resulting from excessive growth of the lower side. This could be due to the change in stability of dorsoventral polarity (Waites and Hudson, 1995). Some of the transgenic lines have filamentous leaves. Some of these plants showed lack of apical dominance (Figure 81).

Figure 70 Phenotype of the *as1* mutant.

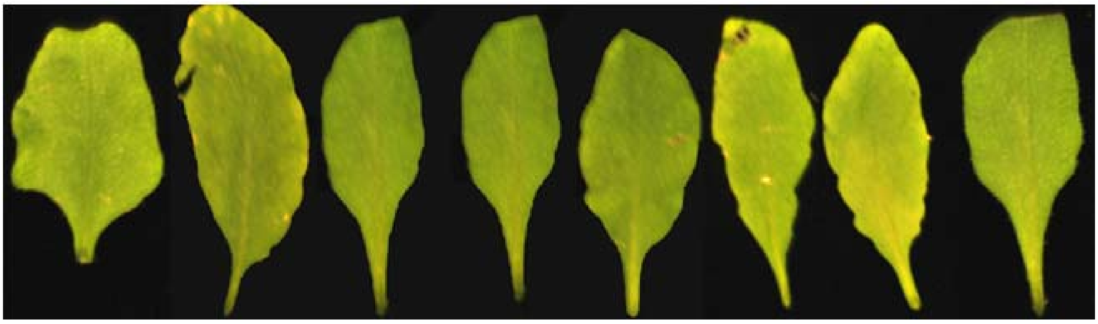


Figure 71 Leaves of as*1* Columbia-0 plants containing *AS1* transgene.
Each leaf is representing independent T1 line. 35S *AS1* has rescued the as*1* mutant phenotypes.

Figure 72 Leaves of as*1* Columbia-0 plants containing *BARP1* transgene.
Each leaf is representing independent T1 line. 35S *BARP1* has rescued the as*1* mutant phenotypes.

Figure 73 Leaves of *as1* Columbia-0 plants containing modified *AS1* transgene.
Each leaf is representing independent T1 line. 35S modified *AS1* has rescued the as*1* mutant phenotypes.

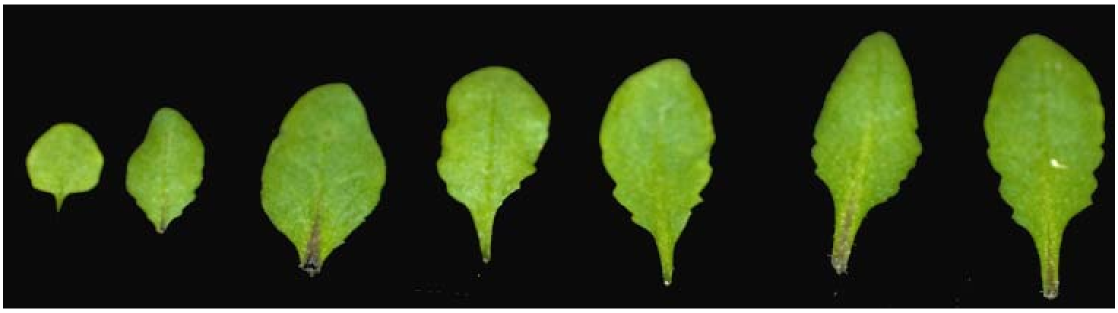


Figure 74 Leaves of wild type Columbia-0 plants.

Figure 75 Leaves of wild type Columbia-0 plants containing the *AS1* transgene.

Each leaf is representing independent T1 line. Plants have longer petioles and narrower leaves as compare to wild type plants.

Figure 76 Leaves of wild type Columbia-0 plants containing a *BARP1* transgene.
Each leaf is representing independent T1 line. Over expressing lines have longer petioles and narrower leaves as compare to wild type plants.

Figure 77 Leaves of wild type Columbia-0 plants containing a *BARP1* transgene (T2 lines).
Each leaf is representing independent T2 line. Over expressing lines have longer petioles and narrower leaves as compare to wild type plants.

Figure 78 Leaves and whole plant of wild type Columbia-0 plants containing modified *AS1* transgene. Each leaf is representing independent T1 line. Phenotypes varied from *as1* like to wild type plants which indicate that the transgene has dominant negative effects.

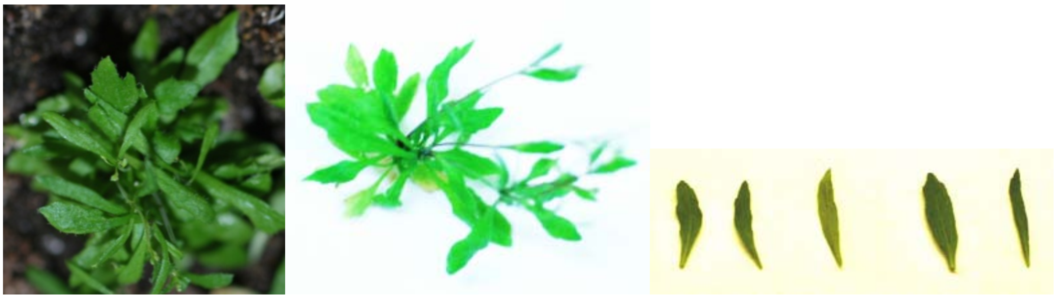


Figure 79 Leaves and whole plants of wild type Columbia-0 plants containing a modified *AS1* transgene (T2 lines). Phenotypes resemble 35S RS2△Myb.

Figure 80 35S RS2 △Myb phenotype in wild type back ground. (Theodoris *et al*., 2003).

Figure 81 Some representatives of novel leaves among T1 transgenic lines in Wild type background.
From Top left to bottom right: 1-4 35S *BARP1*, 5-7 35S *AS1* and 8-15 35S Modified *AS1*.

5.4. Discussion

Previous work has shown that *ARP* genes function is conserved between lycophytes (*Selaginella kraussiana*), monocots (*Zea mays*) and eudicot plants (*Arabidopsis thaliana*, *Cardamine hirsuta*, *Lotus japonicus*) (Harrison *et al*, 2005; Theodoris *et al*., 2003; Hay and Tsiantis, 2006; Luo *et al*., 2005). Likewise *BARP1* rescued *as1* mutant phenotypes indicates that *BARP1* is functionally equivalent to other *ARP* genes. Modified *AS1* rescued *as1* mutants partially which may suggests that modified *AS1* is partially functional.

ARP genes are member of a small unique MYB-related gene family that are required for repressing expression of certain *KNOX* (*KNOTTED1-like homeobox*) genes in leaves and consists of two protein domains. The Myb domain presumed to be involved in nucleic acid binding is at the N-terminus. The Myb domain is highly diverged compared with that in other Myb proteins and the DNA recognition helix in Myb repeat R3 is completely unique. ARP proteins bind DNA and cofactors are required for nucleic acid interaction. The C-terminal domain is involved in homodimerization and is highly conserved among *ARP* genes family and does not show homology to any other known protein sequence.

Transformation of wild type plants with modified *AS1* yielded *as1* like mutant plants which indicates that modified *AS1* is a dominant negative locus. *AS1* and modified *AS1* may have affinity to bind similar nucleic acids and binding of modified *AS1* with those nucleic acids may have interfere with the function of *ARP* genes but this is unlikely as modified *AS1* rescued *as1* mutants partially. Another possibility is that homodimerization of *AS1* and modified *AS1* may have yielded a protein structure which was unable to perform normal *ARP* gene function and yielded mutant like phenotypes, as RS2 protein is able to make dimers with AS1 proteins. Further *AS1* and modified *AS1* may have various levels of interactions as modified *AS1* over expressed lines have yielded a range of phenotypes ranging from *as1* like plants to wild type like plants.

Class III *HDZip* genes have a fundamental role in the shoot in establishing a functional apical meristem and polarity in lateral organs (Floyd and Bowman, 2006). The novel phenotypes observed among transgenic plants resembles those of *HDZip III* mutants. *ARP* genes work upstream of *HDZip III* genes. The fact that leaf form has altered in novel ways lends support to more complex models in which the ARP protein interacts with and possibly titrates factors that normally interact with *HDZip III* genes to modulate leaf morphology.

5.5. Conclusion

Both *BARP1* and *BARP2* like *AS1* (modified *AS1*) genes are functionally equivalent to *ARP* genes for regulation of leaf morphology as both can complement *as1* mutant plants. Modified *AS1* genes over expression in *Arabidopsis* suggested that the modified *AS1* is a dominant negative locus.

Chapter 6

Gene Expression Analysis

6.1. Introduction

There is a great debate about whether "Evolution of form is very much a matter of teaching very old genes new tricks" (Carroll, 2005) as Carroll believes or whether the evolution of form is very much a matter of teaching old genes to make new genes (Coyne, 2005).

According to Carroll, the main source of evolutionary changes is in the switches that control proteins instead of change in protein coding sequence. These switches are the promoters and enhancers in DNA that regulate transcription. They promote evolution by causing existing genes to be expressed at new times and in new places. Carroll also claims that proteins are resistant to evolutionary change because they are often involved in many pathways, and therefore a change in protein sequence, while enhancing one aspect of the protein's many functions, could damage several others. In contrast, changing an enhancer or promoter can affect the expression of a single protein without altering its structure, so such changes are more likely to be adaptive. He denies the idea that new genes underlie the evolution of diversity in most animal groups and hypothesises that changes in the expression patterns in same set of genes enable very different species to develop using essentially the same tool kit.

Carroll points out that dissimilar species may be genetically similar. For example mice and humans share about 25,000 genes and chimps and humans are almost 99 percent identical at the DNA level. Since the coding sequences of genes are so widely shared, the differences may arise through the evolution of non coding regulatory elements (Carroll, 2005). So the differences in phenotypes arise due to the changes in promoters and enhancers acitivity (Carroll, 2005).

The studies of genetic variation in other species also reveal the correlation between the variation in spatial patterns of gene expression and phenotype. For example, expression of *Distal-less* is linked with the formation of eye spots on butterfly wings (Brakefield *et al*., 1996), so a change in the spatial pattern of *Distal-less* expression can confer a change in the size and distribution of eye spots.

Comparative studies of *Hox* gene expression between primitive wingless and advanced winged insects proposes that *cis*-regulatory elements of wing formation genes are under the negative control of *Hox* genes in some body segments (Carroll *et al*., 2005). In crustaceans, thoracic limbs are transformed into feeding appendages due to the changes in *Hox* gene expression patterns (Averof and Patel, 1997).

On the other hand Coyne claims that changes in proteins are the cause of diversity of form on earth (Coyne, 2005). He argues that humans have about 32,000 protein-coding genes while fruitflies only 13,000 and between 40% and 50% of humans protein-coding genes have no known homologues in flies. Clearly, new proteins are responsible for the difference between these species. Further humans and chimps have different amino-acid sequences in at least 55% of their proteins, a figure that rises to 95% for humans and mice (Mikkelsen *et al*., 2005; Chinwalla *et al*., *2002*). Thus we cannot exclude protein-sequence evolution as an important reason of evolution of form. Coyne rejects Carroll's idea that "change in protein-coding sequence can destroy its one of several functions" by describing the processes of protein evolution which do not have any injurious side effects. These include gene duplication and whole genome duplication events. Extra copies of a gene can arise by unequal crossing over or by reverse transcription, allowing one copy to retain its function while the other assumes a new function. This process is thought to have been a major source of novelty during evolution (Ohno, 1970; Zhang *et al*., 2003). A large fraction of genes are members of families derived from repeated duplications and diversification of ancestral genes, a process that has yielded many evolutionary novelties. These families include the globins; immunoglobulins; opsins (which led to colour vision in Old World primates); and olfactory receptors (almost certainly involved in the evolution of a keen sense of smell in land animals). Lactalbumin, which helps to produce milk in mammals, resulted from a duplication of lysozyme, and the crystallins of our eye lenses are ultimately derived from heat-shock genes (Walsh and Stephan, 2001; Thornton and De Salle, 2000).

It is possible that gene function evolved, giving rise to distinct morphological traits in different species, either by changes to upstream elements or by changes to the properties of the gene product. Also, recruitment of new target genes could change the output of the original gene and, as a result give rise to new phenotypes.

Changes in the expression pattern of key regulators are important in plants as well. An excellent example is genetic regulation of flower development which is quite well understood in *Arabidopsis* and snapdragon. The ABC model posited that floral organ identity is controlled by three gene functions, A, B and C that act in combination; A-function alone specifies sepal identity, A- and B-functions together control petal identity; B- and C-functions together control stamen identity; C-function alone specifies carpel identity (Coen and Meyerowitz, 1991; Bowman *et al*., 1991).

C- function gene *AGAMOUS* (*AG*) in *Arabidopsis* regulates both the male and female organ identity. Whereas in *Antirrhinum* two C-function genes *FARINELLI* (*FAR*) and *PLENA* (*PLE*) have undergone sub-functionalization and contribute unequally to specify the male and female organs. Both 35S *AG* and 35S *PLE* were able to convert sepals to the carpels and petals to stamens in *Arabidopsis* and *Antirrhinum* respectively. However 35S *FAR* could convert petals into stamens in *Antirrhinum* but has no affect on sepal identity (Causier *et al*., 2005). The difference in the function of *AG*, *PLE* and *FAR* was due to the presence of an extra amino acid in *FAR* proteins which has altered the protein-protein interactions (Airoldi *et al*., 2010).

Specific adaptations are correlated with the changes in the pattern of gene expression in the evolution of physiological traits. Flowers of *Clarkia breweri* which emit a strong sweet scent have evolved from an extant nonscented species, *C. concinna*. The scent is controlled by the production of S-linalool, an acyclic monoterpene. *Lis*, the gene encoding S-linalool synthase, is highly expressed in *C.*

breweri, as compare to in the non scented *C. concinna*. The differential expression of *Liz* regulates the scent emission in these species (Dudareva *et al*., 1996).

Expression patterns of a suite of enzymes, normally used for housekeeping functions, are altered in C4 species which is accompanied by alterations in internal histology and chloroplast structure of the leaf. Expression patterns common to all C4 lineages are central to the evolution and development of the pathway and patterns that vary are lineage specific (Sinha and Kellogg, 1996).

Changes in the expression pattern of *ARP* and *KNOX* genes are reported to be associated with leaf form in a range of species. Mutations at the *Phantastica* (*Phan*) locus of *Antirrhinum majus* have shown that subtle changes in the level or pattern of *Phan* activity can give rise to a variety of organ morphologies including needle like leaves, cup shaped leaves and peltate leaves (Waites and Hudson, 1995). The expression domain of *ARP* in a range of species with compound leaves correlates with the type of compound leaf i.e. pinnate, palmate or peltate palmate (Kim *et al*., 2003a; Kim *et al*., 2003b). This suggests that the convergent evolution of *ARP* expression may be responsible for leaf form variation in species.

In tomato compound leaves have evolved through a change in expression of a homeobox-containing gene in the leaf primordia (Hareven *et al*., 1996). An up regulation of homeobox-containing gene *LeT6* resulted in the conversion of unipinnately compound leaves into three- or four fold pinnately compound leaves (Chen *et al*., 1997). Species-level differences in leaf form in the native tomatoes of Galapagos Islands are also due to changes in the *KNOX* (*PETROSELINUM* (*PTS*)) gene expression where the expression of the *KNOX* gene is up regulated in the leaves of highly dissected *Solanum galapagense* in comparison to its expression levels in the less dissected sister species *Solanum cheesmaniae* (Kimura *et al*., 2008). The compound leaf character of *Elaeis guineensis* (palms) is found to be dependant on the expression of *KNOX1* genes (Stefan *et al*., 2007) and reactivation of *KNOX* genes expression after leaf formation in the basal meristem of *Welwitschia mirabilis* is associated with indeterminate leaves which can grow for 400 to 1500 years (Pham and Sinha, 2003).

Differential expression of *KNOX* genes between pinnately compound leaved *Cardamine hirsuta* and simple leaved *Arabidopsis thaliana* correspond to the natural variation in the leaf form of these two closely related species. The difference in the expression pattern of *KNOX* genes between these two species is driven by the variation in the promoter region of *KNOX* genes between two species (Hay & Tsiantis, 2006).

A model has been proposed correlating the expression pattern of *ARP* and *KNOX* genes with leaf form (Champagne *et al.* 2004; Kim *et al.*, 2003). Expression of *ARP* genes along the adaxial domain of developing leaves mostly generates simple leaves. *ARP* expression confined to distal regions of developing leaves results in the formation of peltate leaves and *ARP* expression along the boundary of adaxial domain depict the leaflet placement in compound leaf formation. Lack of *KNOX* gene expression in developing leaves mostly generates simple leaves and *KNOX* expression reactivation in developing leaves results in the formation of compound leaves (Champagne *et al.* 2004; Kim *et al.*, 2003). I have tested this hypothesis in *Begonia* section Gireoudia species through in situ hybridization in a range of species. This section has a number of simple, peltate and compound leaved sister species making it an excellent model to study the inter specific variation of leaf form.

6.2. Results

6.2.1. Fixation of vegetative buds

Vegetative buds of different species were fixed and sectioned as described in chapter 2. I tried to fix the vegetative buds of the same age from each species in order to compare the expression pattern of genes at the same stage of development. In order to make the judgement about the age and orientation of vegetative buds to be used for in situ hybridization, I initially dissected vegetative buds of different species under a dissection microscope. Close observation of different species gave me a general idea about the time differences in the initiation of two successive primordia. I have observed the smallest time difference in the initiation of two successive primordia was for *B. mazae* and longest was for *B. thiemei.* This could have been dependant on the number of leaves each species generated in a specific time period or could have been due to the size of the leaves. *B. mazae* clearly generates more leaves during a given time period as compared to *B. thiemei* and produces smaller leaves than *B. thiemei* as well.

For hybridization I harvested the buds of the same size from different specimens which were approximately of the same age. I did a quick dissection of the vegetative buds to remove larger and unwanted parts while they were still attached to the plant and directly fixed them into PFA. Older stipules wrapped up the developing leaves very tightly and could hinder the fixation of enclosing material, so I dissected the material further after 2 hours of fixation with fine needles and forceps to reveal the youngest stipules enclosing youngest primordia and shoot apical meristem. *Begonia* species have a large amount of trichomes which have created problems during sectioning. I have tried to remove them after buds had undergone ethanol and histoclear treatments but doing this left material unintegrated. So the only way to generate good fixed material was disection of lots of buds at the same time. In *Begonia* the leaf primordium occupies a significant portion of the shoot apex. At the base of each leaf a pair of stipules is present. Both stipules nearly enclose the next leaf primordium that is initiated on the SAM apex (Figure 84) (Barabe *et al*., 2007).

6.2.2. Transcription of probes

The C-terminal domain (240 amino acids) of BARP1 which is downstream of the Myb domains was used to make probes for detection of RNA by DIG in insitu hybridization. The reason for using the C-terminal domain only is that the Myb domains of *ARP* genes show significant similarity with most plant Myb proteins and the use of the whole *BARP1* gene as a probe would have given misleading results by hybridizing with other Myb proteins. Myb Domains of ARPs are highly similar with other Myb like proteins in plants. For example Myb domains of *ARP* genes in snapdragon had only one amino acid upstream of the first repeat, and the first repeat was 2 or 3 amino acids longer than that of other MYB proteins (Waites and Hudson, 1998).

PCR based sense and antisense probes were made using *BARP1*-C-TerT3, *BARP1*-C-TerT7 and *BARP1*-C-TerSP6 primers to amplify C-terminal domain of *BARP1* genes with High Fidelity Taq polymerase as described in chapter 2. Genomic DNA of different species was used as templates and amplified fragments were sequenced to confirm the region.

For *KNB1* and *KNB2* gene specific primers were used for making probes as described in chapter 2. *KNB1* and *KNB2* have high homology so a region in the homeodomain which was less conserved between *KNB1* and *KNB2* was selected. Species specific probes were used for hybridization. The positions of primers on *BARP1* gene, *KNB1* genes and *KNB2* genes is described in Figure 82 and Figure 83.

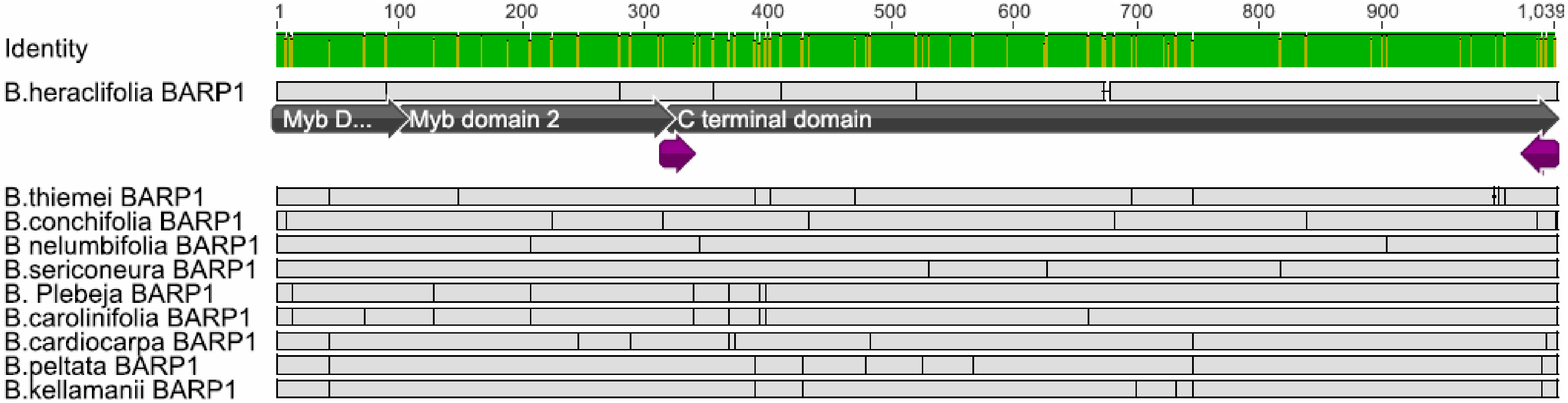


Figure 82 Position of primers for making *BARP* genes probes for in situ hybridization.
The C-terminal domain was used to detect the signals for *BARP* probes in insitu hybridization.

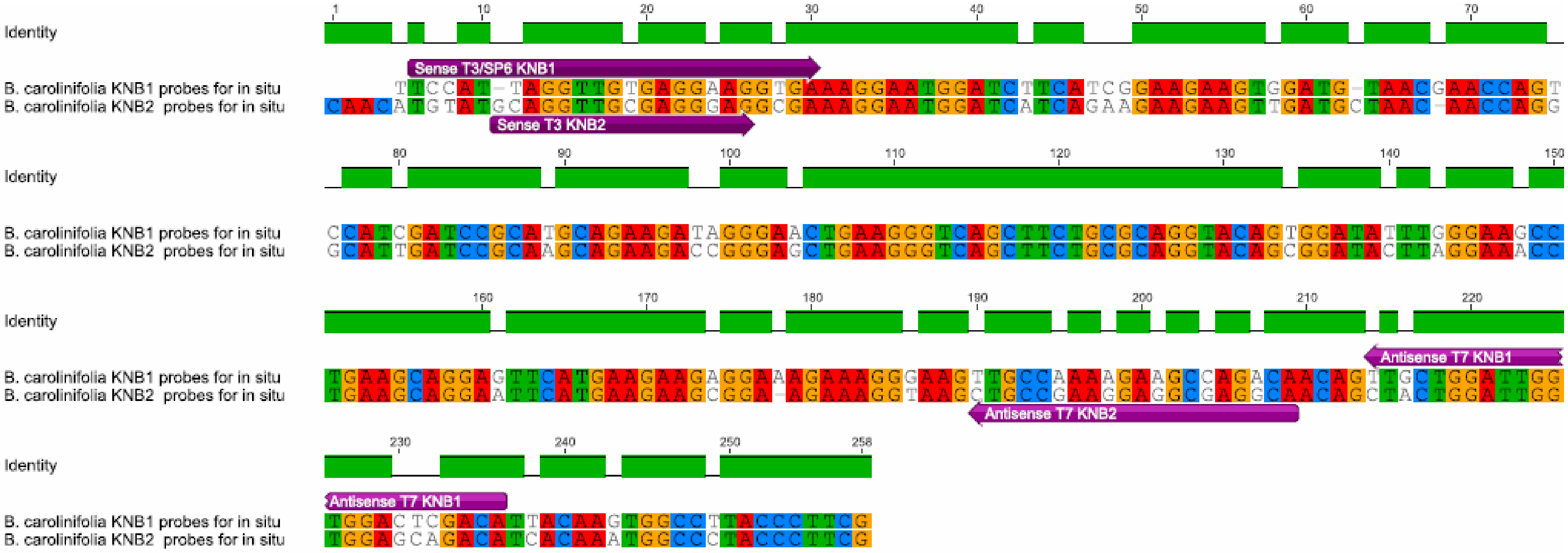


Figure 83 Position of primers for making *KNB1* and *KNB2* genes probes for in situ hybridization.

6.2.3 Expression of *BARP1* in *Begonia*

Different regions of meristems are defined on the basis of histological studies performed by Kidner (Figure 12); McLellan (1990) and Barabe (Barabe *et al*., 2007) (Figure 84). The expression of *BARP1* is described in sequential transverse sections of *B. mazae* in Figures 85-87. Figure 85 shows the sections from stem to leaf primordia, Figure 86 shows sections from petiole to blade tip and Figure 87 are comprised of section from axillary meristem to the leaf primordia. In *B. mazae BARP1* is expressed in leaf primordia, vascular bundles, and stipule primordia and at the distal tips of developing laminae and developing stipules (Figure 85-87). *BARP1* is also expressed in adaxial regions of petioles in *B. mazae* (Figure 86). There is strong expression of *BARP1* in the dormant axillary meristem and there was no *BARP1* expression in the active meristem in *B. mazae* (Figure 85-86).

In *B. carolineifolia BARP1* is expressed at the tips of leaflet primordia, vascular bundles and at the tips of developing stipules (Figure 88). Similar to *B. mazae, BARP1* in *B. kellemanii* and *B. heracleifolia* is expressed in leaf primordia, stipule primordia and vascular bundles (Figure 90 and 91). *BARP1* is expressed all over in the younger leaf primordia and expression becomes confined to the tips of developing leaves in *B. kellemanii* and *B. heracleifolia* (Figure 90 and 91). *BARP1* is also expressed on the adaxial side of the petiole in *B. heracleifolia* (Figure 91 I).

6.2.4 Expression of *KNB1* in *Begonia*

KNB1 in situs were not very clear but *KNB1* is expressed in the shoot apical meristem, tips of developing leaves and in the developing stipules of *B. mazae, B. kellemanii* and *B. heracleifolia* (Figure 89-91). In *B. carolineifolia KNB1* expression is very strong at the tips of developing leaflets (Figure 88).

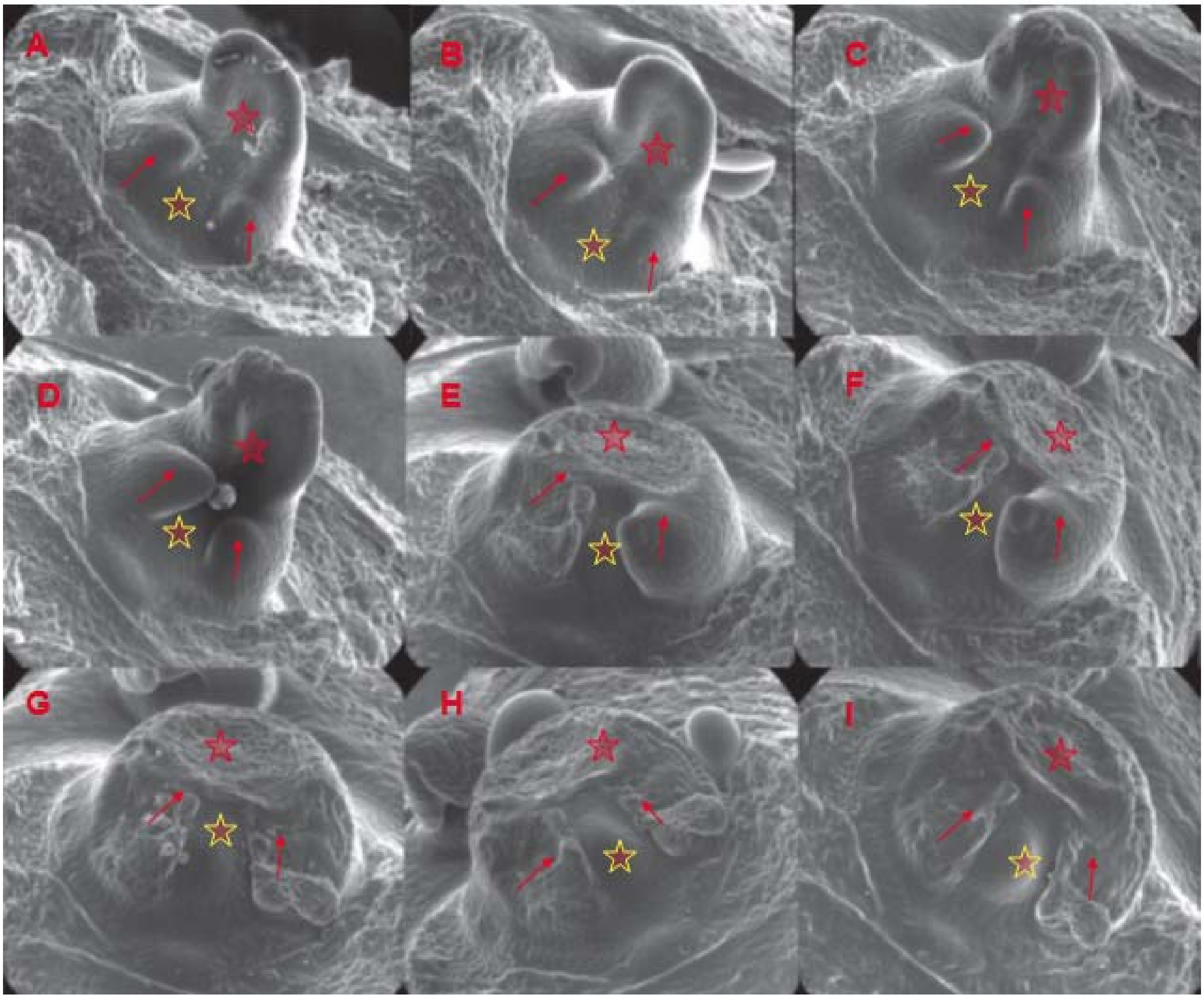
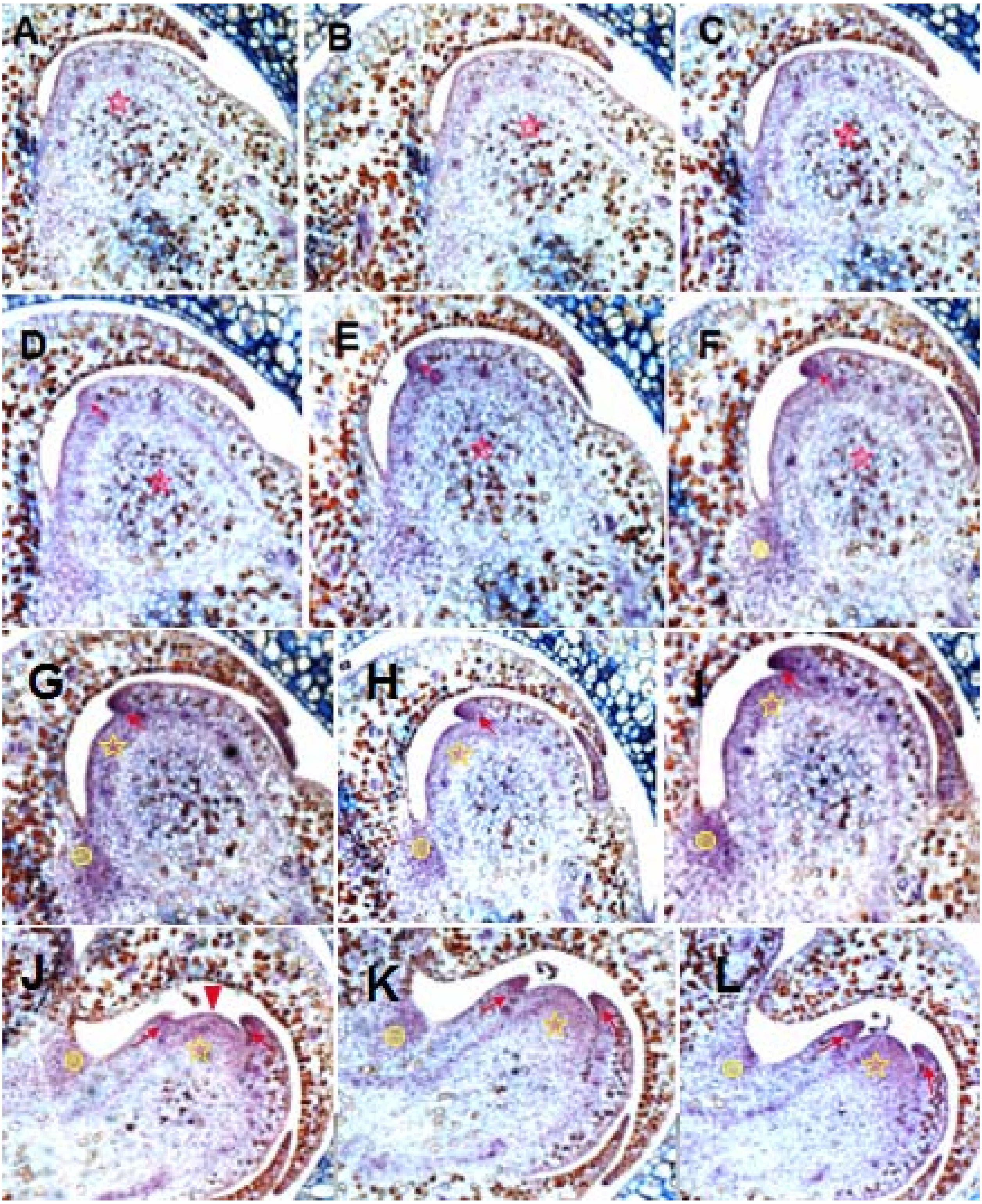


Figure 84 Developmental sequence of a single shoot apical meristem in *Begonia* (A-I).
It shows the development of a leaf primordium followed by the initiation of a new leaf primordium; arrows indicate the stipules surrounding the leaf, yellow star indicates the region of SAM and red star indicates the region of developing leaf (Barabe., *et al*., 2007). The stipules and shoot apical meristem are visible at the base of the leaf.


A
B
C
D
E
F
G
H
I
J
K
L

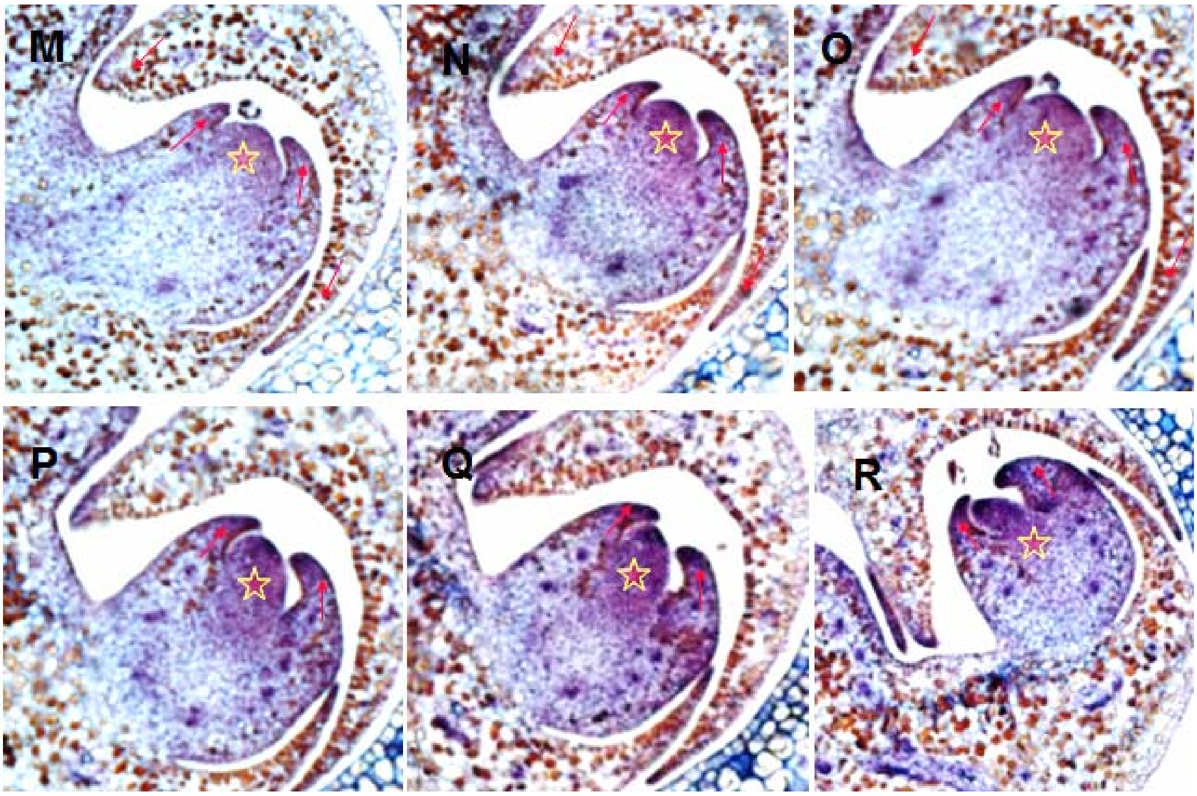


Figure 85 Sequential longitudinal sections of *Begonia mazae* vegetative bud (A-R).
The red star indicates the region of stem and emerging stipules are shown by red arrows. Yellow stars indicate the region of SAM. Red triangles represent the developing leaf. *BARP1* is expressed at the tips of developing stipules (D-R). *BARP1* is expressed at the sites of developing leaf primordia (J-R). *BARP1* is also expressed at the sites of silent axillary meristems indicated by yellow circles (F-L).

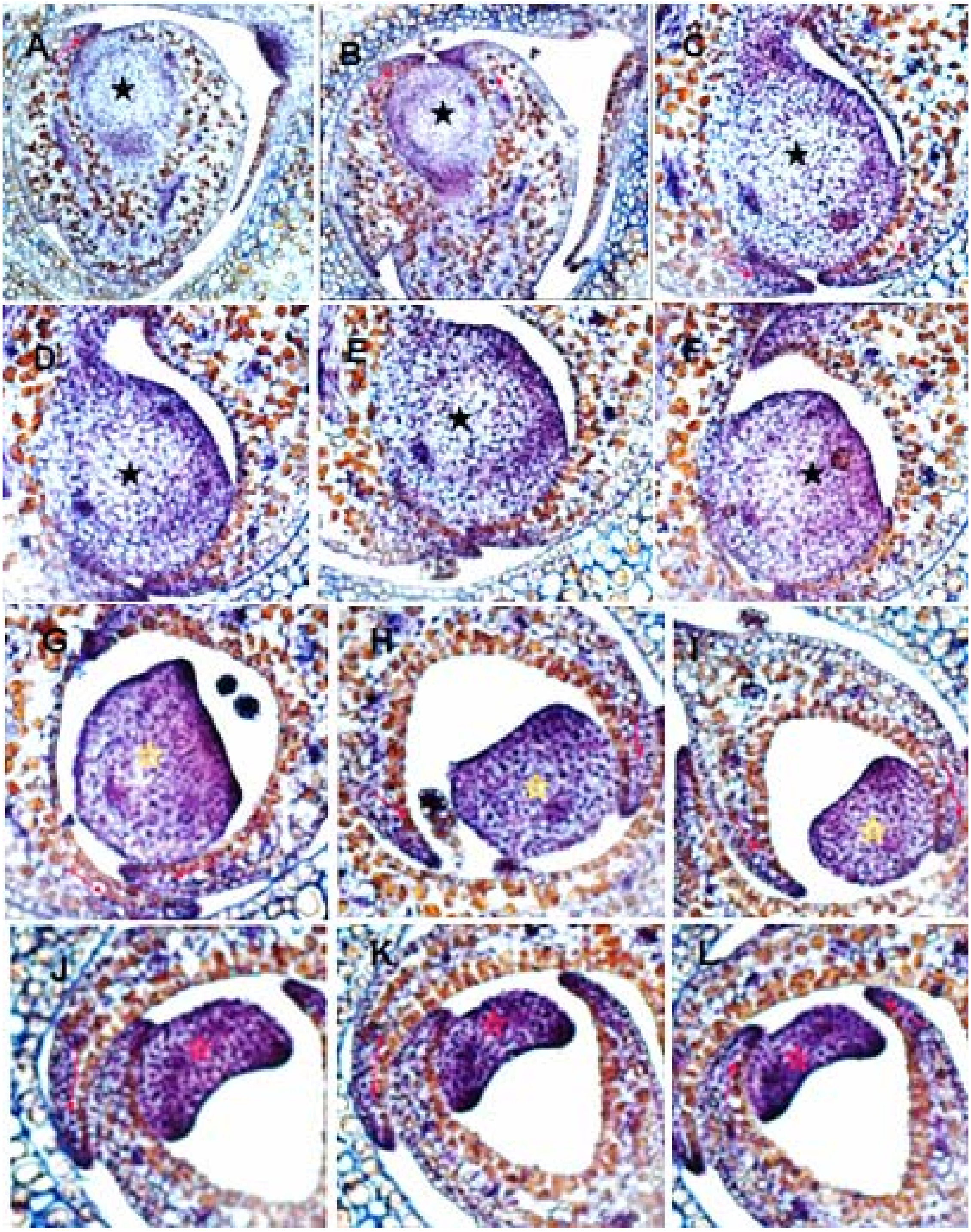


Figure 86 Sequential Transverse sections of *Begonia mazae* vegetative bud.
The sections are through the stems to the developing blade. The stipules are shown by red arrows, stems by black stars, petioles by yellow stars and developing blade by red stars. The petioles are adaxialized at younger stage of development and *BARP1* is expressing at the

adaxial region of petioles (G). *BARP1* expression is confined to the distal margins of the developing blade (I-L).

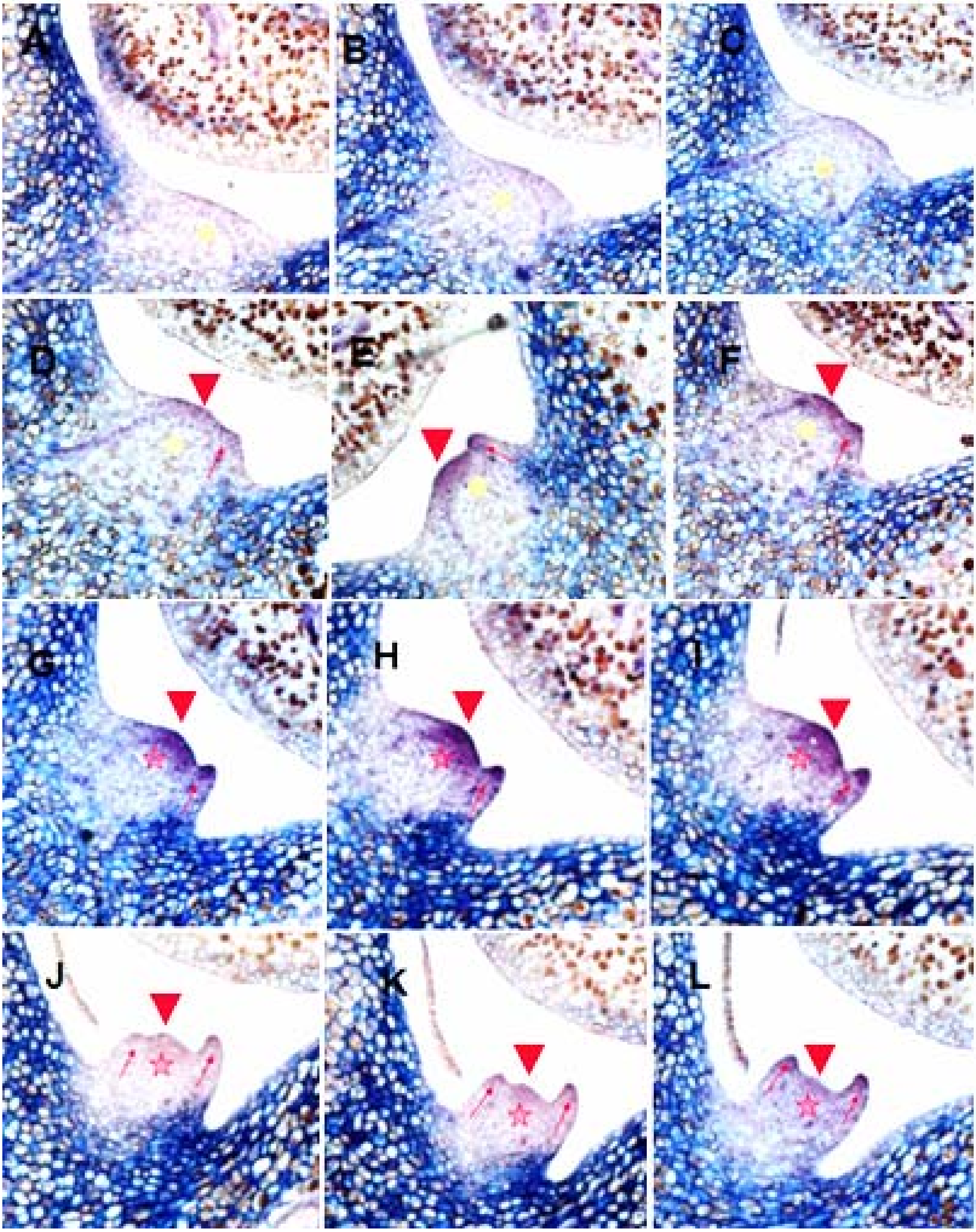


Figure 87 Sequential Transverse sections of *Begonia mazae* vegetative bud (A-L).
The yellow circles indicate the region of SAM and emerging stipules are shown by red arrows. Red triangles represent the developing leaf. *BARP1* is expressed at the tips of developing stipules (J-L). *BARP1* is expressed at the sites of developing leaf primordia (D-L). *BARP1* is not expressed in active axillary meristems indicated by yellow circles.

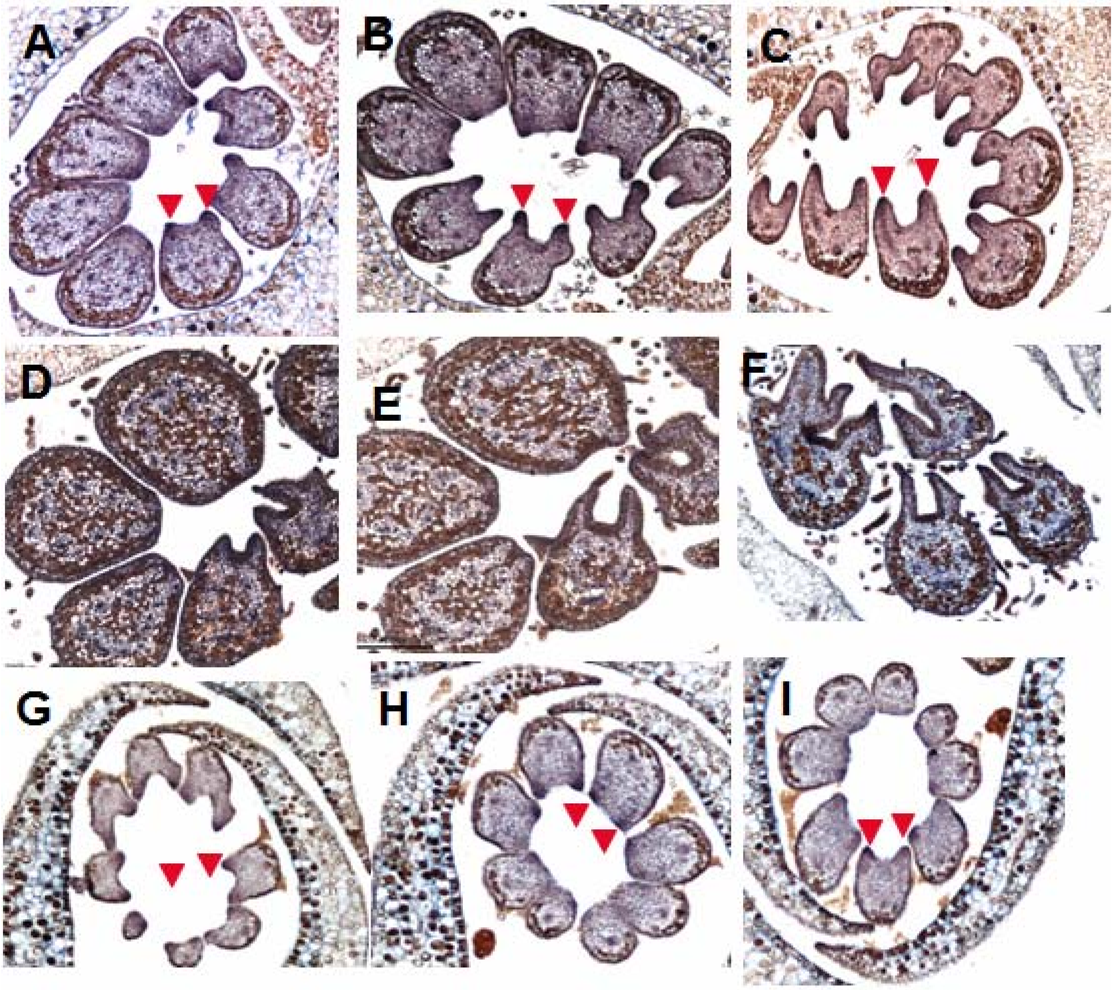


Figure 88 Expression of *KNB1*, *KNB2* and *BARP1* in *B. carolineifolia*.
***KNB1* (A-C), *KNB2* (D-F) and *BARP1* (G-I) in *B. carolineifolia*. *KNB1* and *BARP1* are co-expressed at the distal parts of leaflet primordia as indicated by red triangles at distal margins of one of leaflets primordia in A-C and G-I. *KNB2* is expressed everywhere in leaflet primordia.**

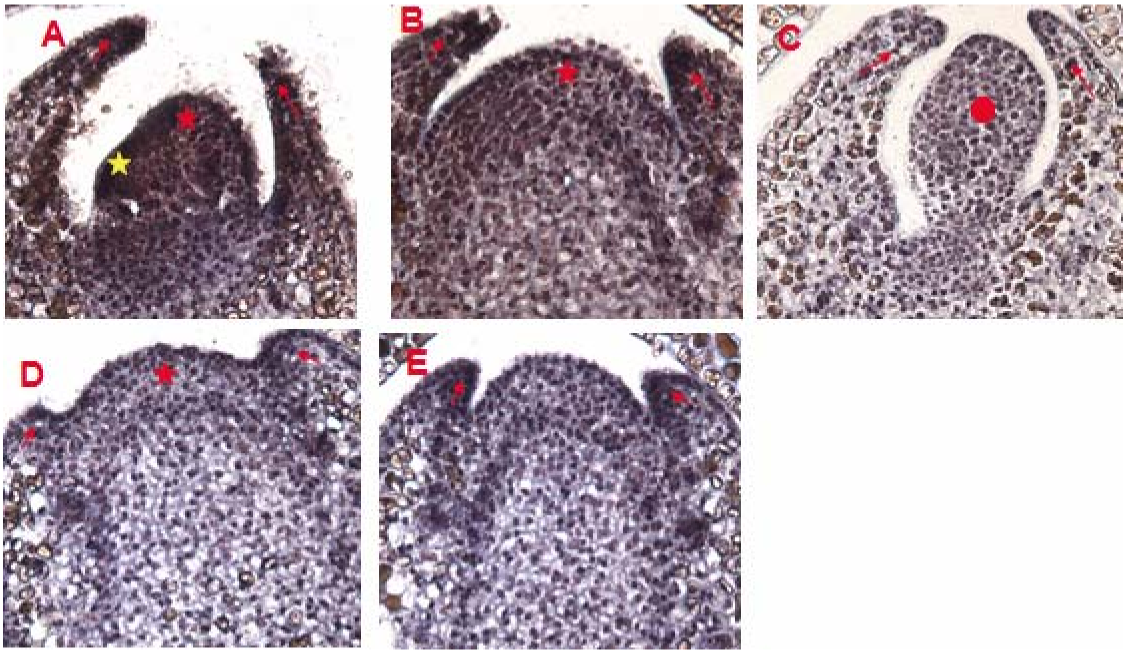


Figure 89 Expression of *KNB1* and *KNB2* in *B. mazae*.
***KNB1* (A-C), and *KNB2* (D-E) in *B. mazae*. *KNB1* and *KNB2* are expressed everywhere (A-E). Red arrows indicate the distal parts of stipule primordia, red stars indicate the leaf primordia, red circles show developing leaf blade and yellow star indicates the shoot apical meristem.**

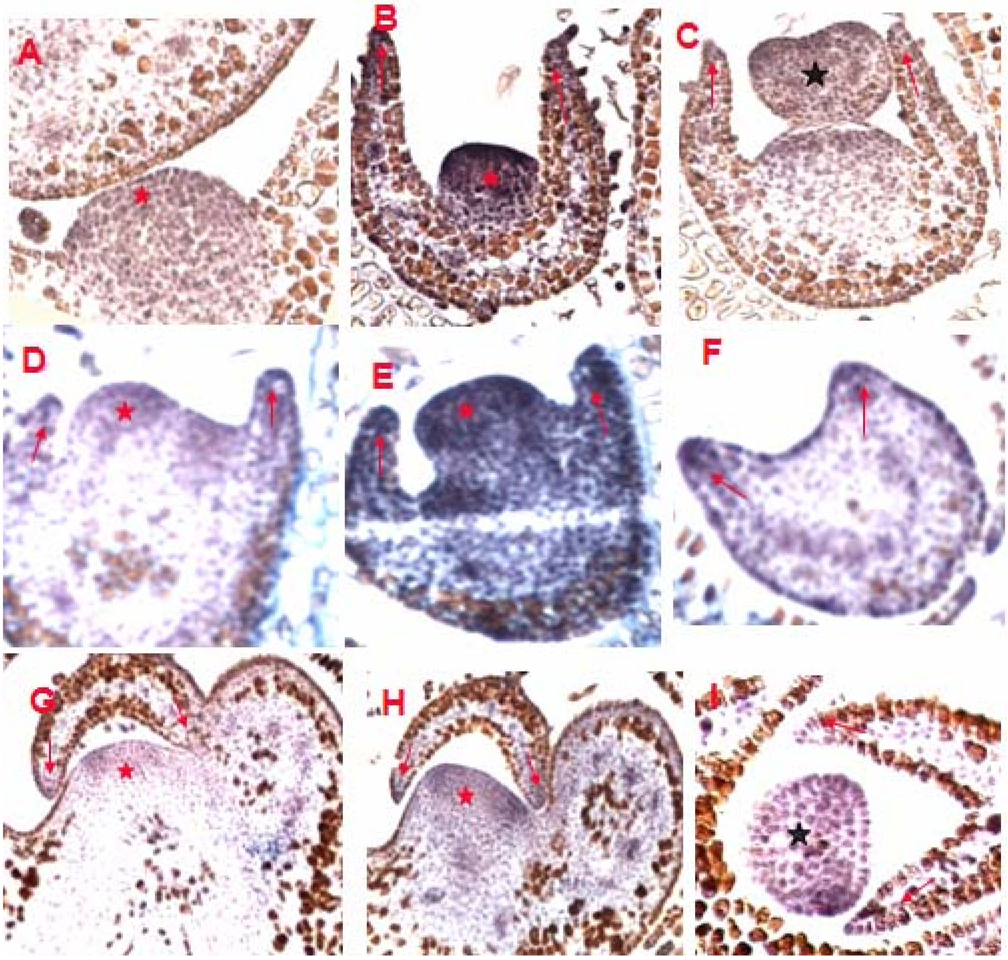


Figure 90 Expression of *KNB1*, *KNB2* and *BARP1* in *B. kellemanii*.
***KNB1* (A-C), *KNB2* (D-F) and *BARP1* (G-I) in *B. kellemanii*. *BARP1* is expressed at the distal parts of stipule primordia as indicated in G-I. *BARP1* is also expressed at leaf primordia (G,H) and in the developing leaves (I). *KNB1* and *KNB2* are expressed everywhere (A-F) and *KNB2* expression is stronger at the distal margin of stipule primordia (D-E). Red arrows indicate the distal parts of stipule primordia, red stars indicate the leaf primordia and black stars show developing leaf blade.**

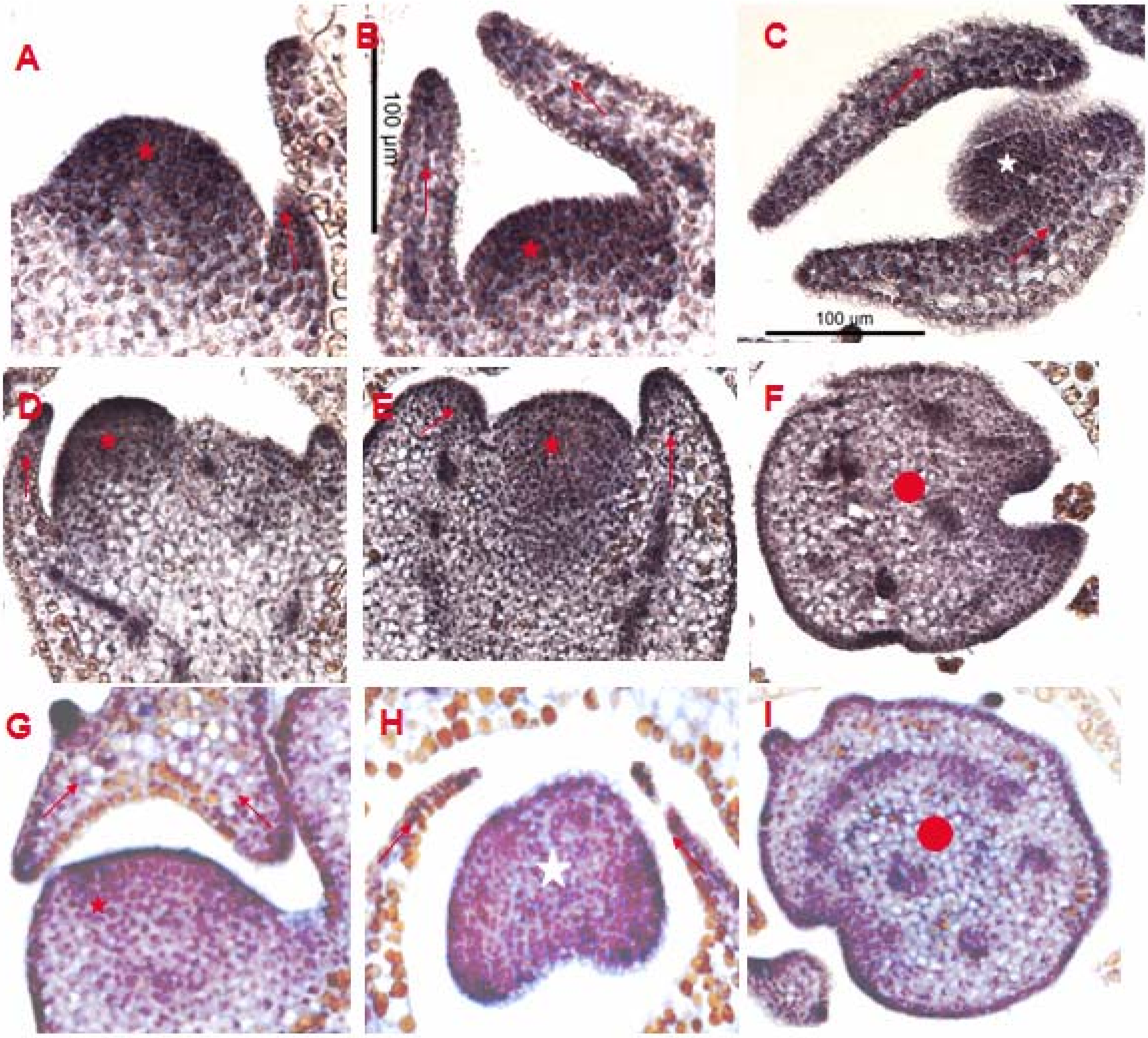


Figure 91 Expression of *KNB1*, *KNB2* and *BARP1* in *B. heracleifolia*.
***KNB1* (A-C), *KNB2* (D-F) and *BARP1* (G-I) in *B. heracleifolia*. *BARP1* is expressed at the distal parts of stipule primordia as indicated in G-H. *BARP1* is also expressed at leaf primordia (G) and in the developing leaves (H). *BARP1* is also expressed along the adaxial margins of petiole and in vascular bundles of petiole (I). *KNB1* and *KNB2* are expressed everywhere (A-F). Red stars indicate the leaf primordia, arrows indicate stipule primordia and circles indicate petioles sections.**

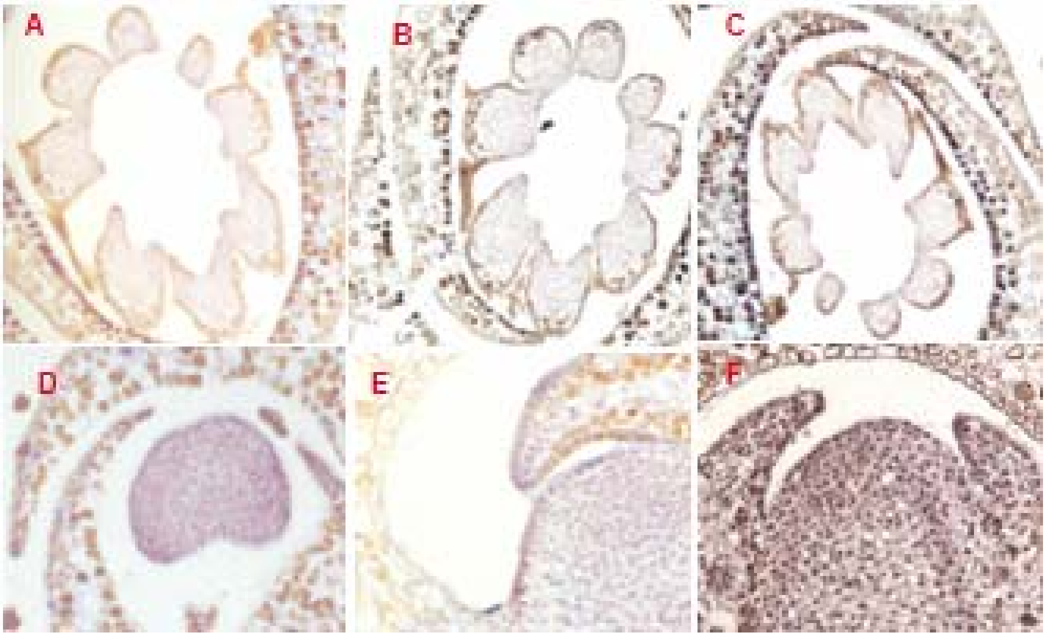


Figure 92 Sense probes for *KNB1*, *KNB2* and *BARP1*.
***KNB1* (A) and *BARP1* (B-C) in *B. carolineifolia*. *BARP1* in *B. mazae* (D) and *B. kellemanii* (E) and *KNB1* in *B. mazae* (F).**

6.2.5 Expression of *KNB2* in *Begonia*

KNB2 probes failed to give a clear signal but *KNB1* expression appears to be stronger than *KNB2* in *B. mazae* (Figure 89). As in situs, these are not very quantitive but the probes are derived from the same regions and the exposure times were the same. *KNB2* appears to be expressed in SAM, leaf primordia, vascular bundles, at the tips of developing leaves and in stipules of *B. kellemanii* (Figure 90).

Most of the time no signals were detected for sense probes of *BARP1*, *KNB1* and *KNB2* (Figure 92). However sometimes signals were detected in sense probes similar to anti semse probes expression patterns.

6.3 Discussion

Unlike in maize, *Anthirrinum* and *Arabidopsis*, tomato *ARP* (*LePHAN*) and *KNOX* transcripts are co-localized within the shoot apex (Koltai and Bird, 2000). In the compound leaved plants *Senna actinophylla, Acacia hindisii, Vitex cannabifolia, Dizygotheca elegantissima, Oxalis regnellii, Koelreuteria paniculata, Aquilegia formosa* and *Pachira aquatica ARP* genes are expressed in the shoot apical meristem, stem and leaf vascular traces (Kim *et al*., 2003). In barley *KNOX* genes are expressed in the SAM and young leaves (Muller *et al*., 2001). Whereas *BARP1* and *KNOX* genes are coexpressed in the stem and vascular bundles of *Begonia* section Gireoudia species.

ARP genes are required for the establishment of dorsal identity in *Antirrhinum majus* as it is responsible for regulating all aspects of dorsoventrality in leaves, bracts and petals from specifying the position of laminal initiation early in organ development, to determination of dorsal cell types at a later stage in this plant (Waites and Hudson, 1995). The presence of *BARP1* expression on the adaxial side of petiole of *B. mazae* and *B. heracleifolia* (Figure 86 G and 91 I) indicates that *BARP1* could be specifying the position of laminal initiation and indicates dorsoventral patterning in *Begonia* petioles early in development which later on become completely abaxialized.

The presence of a notch in the leaves at P1 stage of peltate *Begonia* species also indicates that establishment of peltateness is a late event in *Begonia* leaf development (chapter 3). The presence of *BARP1* expression on adaxial side of petioles may indicate that *BARP1* is promoting adaxial/abaxial polarity earlier in development and its confinement to the distal tips of laminae later may be the cause of establishing peltateness. The restriction of the *LePHAN* expression to the distal end of the leaf primordium in tomato plants has resulted in the production of peltately palmate leaves (Kim *et al.*, 2003). Further *ARP* expression was confined to the distal region of the leaf primordium in a range of peltate compound-leafed species (Kim *et al.*, 2003). Likely *BARP1* is expressed at the distal tips of laminae in *Begonia* species and all *Begonia* section Gireoudia species are peltate to some degree (abaxialized petioles and laminar outgrowth at the lamina- petiole attachment point) (chapter 3).

Stipules are attached to the main stem in pea and are flattened laminae that are conventionally described as lateral organs of the pea compound leaf (Sachs, 1972) and the *CRI* (*ARP* orthologue in pea) regulates stipule initiation (Tattersall *et al.*, 2005). The presence of *BARP1* expression in stipule primordia and developing stipules may indicate a role for *BARP1* in stipule initiation and stipule development in *Begonia*.

KNOX genes are linked with indeterminacy (Long *et al* 1996; Volbrechet *et al*, 2000). Their expression is deactivated in simple leaved species but reactivated in compound leaved species during leaf development for leaflet formation (Shani *et al*, 2009; Hay & Tsiantis, 2006; Harevan *et al*, 1996). *KNOX* expression patterns corresponded to the developmental stage of the leaf primordia and not necessarily with the final leaf morphology (Bharathan *et al*., 2002). *KNOX* expression is correlated with complex leaf primordia such as in *Lepidium oleraceum KNOX* proteins are expressed in the complex leaf primordium which undergoes secondary morphogenesis to form simple leaves (Bharathan *et al*., 2002). *KNOX* genes are expressed in the leaf primordia and developing leaves of *Begonia* section Gireoudia species. Their expression in leaves may be required for the formation of laminae outgrowth later in leaf development which may have required indeterminate environment as peltateness is established later in leaf development in *Begonia*.

KNOX independent mechanism of compound leaf formation has been reported for pea where *UNIFOLIATA* (ortholog of *Arabidopsis LEAFY)* regulates compound leaf formation (Gourlay *et al*., 2000). And *NO APICAL MERISTEM/ CUP-SHAPED COTYLEDONS3* (*NAM/CUC3*) family are required for proper expression of *KNOX* and *UFO*-like genes during compound leaf formation in *Solanam lycopersicon, Cardamine hirsuta* and *Pisum sativum* (Blein *et al*. 2008). These genes may be the key regulators for controlling compoundness in *Begonia*.

6.4 Limitations

I optimized the protocol of in situ hybridization for *Begonia* section Gireoudia species as described in chapter 2. I used PCR based probes for hybridization. Firstly PCR optimization for primers with T3, T7 and SP6 adapters was time consuming. I did different PCR based techniques and in my hands PCR worked best with longer primers. Secondly, the correct amount of RNA probes for hybridization was different for different genes and also varied for different species so I tested several concentrations of every RNA probe (300, 500, 700, 800 and 1000 ng/slide) as described in chapter 2.

I did get signals with sense probes at some times and sometimes these signals were in the same places as antisense and sometimes at random places. It may be because of using higher amount of RNA probes for *Begonia* as compare to other species as lower amounts of RNA probes were unable to detect signals. Every time fresh PCR products were used to transcribe probes and four PCR reactions (each in 50ul total volume) were pooled to get 800 ng of RNA probes for hybridization. I did not sequence all the PCR reactions each time and there is possibility that sometimes non specific products could have been generated and used for making probes which gave signals in sense probes. Idealy a positive control should be used alongside using negative control. I used *STM*-like *KNOX* genes as a positive control but could not get the probes working.

Finding the correct orientation of *Begonia* vegetative buds of fixed material was difficult task. For me transverse sections worked better; longitudinal sections material was coming in unintegrated parts through microtome which may be due to poor fixation. Describing different regions of *Begonia* vegetative buds in sections was challenging as little literature is available for *Begonia* histology.

6.5 Conclusion

BARP1 genes are expressed at leaf primordia, developing stipules, developing leaves and vascular bundles in simple, peltate and compound leaves of *Begonia* section Gireoudia. *BARP1* may be the key regulator for controlling peltate leaf form in this section because it is expressed in the distal tips of developing leaves which has been reported for most of peltate leaved species. This supports the association mapping studies where *BARP1* is a major locus for controlling peltateness (chapter 7).

Chapter 7

Genetic association between *ARP* and *KNOX* genes and leaf form in *Begonia* section Gireoudia

7.1. Introduction

Association mapping is a powerful tool to test for linkage of a phenotype with a genotype. It is usually performed by establishing a link between a particular genetic marker and the target trait in a fairly large population of segregating individuals by testing whether that marker is present more often in the target phenotype than expected by chance. This can than be verified by QTL analysis.

There are numerous examples where links between phenotypes and genotypes have been found on the basis of candidate gene association mapping (For example: Aranzana *et al*., 2006; Ehrenreich *et al*., 2009; Gonza´lez-Martı´nez *et al*., 2007; Gonza´lez-Martı´nez *et al*., 2008; Stracke *et al*., 2009; Abdurakhmonov *et al*., 2009; Belo *et al*., 2008; Skøt *et al*., 2005; D'hoop *et al*., 2008; Malosetti *et al*., 2007; Agrama *et al*., 2007; Wang *et al*., 2008; Jun *et al*., 2008; Stich *et al*., 2008 a; Stich *et al*., 2008 b; Yao *et al*., 2009). Usually a candidate gene is selected on the basis of previous knowledge. The gene has been implicated in the trait of interest through mutational analysis or expression data in related plants.

An excellent example of establishing a link between a candidate gene and a trait of interest has come from Yamaguchi's work on flattening in unifacial leaves of *Juncus prismatocarpus* (Yamaguchi *et al*., 2010). The unifacial leaves of *Juncus* are abaxialized and have become flattened despite lacking dorsoventral polarity (Yamaguchi and Tsukaya, 2010). A small number of candidate genes, which had been reported to control different aspects of leaf polarity in other species were tested in *J. prismatocarpus*. These candidate genes included members of the *HD-ZIP III, YABBY* and *KANADI* families, *ARF3/ETT, PRESSED FLOWER* and *ARP* gene families (Yamaguchi *et al*., 2010). First of all the expression patterns of these genes were studied in unifacial flattened leaves of *J. prismatocarpus* and unifacial cylindrical leaves of *J. wallichianus*. Only *DROOPING LEAF* (*DL*) (a member of the *YABBY* gene family) and *PRESSED FLOWER* (*PRS*) genes were differentially expressed in leaves of the two species (Yamaguchi *et al*., 2010). To uncover the gene responsible for the differences in leaf form F1 and 289 F2 interspecific hybrids were generated. Genetic analysis revealed that leaf flatness is a polygenic trait and two or more loci are controlling the trait alongside with dominant or semi dominant factors in *J. prismatocarpus* (Yamaguchi *et al*., 2010). CAPS (Cleavage amplified polymorphism sequence) markers were designed to determine the association between *DL* or/and *PRS* genes and laminae growths. CAPS markers are based on differences in restriction enzyme digestion patterns of PCR fragments caused by nucleotide polymorphism between species. Their results indicated that the *DL* locus or a locus tightly linked to *DL* is responsible for differential lamina growth between *J. prismatocarpus and J. wallichianus* (Yamaguchi *et al*., 2010).

As candidate genes for the variation in leaf form in *Begonia* section Gireoudia we have chosen Class 1 *KNOX* genes, the *NAM*/*CUC* and *YABBY* families and *ARP* genes. Differential expression of class 1 *KNOX* genes underlies the natural variation in the leaf form between pinnately compound leaved *Cardamine hirsuta* and simple leaved *Arabidopsis thaliana* (Hay & Tsiantis, 2006). The level of *KNOX* gene expression is also linked with the degree of compoundness in other groups. *KNOX* genes are expressed at higher level in the leaves of highly dissected *Solanum galapagense* in comparison to its expression levels in the less dissected sister species *Solanum cheesmaniae* (Kimura *et al*., 2008).

NAM/CUC3 genes are expressed at the distal boundaries of the leaf rachis at the points where leaflets are formed in *Solanam lycopersicon, Cardamine hirsuta* and *Pisum sativum*. They have been shown to regulate the expression of *KNOX* and *UFO*-like genes through a feed forward regulatory loop during compound leaf formation in these species (Blein *et al*., 2008). Class 1 *KNOX* like genes and *NAM/CUC3* are good candidates for the loci regulating leaf dissection in *Begonia* section Gireoudia.

The expression of *ARP* genes is linked with peltate leaf form in a range of species where *ARP* genes control the extent of lamina outgrowth and also regulate the number and placement of leaflets in compound leaf species (Kim *et al*., 2003). *YABBY* expression is associated with the development of the peltate leaf form in *Tropaeolum majus* (Gleissberg *et al*., 2005). This makes *ARP* and *YABBY* genes good candidates for loci regulating peltate leaf form in *Begonia*.

We have generated an F1 and back cross progenies for *Begonia* section Gireoudia species as described in chapter 3 to carry out genetic studies for leaf form. Genetic analysis has revealed that peltate and compound leaf forms are complex traits and may be controlled by more than one locus (Chapter 3). I isolated *ARP*, *KNOX* and *NAM/CUC3* like genes from *Begonia* species by degenerate PCR. *YABBY* genes were retrieved from *Begonia* transcriptome data for *B. conchifolia* and *B. plebeja*. The results of association mapping studies based on CAPS markers for *ARP*, *KNOX, YABBY* and *NAM/CUC3* like genes are described in this chapter.

7.2. Association between *BARP1* and leaf disection in (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia*

To test for association between *BARP1* genes and compound leaf form in *B. carolineifolia*, I genotyped the back cross population of (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* for *BARP1* genes based on Cleaved Amplified Polymorphic Sequence (CAPS) markers. Examination of *BARP1* sequences from *B. carolineifolia* and *B. heracleifolia* showed a restriction polymorphism for *Hpa*II (CCGG). *Hpa*II cleaved *B. carolineifolia* and *B. heracleifolia* products at different sites generating different sized restriction fragments (For *B. heracleifolia* the expected sizes of resctriction fragments are 556bp and 744bp and for *B. carolineifolia* the expected sizes are 110 bp and 1190bp) (Figure 93 B & C). PCR was carried out with *BARP*-36F and *BARP*1265R primers on DNA from 28 back cross plants and expected size (1300 bp) products were observed. The digested products were analysed by gel electrophoresis (Figure 93 D). I did get some restriction fragments of unexpected sizes which may be the result of recycling GFX PCR columns for several times (4 times) or due to the presence of primer dimers.

Box plots of dissection index (dissection index is calculated as described in chapter 3) from the back cross population of 28 individuals divided on the basis of *BARP1* showed no link between *BARP1* and compound leaf form in *B. carolineifolia* (Figure 94).

7.3. Association between *KNB1* and leaf dissection in (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia*

In order to test whether *KNB1* genes are associated with compound leaf form in *B. carolineifolia*, I genotyped a back cross population of (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* for *KNB1* genes through Cleaved Amplified Polymorphic Sequence (CAPS) markers. *KNB1* sequences of *B. carolineifolia* and *B. heracleifolia* (chapter 4) showed a restriction polymorphism for *Rsa*1 (GTAC). *Rsa*1 cleaved the *B. carolineifolia* fragment 4 times and *B. heracleifolia* 5 times (Figure 95 A) generating different sized products. 28F and 1087R primers were used to amplify PCR products from the DNA of back cross population. The digested products were analysed by gel electrophoresis (Figure 95 B). Box plots of dissection index (dissection index is calculated as described in chapter 3) from back cross population of 28 plants divided on the basis of *KNB1* genotype showed no link between *KNB1* and dissection index in *B. carolineifolia* (Figure 96).

7.4. Association between *KNB2* and leaf dissection in (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia*

To test the linkage of *KNB2* with compound leaf form in *B. carolineifolia*, I genotyped the back cross population of (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* based on Cleaved Amplified Polymorphic Sequence (CAPS) markers. *KNB2* sequences from *B. carolineifolia* X *B. heracleifolia* (chapter 4) revealed a restriction polymorphism for *Xba*1 (TCTAGA). *Xba*1 cleaved *B. carolineifolia* and *B. heracleifolia* at different sites generating different sized fragments revealing polymorphism (Figure 97 B). *KNB2*FGENO and *KNB2*RGENO primers were used to amplify PCR products from DNA of back cross progeny and digested with *Xba*1. The digested products were analysed by gel electrophoresis (Figure 97 C). Box plots of dissection index (dissection index is calculated as described in chapter 3) from the back cross population of 28 plants divided on the basis of *KNB2* genotype showed no link between *KNB2* and dissection index in *B. carolineifolia* (Figure 98).

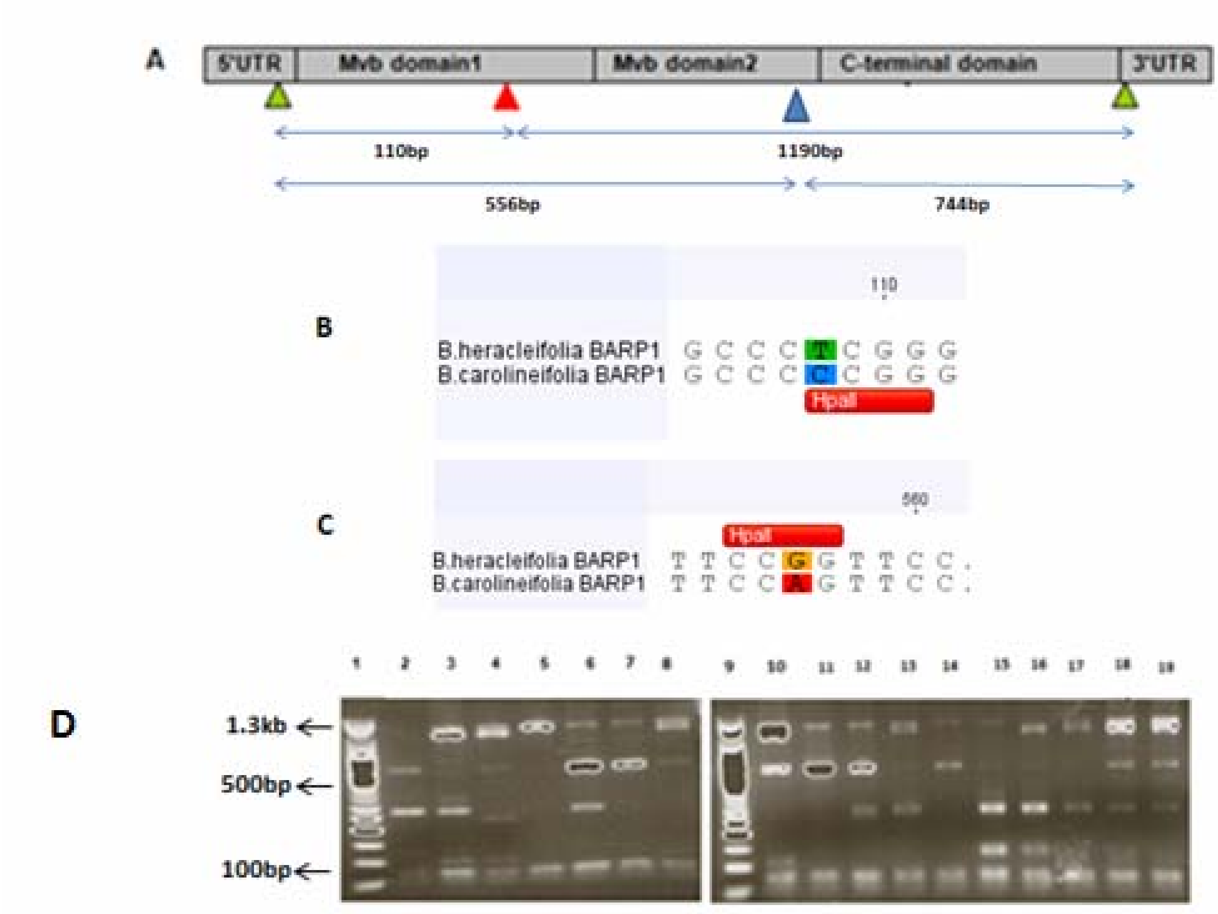


Figure 93 Genotyping for *BARP1.*
(A) Diagrammatic representation of BARP1 showing the positions of primers (green triangles) representing -36F and 1265R used to amplify PCR products and polymorphic sites (red and blue triangles). (B) & (C) Part of sequence alignment showing CAPS marker used to score for *BARP1* segregation. (D) Digestion of backcross population and parents showing polymorphism for *BARP1* Lane 1,9 50bp DNA ladder, Lane 2 *B. heracleifolia*, Lane 3 *B. carolineifolia*, Lane 4 to 19 some representatives from backcross population. The upper fragment may be representing the partially digested products. The unexpected cleaved fragments may be the primer dimers or some left over products from recycled GFX columns.

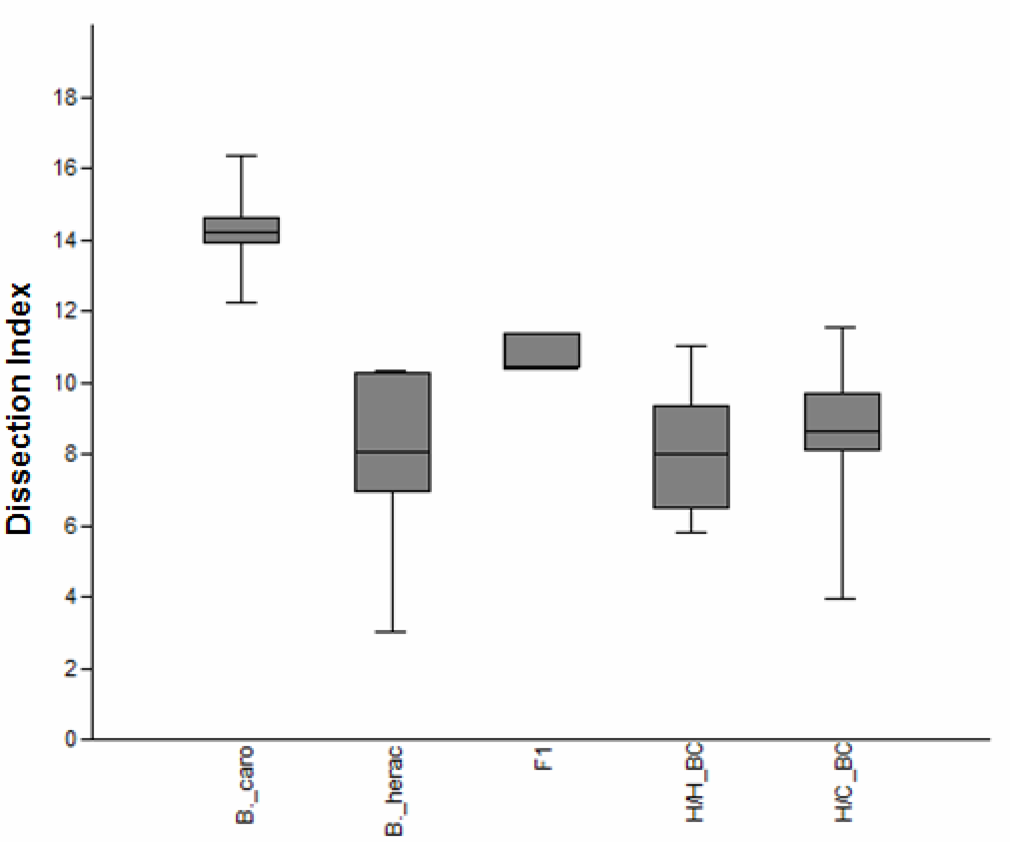


Figure 94 Association mapping for disection index in *B. carolineifolia* and *B. heracleifolia* for *BARP1.*

The graph shows the dissection index from 28 back cross plants of (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* depending on *BARP1* locus. The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median.

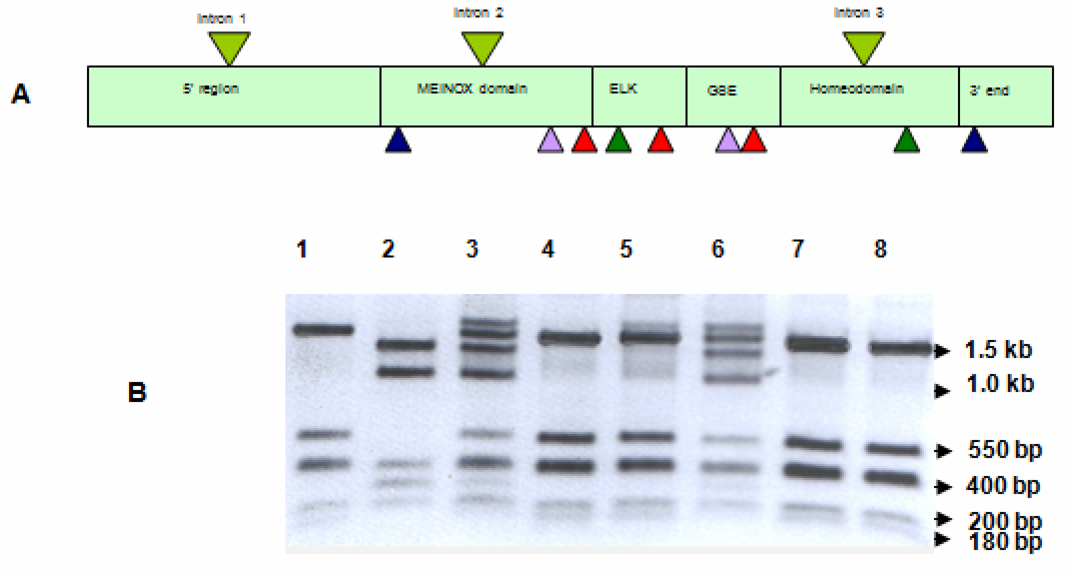


Figure 95 Genotyping for *KNB1.*

(A) Diagrammatic representation of *KNB1* gene showing positions of primers (blue triangles) used to amplify PCR products and polymorphic sites (red triangles show *Rsa1* sites for *B. heracleifolia*, light purple triangles for *B. carolineifolia* and dark green triangles for both *B. carolineifolia* and *B. heracleifolia;* (B) Digestion of backcross population and parents PCR fragments showing *KNB1* polymorphism. Lane 1 *B. heracleifolia* (expected sizes are 550bp, 400bp, 200bp, 200bp, 198bp and 180bp), Lane 2 *B. carolineifolia* (expected sizes are 400bp, 305bp, 300bp, 200bp and 190bp), Lane 3 to 7 some representatives from backcross population. Upper bands in lane 1 and lane 2 indicates the partially digested fragments.

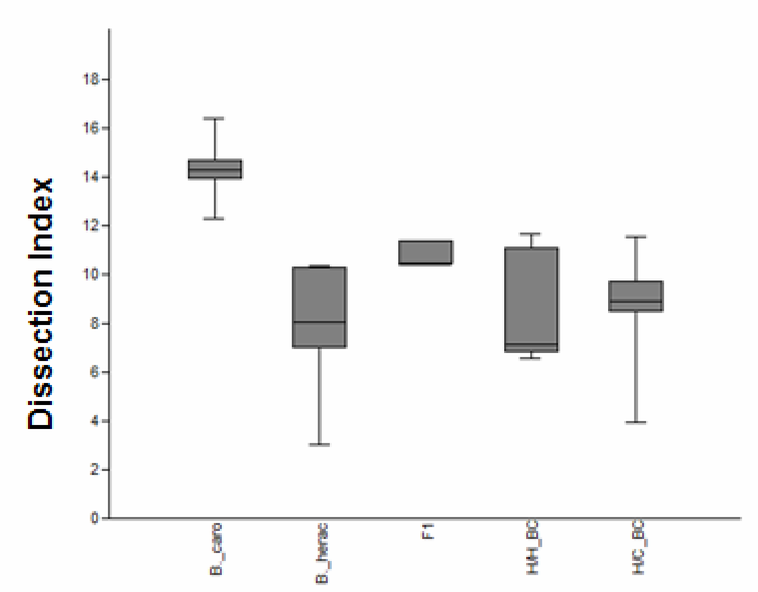


Figure 96 Association mapping for dissection index in *B. carolineifolia* and *B. heracleifolia* for *KNB1.*

The graph showing the dissection index from 28 back cross plants of (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* based on *KNB1* locus. The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median.

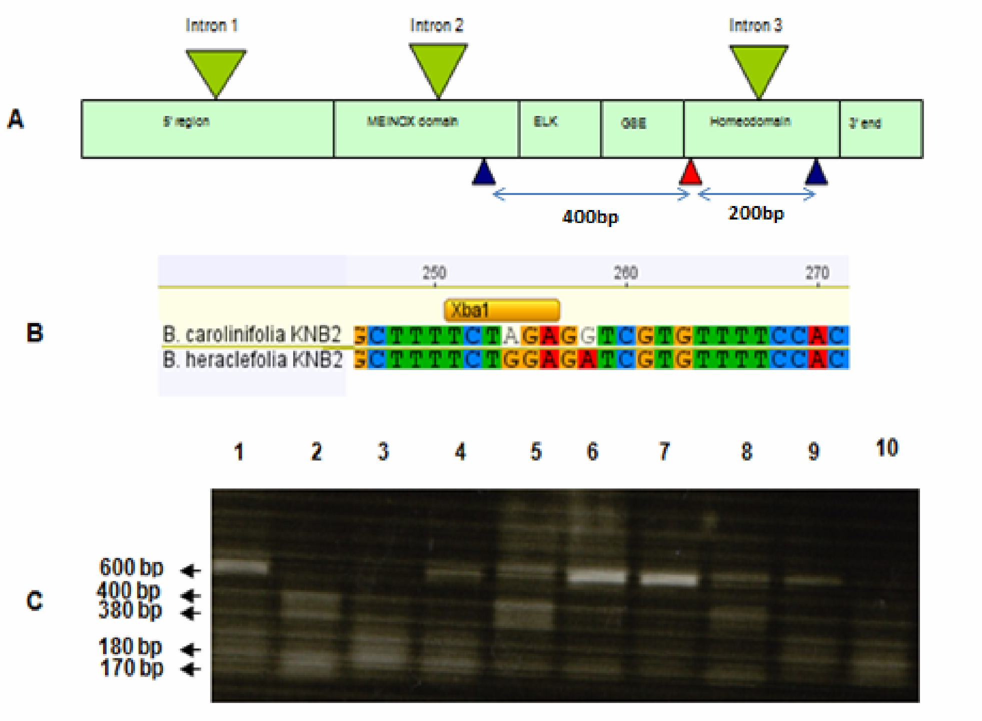


Figure 97 Genotyping *KNB2.*
(A) Diagrammatic representation of *KNB2* gene showing positions of primers (blue triangles) used to amplify PCR products and polymorphic site (red triangle). (B) Part of sequence alignment showing CAPS marker used to score *KNB2* segregation. (C) Digestion of backcross population and parents showing polymorphism for *KNB2*. Lane 1 *B. heracleifolia*, Lane 2 *B. carolineifolia*, Lane 3 to 10 some representatives from backcross population. Upper band may be the result of partial digestion; unexpected fragments may be primer dimers or some unknown fragments from recycled GFX Colums.

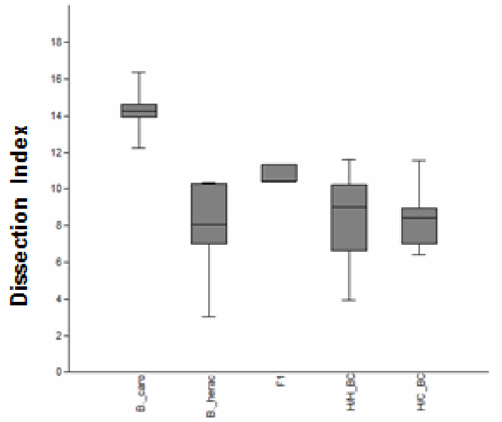


Figure 98 Association mapping for dissection index in *B. carolineifolia* and *B. heracleifolia* for *KNB2.*
The graph showing dissection index of 28 back cross plants of *B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* based on *KNB2* locus. The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median.

7.5. Association between *CUC* genes and leaf disection in (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia*

NAM/CUC3 (*NO APICAL MERISTEM* / *CUP-SHAPED COTYLEDON*) genes are expressed at leaflet initiation sites at the distal boundaries of leaf rachis in *Solanum lycopersicon, Cardamine hirsuta* and *Pisum sativum* and also regulate the expression of *KNOX/UFO* like genes through a feed forward regulatory loop during compound leaf formation in these species (Blein *et al*., 2008). To test *CUC* (*Cup-shaped Cotyledons*) like genes association with compound leaf form in *B. carolineifolia*, I amplified these genes by designing primers on *BNAM* genes of *B. plebeja* and *B. conchifolia* retrieved from transcriptomes sequences. I genotyped the backcross population of (*B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* for *BNAM* (*NAM* orthologs in *Begonia*) genes based on Cleaved Amplified Polymorphic Sequence (CAPS) markers. Sequencing the *BNAM* genes from *B. carolineifolia* and *B. heracleifolia* showed a restriction polymorphism for *Zra*1 (GACGTC).

*Zra*1 cleaved *B. heracleifolia* and does not digest *B. carolineifolia* PCR fragments (Figure 99 A). PCR was carried out with *CUC*55F and *CUC*560R primers on DNA of back cross population of 28 plants. The digested products were analysed by gel electrophoresis (Figure 99 B). Box plots of dissection index (Dissection index is calculated as described in chapter 3) from back cross population of 31 individuals divided on the basis of *BNAM* genes showed no link between these genes and compound leaf form in *B. carolineifolia* (Figure 100). However phylogenetic analysis has revealed that the *BNAM* genes I amplified from *B. carolineifolia* or *B. heracleifolia* were not *CUC3* orthologs but other *NAC* like genes (Figure 101).

7.6. Association between *BARP1* and leaf form in (*B. nelumbiifolia* X *B. heracleifolia*) X *B. nelumbiifolia*

In order to check whether the *BARP1* locus is linked to peltate leaf form in *Begonia* section Gireoudia, I genotyped a back cross population of (*B. nelumbiifolia* X *B. heracleifolia*) X *B. nelumbiifolia* for *BARP1* based on Cleaved Amplified Polymorphic

Sequence (CAPS) markers. *BARP1* sequences from *B. nelumbiifolia* and *B. heracleifolia* (chapter 3) revealed a restriction polymorphism for *BssS*1 (CTCGTG). *BssS*1 cleaved the *B. nelumbiifolia* fragment while *B. heracleifolia* did not have the restriction site for *BssS*1 (Figure 102 B). *BARP*-120F and *BARP1*245R primers were used to amplify PCR products from DNA of the back cross progeny and the products, cleaned up than digested with *BssS*1 and analysed by gel electrophoresis (Figure 102 C). Results showed that there is no link between *BARP1* and peltateness as all the back cross progeny were with mature peltate leaves and 13/20 plants were heterozygous at *BARP1* locus as was the non peltate F1.

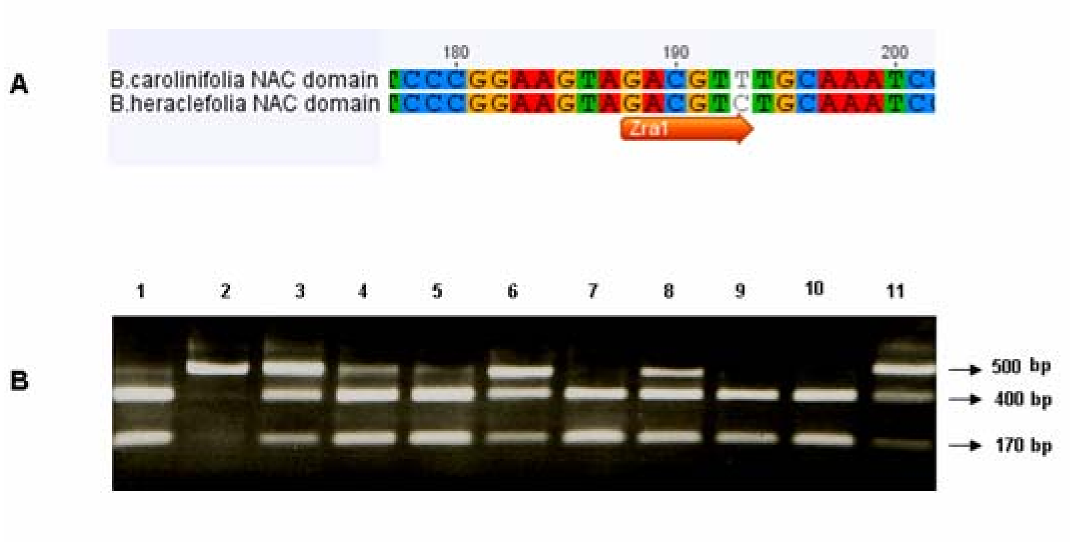


Figure 99 Genotyping *BNAM*.
(A) Part of sequence alignment showing CAPS marker used to score *BNAM* segregation. (B) Digestion of backcross population and parents showing polymorphism for *BNAM* genes. Lane 1 *B. heracleifolia*, Lane 2 *B. carolineifolia*, Lane 3 F1 hybrid and Lane 4 to 11 some representatives from backcross population.

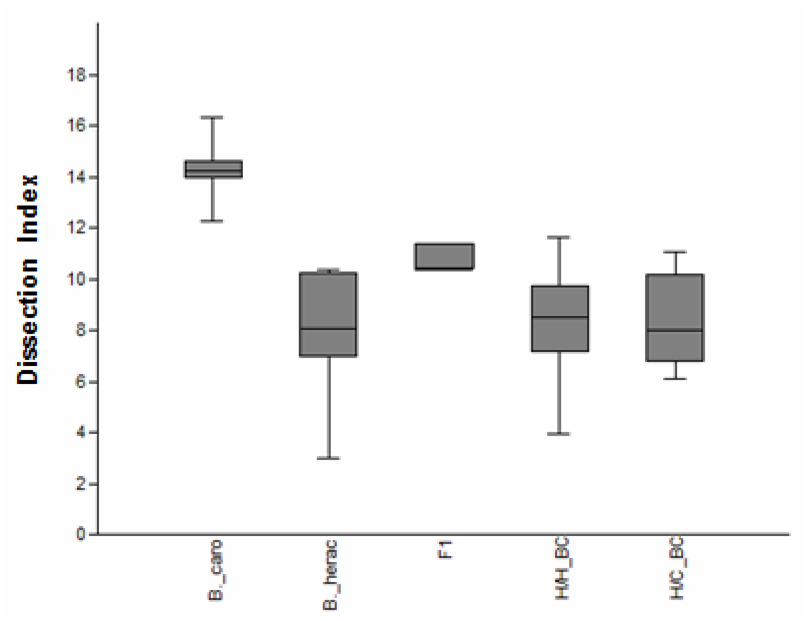


Figure 100 Association mapping for dissection index in *B. carolineifolia* and *B. heracleifolia* for *BNAM*.
The graph showing dissection index of 28 back cross plants of *B. carolineifolia* X *B. heracleifolia*) X *B. heracleifolia* based on NAC genes like locus. The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median.

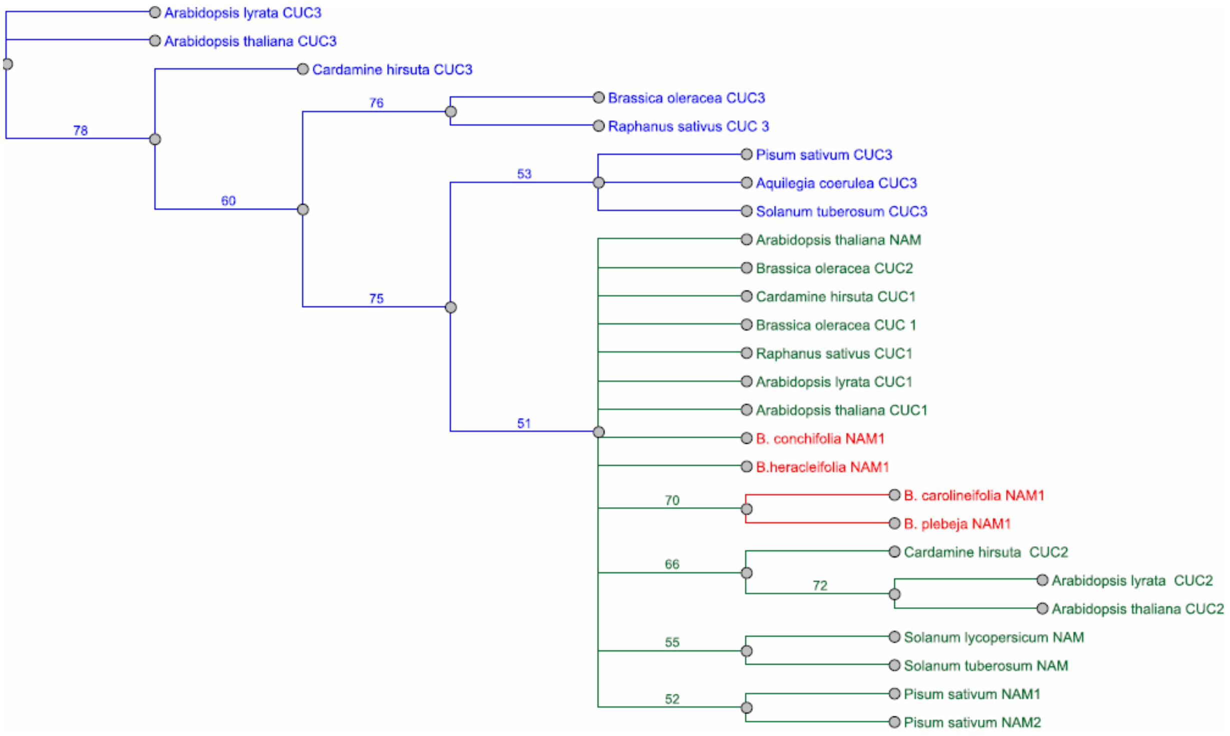


Figure 101 Phylogenetic relationship for *CUC*/*NAM* gene family.
Neighbour joining tree based on protein sequences for *CUC*/*NAM* genes. The alignment was carried out in Geneious align. The tree suggests that *NAC*-like genes studied are not *CUC*3 orthologs.

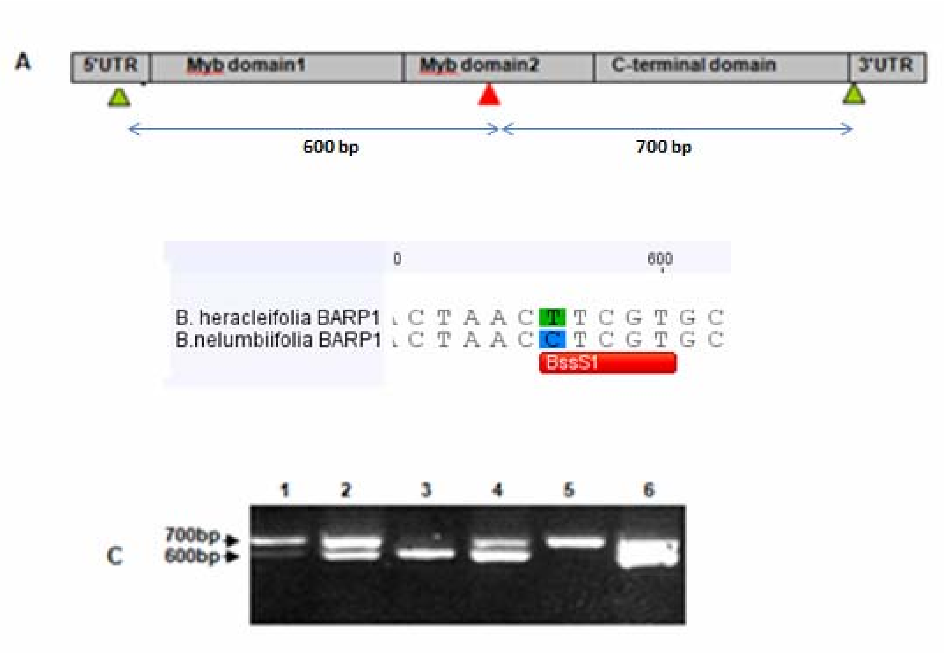


Figure 102 Genotyping for *BARP1.*
(A) Diagrammatic representation of BARP1 showing positions of primers (green triangles) used to amplify PCR products and polymorphic site (red triangle). (B) Part of sequence alignment showing CAPS marker used to score *BARP1* segregation (C) Digestion of backcross population and parents showing segregation for *BARP1*. Lane 5 *B. heracleifolia*, Lane 6 *B. nelumbiifolia*, Lane 4 F1 hybrid and Lane 1 to 3 some representatives from backcross population.

7.7. Association between *BARP1* and leaf form in (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* and (*B. conchifolia* X *B. plebeja*) X *B. plebeja*

To check the linkage between *BARP1* genes and peltate leaf form in *B. conchifolia*, Cleaved Amplified Polymorphic Sequence (CAPS) markers were used to genotype the back cross progeny of (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* (ARB 302). *BARP*-120F and *BARP1*245R primers were used to amplify PCR products from DNA of back cross plants. Sequences of *BARP1* from parents (chapter 4) revealed a restriction polymorphism for *Hph1* (TCACC). *Hph*1 cuts *B. plebeja* fragment but not *B. conchifolia* (Figure 103 B). The digested products were analysed by gel electrophoresis (Figure 103 C). 8/10 plants with peltate leaf form have c/c genotype and 9/11 plants with c/p genotype were non peltate. The two-tailed P value from Fisher exact test with null hypothesis “There is no difference in genotype of peltate and non peltate leaf form” equals 0.0089. Thus an association between genotypes and leaf forms is considered to be very statistically significant (table 23). Box plots of back cross population divided on the basis of the *BARP1* genotype also indicate a link between *BARP1* and peltate leaf form in *B. conchifolia* (Figure 104).

In another backcross progeny of 115 plants from (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* (ARB 312) and 129 plants from (*B. conchifolia* X *B. plebeja*) X *B. plebeja* (BOB 360) a link between *BARP1* genes and peltateness is seen. This link is established on the basis of the genotyping data carried out by KB Biosciences. All plants in back cross progeny of (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* are peltate and plants homozygous for *BARP1 B. conchifolia* allele are clearly more peltate than heterozygotes (Figure 105). In back cross plants of (*B. conchifolia* X *B. plebeja*) X *B. plebeja* all peltate leaved plants are heterozygous at *BARP1* (Figure 106 and table 24).

Table 23 Genotyping segregation for *BARP1* in ARB 302

	Peltate	Non Peltate	Total
CC	8	2	10
CP	2	9	11
Total	10	11	21

Back cross progeny of (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* (ARB302) segregation on the basis of *BARP1* CAPS marker screening. P-value for Fisher exact test is 0.0089 with null hypothesis "There is no difference in genotype of peltate and non peltate leaf form". This indicates that *BARP1* genes are linked with peltate leaf form in *B. conchifolia*.

Table 24 Genotyping segregation for *BARP1* in BOB 360

	PELTATE	NON PELTATE	TOTAL
PP	0	68	68
CP	12	49	61
	12	117	129

Back cross progeny of (*B. conchifolia* X *B. plebeja*) X *B. plebeja* (BOB 360) segregation on the basis of *BARP1* KB Biosciences data. Chi square test value 0.53 > 0.05 with null hypothesis "there is no significant difference between expected and observed values".

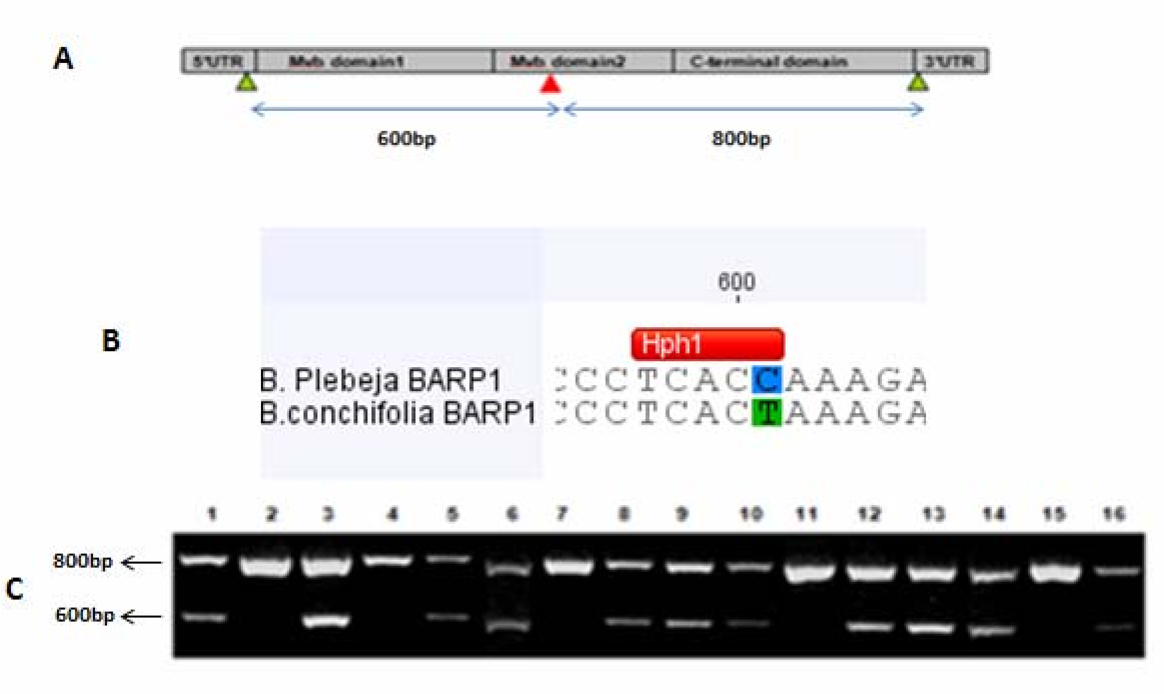


Figure 103 Genotyping for *BARP1* in ARB 302.
(A) Diagrammatic representation of BARP1 showing positions of primers (green triangles) used to amplify PCR products and polymorphic site (red triangle). (B) Part of sequence alignment showing CAPS marker used to score *BARP1* segregation. (C) Digestion of backcross population and parents showing polymorphism at *Hph*1 site. Lane 15 *B. conchifolia* Lane 16 *B. plebeja*, Lane 14 F1 hybrid, Lane 1-13 some representatives from backcross population.

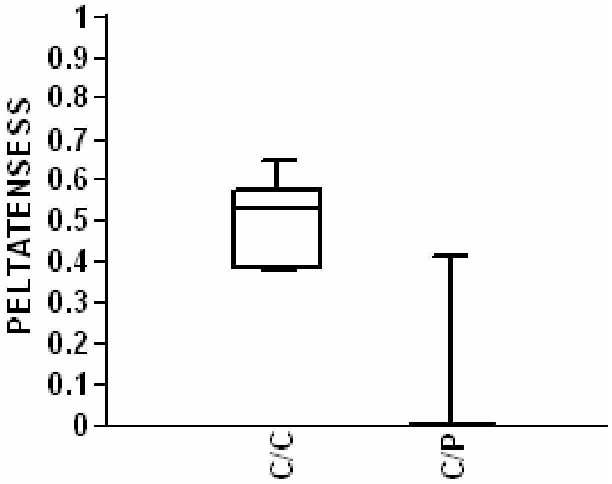


Figure 104 Association mapping for peltateness in *B. conchifolia* with *BARP1*.
The graph with standard deviation showing difference in the leaf form from 21 back cross plants of (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* (ARB 302) depending on *BARP1* locus. The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median. The graph shows a link between *BARP1* locus and peltateness.

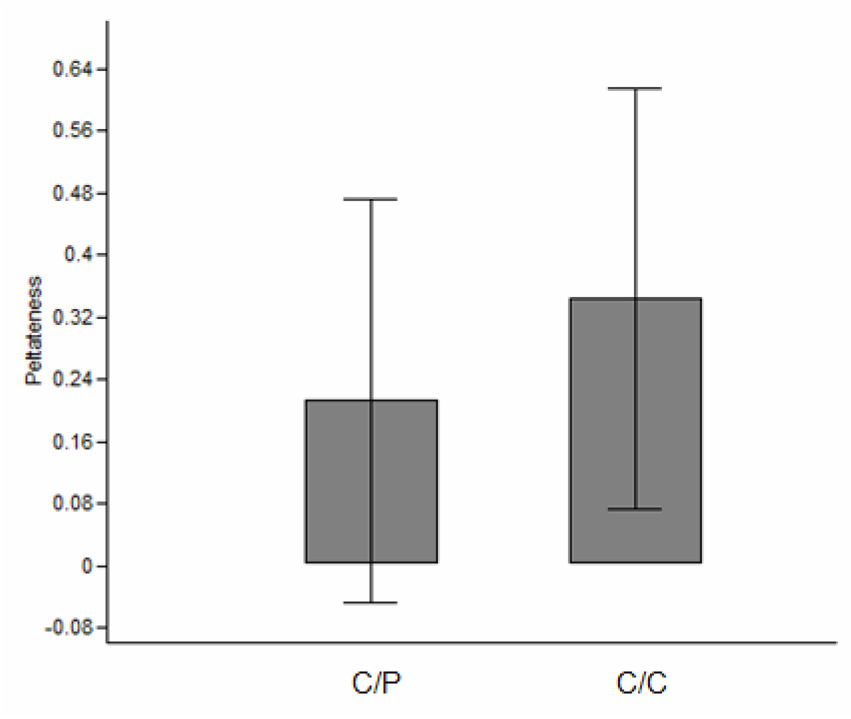


Figure 105 Association mapping for peltateness in *B. conchifolia* with *BARP1*.
The graph showing difference in the leaf form from 115 back cross plants of (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* depending on *BARP1* locus. The box plots represents 25, 75% quartiles. Whiskers represent minimum and maximum range, the middle line in boxes is median.The graph shows a link between *BARP1* locus and peltateness.

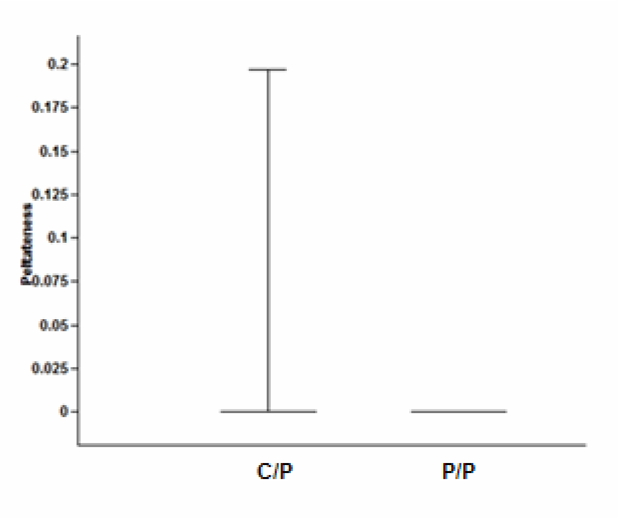


Figure 106 Association mapping for peltateness in *B. conchifolia* with *BARP1*.
The graph showing difference in the leaf form from 129 back cross plants of (*B. conchifolia* X *B. plebeja*) X *B. plebeja* depending on *BARP1* locus. The graph shows a link between *BARP1* locus and peltateness.

7.8. Association between *YABBY* genes and leaf form in (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* and (*B. conchifolia* X *B. plebeja*) X *B. plebeja*

YABBY genes were retrieved from transcriptomes sequences for *B. conchifolia* and *B. plebeja*. Phylogenetic analysis has revealed that we isolated a *YABBY3*/ *FILAMENTOUS FLOWER* ortholog (Figure 107). The genotyping was carried out on 129 plants of back crosses of (*B. conchifolia* X *B. plebeja*) X *B. plebeja* (BOB 360) by KB Biosciences which showed no link between *YABBY* genes and peltateness (Table 25). No link between *YABBY* genes and peltateness in *B. conchifolia* could be established in a back cross progeny of (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* (ARB312) as the F1 used to generate backcross progeny is homozygous at *YABBY* locus and so are the backcross plants.

Table 25 Genotyping segregation for *BYABBY* in BOB 360

	PELTATE	NON PELTATE	TOTAL
AA	5	57	62
AC	7	60	67
	12	117	129

Back cross progeny of (*B. conchifolia* X *B. plebeja*) X *B. plebeja* (BOB 360) segregation on the basis of *YABBY* CAPS marker screening. Chi square test value 0.69 > 0.05 with null hypothesis "there is no significant difference between expected and observed values".

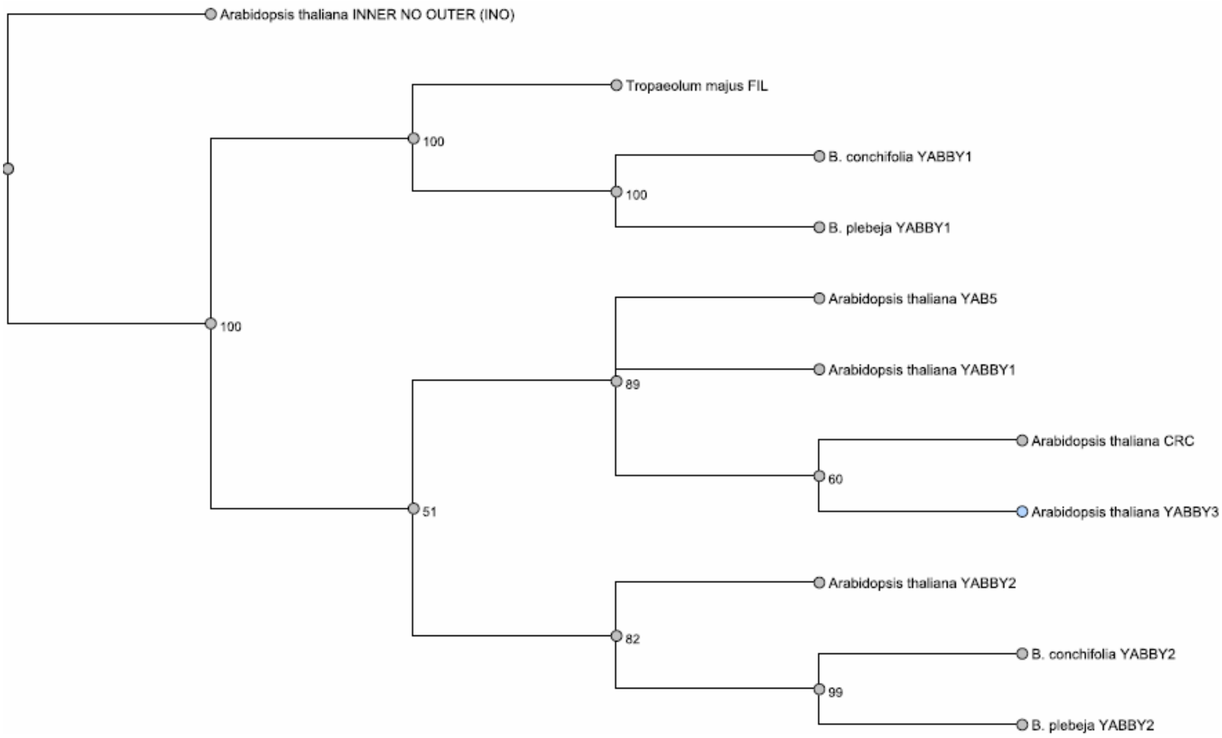


Figure 107 Phylogenetic relationship for *YABBY* gene family.
Neighbour joining tree with 1000 bootstraps based on protein sequences for *YABBY* genes. The alignment of sequences was carried out in Geneious align.

Note:

The genotyping of ARB 312 and BOB 360 for *BYABBY* and *BARP1* was carried out by KB Biosciences.

7.9. QTL analysis for Leaf shape traits

Quantitative trait locus (QTL) analysis is performed to reveal the genetic architecture underlying interspecific differences in leaf traits. Key genes can be missed by using a candidate gene approach but QTL studies have the advantage of screening the whole genome for association with the phenotype. Success depends on the density of the map, the linkage between markers and loci affecting the traits and the number and strength of the QTLs (Thornsberry *et al*., 2001; Zhao *et al*., 2007; Wei *et al*., 2006; Gonzalez-Martinez *et al*., 2006, 2007; Casa *et al*., 2008; Kraakman *et al*., 2006).

For performing QTL analysis previous knowledge of candidate genes is not necessary. However individual genes have been identified on the basis of QTL analysis in a very few cases. One such example is the identification of a gene controlling heading date in rice (Yano *et al*., 1997). Of the four genes predicted within the region of significant QTL, one gene (*CDS25*) was highly similar to the *Arabidopsis Flowering Time* (*FT*) gene. *FT* gene promotes flowering in *A. thaliana* under long day conditions. So it was selected as a likely candidate. Further investigations confirmed that this was the gene responsible for the trait (Kojima *et al*., 2002).

In second example a QTL for early flowering in short day conditions was mapped to the top of chromosome 1 in *A. thaliana* using the L*er* x Cvi RIL (Recombinant Inbred Lines) population. This QTL was fine mapped to a 45kb region representing 15 ORFs. One of these was *CRY2*, a gene known for its function in controlling the flowering time and perception of day length (El-Assal *et al*., 2001). Further experiments including complementation tests and sequencing this gene showed a single nucleotide substitution in the Cape Verdi Islands accession of *A. thaliana*. In this case QTL helped identifying new allele of the *CRY2* gene controlling this trait.

Complex traits are, by definition, controlled by a large number of loci (Barton and Turell 1989). QTL analysis is a powerful method of identifying the genomic location of some of the more important loci regulating a complex trait. In this way a few key loci can be identified even for a very complex trait (Flint and Mott 2001)

I have performed QTL analysis for leaf shape traits in mapping populations of (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* (ARB312) and (*B. conchifolia* X *B. plebeja*) X *B. plebeja* (BOB360). Genetic analysis performed on these populations is described in chapter 3 which indicated that mapping populations exhibit variation for leaf shape traits. For QTL analysis the genetic map of *Begonia* (Brennan *et al*., unpublished data) based on CAPS and AFLPs markers is used. Marker loci are polymorphic, evenly spread and provide good coverage of the genome.

Single marker Regression QTL and Composite Interval Mapping for peltateness, distance from petiole attachment point to boundary (Pet_b) and notchiness in the mapping population of backcrosses progeny (*B. conchifolia* X *B. plebeja*) X *B. conchifolia* (ARB 312) was carried out using QGene software.

CIM (composite interval mapping) QTL identified two significant QTLs associated with peltateness, located on chromosome 8 and 13, the most significant of which appeared near/at *BARP1* on chromosome 13 (Figure 108) at 40.8 cM. This QTL has LOD score 5.93 of which accounts for 21.1% of the total variation for peltateness. Collectively these two QTLs accounted for 37.5% of the total variation for this trait (Table 26). The same peak was identified in a Single marker Regression analysis.

CIM analysis established that notchiness was governed by two QTLs on chromosome 8 and 13. The most significant of which appeared near *BARP1* on chromosome 13 (Figure 108). This QTL has LOD score 4.7 and accounts for 17.3% of total variation for this trait. Collectively these two QTLs accounted for 24.3% of the total variation for this trait (Table 26).

CIM analysis established that distance from petiole attachment point to boundary (Pet_b) was controlled by two QTLs on chromosome 6 and 13. The most significant of which appeared near / at *BARP1* on chromosome 13 (Figure 109). This QTL has LOD score 4.826 and accounts for 17.6% of total variation for this trait. Collectively these two QTLs accounted for 32.2% of the total variation for this trait.

A repeat of the QTL analysis in the other back cross direction (The progeny of (*B. conchifolia* X *B. plebeja*) X *B. plebeja*, BOB360) showed that none of the leaf traits (peltateness, leaf circularity, leaf eccentricity, notchiness, leaf length, leaf width, distance from petiole attachment point to boundary and distance from petiole attachment point to centre) gave significant LOD scores near the *BARP1* locus (see appendix B).

However in this progeny very few plants were peltate. The genotyping data is available for 12/22 peltate plants in a progeny of 274 plants which all are heterozygous at *BARP1* locus. However 62 plants heterozygous at *BARP1* locus are non peltate. This indicates the presence of modifiers from *B. plebeja* to modify the trait. Peltateness has significant QTL on chromosome 6 at 37.3 cM and 65.2 cM where no known candidate leaf trait markers are present (Figure 111-112). This QTL explains 1.1% variation for this trait.

In order to see whether *BARP1* is controlling any other aspect of leaf shape I have performed QTL analysis using leaf shape principle component analysis (PCA) in PAST for BOB360. None of the plain or Elliptic Fourier PCA showed significant QTL at or near *BARP1* locus (see appendix B). Landmark analysis placed 50 points evenly around the outline of each leaf. These co-ordinates were analysed using either shape PCA analysis or fourier analysis followed by PCA using the programme PAST.

Table 26 QTL for peltateness, notchiness and distance from petiole to leaf boundary for ARP312 determined by CIM

Trait		**Chromosome**	**Locus/Position cM**	**CIM LOD**	**% of variance explained**
Peltateness	QTL1	13	BARP/40.8	5.93	21.1%
	QTL2	8	43.8-56.8	4.487	16.4%
Notchiness	QTL1	13	BARP/40.8	4.747	17.3%
	QTL2	8	39.5-41.2	1.927	7%
Pet_b	QTL1	13	BARP/40.8	4.826	17.6%
	QTL2	6	37.3-65.2	4.062	15%

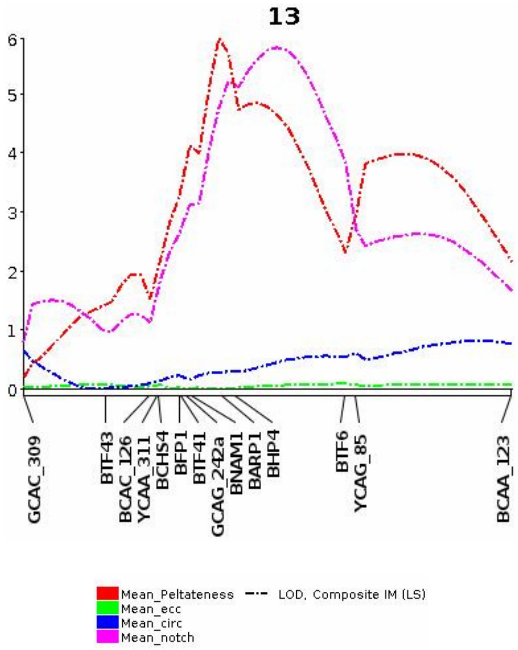


Figure 108 CIM for ARB312 for leaf peltateness, eccentricity, notchiness and circularity.
The graph indicates that peltateness and notchiness may be controlled by *BARP1* locus or locus/loci controlling these traits are very close to it on chromosome 13 (CIM for all 14 chromosomes is presented in appendix B).

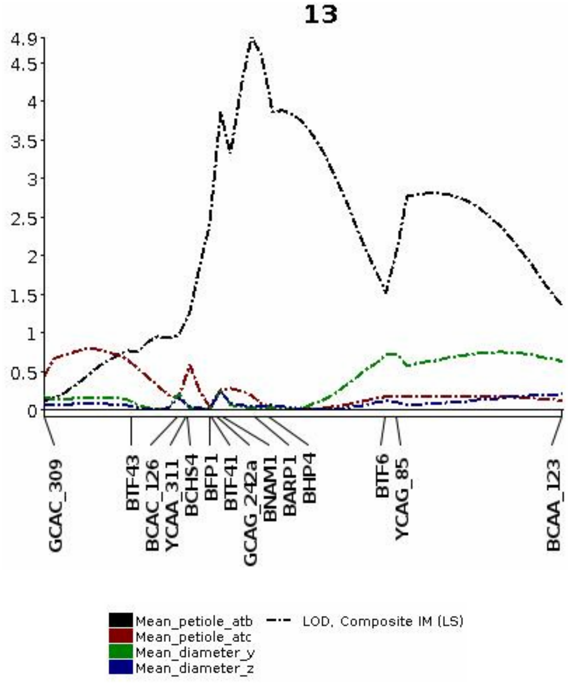


Figure 109 CIM for ARB312 for distance from petiole attachment point to leaf boundary, petiole attachment point to centre, leaf length and leaf width.
The graph indicates that distance from petiole attachment point to leaf boundary may be under control of *BARP1* locus or locus/loci controlling these traits are very close to it on chromosome 13 (CIM for all 14 chromosomes is presented in appendix B)

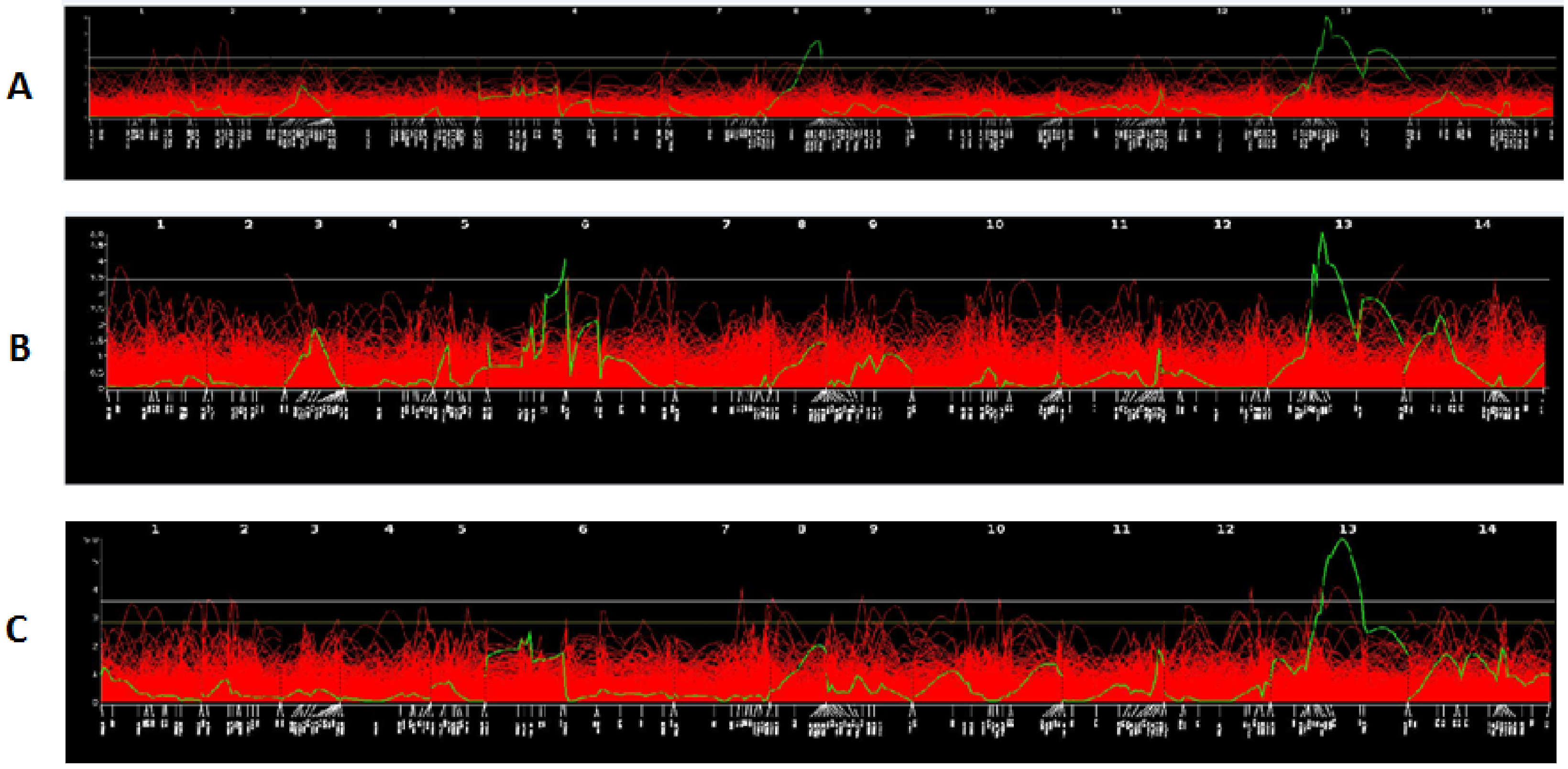
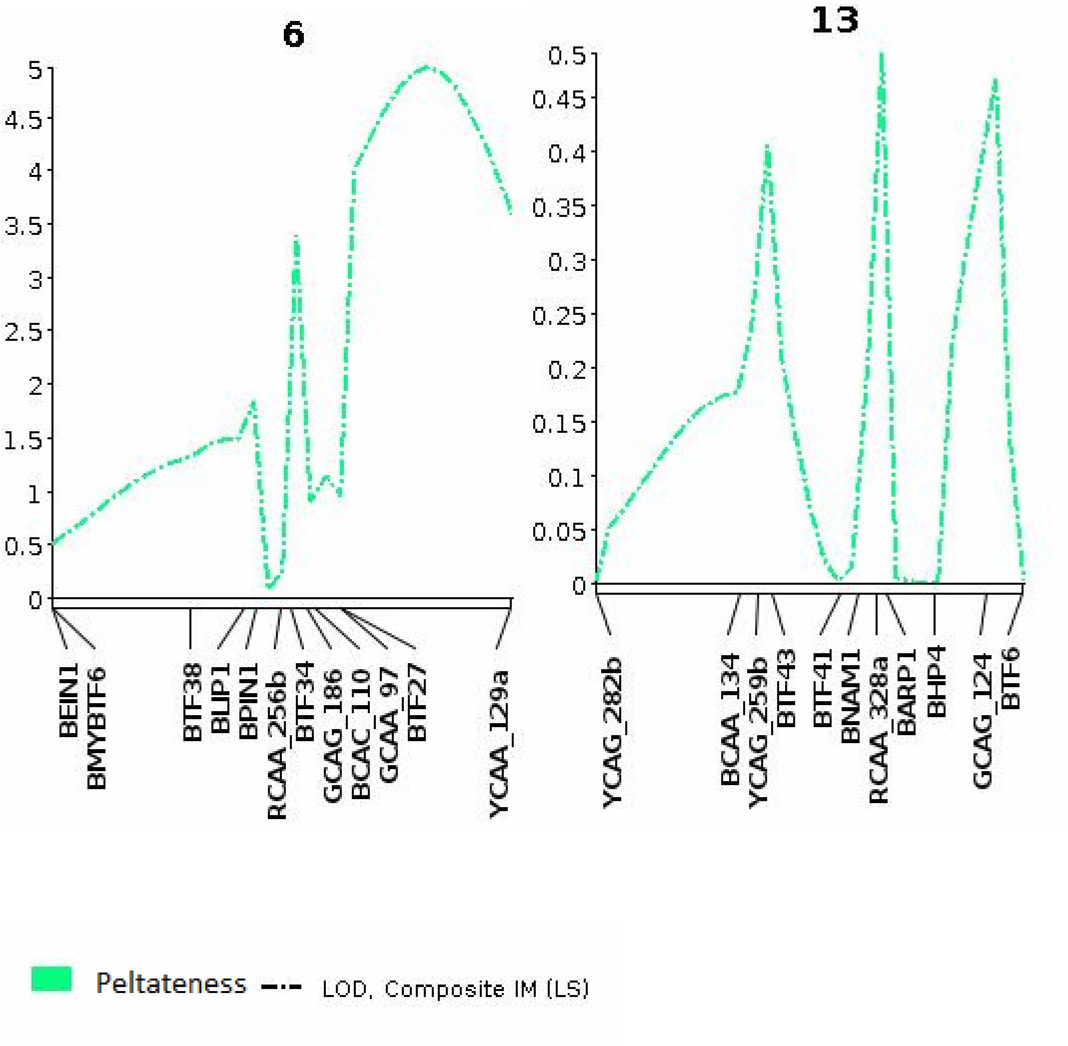


Figure 110 Composite Interval Mapping on ARB312 for peltateness (A), distance from petiole attachment point to leaf boundary (B) and notchiness (C) with 1000 permutation.

Figure 111 CIM for BOB 360 for peltateness. The graph indicates that peltateness is not linked to *BARP1* in this population but may have significant QTL on chromosome6 .



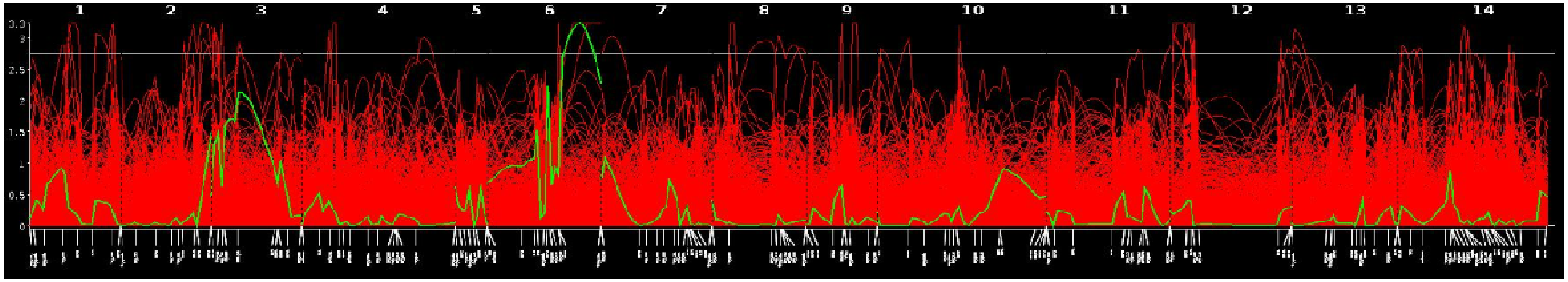


Figure 112 Composite Interval Mapping for peltateness on ARB312 with 1000 permutations.

7.10. Discussion

To date all studies reporting a role of *ARP* like genes for regulating peltateness are based on expression data, transgenics or mutant phenotypes (Kim *et al*., 2003; Waites and Hudson, 1995). For the first time we have showed a direct link of an *ARP1* locus with the recent evolution of peltate leaf form in a lineage. QTL analysis is a statistically robust and reliable method to establish a link between phenotypes and genotypes. The association of *BARP1* with a peltate leaf form in *B. conchifolia* is established in back cross progenies of *B. plebeja* X *B. conchifolia*.

BARP1 is not associated with peltate leaf form in *B. nelumbiifolia*. However I had a small number of plants in a back cross population of (*B. nelumbiifolia* X *B. heracleifolia*) X *B. nelumbiifolia*. This population did not show segregation for peltate leaf form which restricts the possibility of establishing a link between *BARP1* and peltate leaf form in this species. There are very few example of identifying the specific gene controlling the trait of interest in QTL region. This requires identifying a candidate gene located in the region of QTL whose function has been well characterized in other species or exhaustive position based cloning and transgenic analysis (Perez-Perez *et al*., 2002; Yano *et al*., 1997; El-Assal *et al*., 2001). QTL mapping has shown a significant peak on chromosome 13 where *BARP1* is located which indicated that a link between peltate leaf form and *BARP1* in *B. conchifolia*.

FILAMENTOUS FLOWER (member of *YABBY* gene family) is associated with the development of a peltate leaf form in *Tropaeolum majus* by prefiguring the position of lamina outgrowth (Gleissberg *et al*., 2005). *FILAMENTOUS FLOWER* orthologs in *Begonia* are not linked with peltate leaf form in *Begonia conchifolia* as no association between *YABBY* genes and peltate leaf form in *B. conchifolia* has been detected.

However *FILAMENTOUS FLOWER* orthologs in *Begonia* may be good candidates for controlling peltate leaf form in other *Begonia* species in which peltateness has originated independently. This could be tested by performing association mapping studies in a fairly large population of back cross plants or F2 plants with other peltate species in *Begonia*.

BARP1, *KNB1* and *KNB2* genes are not associated with leaf dissection in *Begonia* section Gireoudia species. The *KNOX* gene family is divided into two classes. Each class has undergone duplication multiple times and comprises of many genes (Hake *et al*., 2004; Hay and Tsiantis, 2010, Mukherjee *et al*., 2009). A third class has been reported in *KNOX* genes family recently. This class lacks Homeodomain and has been characterized in *Arabidopsis* (*KNATM*) and tomato (*PTS*) (Magnani and Hake, 2008; Kimura *et al*., 2008). I have amplified only *STM* like *KNOX* genes from *Begonia*. *BP* (*BREVIPEDICELLUS*) like class 1 *KNOX* genes control natural variation in leaf form between *Arabidopsis* and *C. hirsuta* and *PTS* over expression in *Solanum cheesmaniae* is the cause of super compoundness of this species as compare to *Solanum galapagense* (Hay and Tsiantis, 2006; Kimura *et al*., 2008). Studying other *KNOX* like genes in *Begonia* might therefore be worthwhile to study the evolution of compound leaf form in this genus.

Other genes as well as *ARP* and *KNOX* affect the leaf form in other species. In pea (*Pisum sativum*) *UNIFOLIATA* (*UNI)* is required for compound leaf formation (Hofer *et al*., 1997). Leaflet formation in compound leaves of *Solanam lycopersicon, Cardamine hirsuta* and *Pisum sativum* is controlled by feed forward regulatory loop between *NAM/CUC3* genes and *KNOX/UNI* genes (Blein *et al*., 2008). I could not amplify the orthologue of *NAM/CUC3* genes from *Begonia* species instead I got other *CUC*-like genes, which showed no association with the leaf disection index in *Begonia*. The *CUC3* genes may be good candidates to study the evolution of leaf dissection in *Begonia.*

7.11. Conclusion

BARP1, *KNB1* and *KNB2* and *BNAM* are screened for association with peltateness and leaf dissection in back cross populations of *Begonia* section Gireoudia using CAPS markers. *BARP1* is not linked with peltateness in *B. nelumbiifolia* but may be linked with peltateness in *B. conchifolia*. There is no genetic link between disection index and *BARP1*, *KNB1, KNB2* or *NAC* like genes in *B. carolineifolia*. Other candidate genes like *BP*, *PTS* and *BARP2* are worth considering as candidate for leaf form in *Begonia*.

Chapter 8

General discussion

8.1. Introduction

Leaves have evolved multiple times in the lineages of land plants (Tomescu, 2008; Singer and Ashton, 2007; Cronk, 2001). Though leaf-like structures exist in moss and liverworts they are anatomically different from tracheophyte leaves (Tomescu, 2008). Leaves also evolved independently in lycophyte and euphyllophyte lineages as their common ancestors were leafless, (Kenrick and Crane, 1997; Harrison *et al*., 2005). Euphylophytes leaves are assumed to have originated from bifacial determinate lateral branch complexes and those of lycophytes from progressive vascularisation of surface emergences from the stem or from sporangia (Boyce and Knoll, 2002; Tomescu, 2008).

Interestingly both lycophyte and euphyllophyte leaves have recruited the same genetic mechanism to control early leaf development (Harrison *et al*., 2005). *KNOX* genes maintain indeterminacy in the shoot apical meristem and are repressed by *ARP* genes in the leaf primordia in both seed plants and lycophytes (Harrison *et al*., 2005). *ARP* genes are therefore responsible for determinacy which is a prerequisite for the identity of leaf primordium founder cells at the periphery of the apical meristem. The *KNOX-ARP* interaction appears to be a common denominator of leaf development in all extant tracheophytes (Tomescue, 2008). Besides this *KNOX* and *ARP* genes have been reported as regulating leaf form in different seed plants (Hay and Tsiantis, 2006; Kimura *et al*., 2008). *KNOX* genes are permanently repressed in the leaf primordia of simple leaved plants but are up regulated after initial repression in compound leaved plants (Harevan *et al*., 1996; Hay & Tsiantis, 2006; Bharathan *et al*., 2002; Shani *et al*., 2009). The level of *KNOX* expression in the leaf primordia is positively correlated with the complexity of compoundness (Kimura *et al*., 2008). Variation in expression pattern of *ARP* genes is associated with the development of peltate leaves through the regulation of lamina growth (Kim *et al*., 2003).

The purpose of my studies was to determine the genetics underlying the leaf form variation in *Begonia* section Gireoudia. In order to do this I intended to find out whether leaf form is a genetically simple trait or whether many loci contribute to the differences between species. *ARP* and class 1 *KNOX* like genes have been well characterized as regulating the natural variation of leaf form between *Arabidopsis* and *C. hirsuta*, and between different *Solanum* species (Hay and Tsiantis, 2006; Kimura *et al*., 2008). I aimed to discover whether *ARP* and *KNOX* genes were genetically associated with leaf form variation between different *Begonia* species.

8.2. Leaf form is a complex genetic trait in *Begonia*

Peltateness has independently evolved five times in *Begonia* section Gireoudia (Harrison and Kidner, unpublished data). To test whether the same mechanism was recruited each time complementation tests were performed. All the F1 hybrids between peltate species have a peltate leaf form. These results indicate that either peltateness is a dominant trait or the same gene is regulating peltateness in different species. To investigate this five peltate species were crossed with twelve non peltate species and vice versa. Some reciprocal crosses generate the same phenotypes indicating lack of involvement of cytoplasmic factors in leaf form determination. Peltateness behaved as a recessive, dominant or semi dominant trait in these crosses. Furthermore the quantification of peltateness suggests that peltateness is a continuous trait. This suggests the trait is complex. The segregating ratios in back cross progenies also support the involvement of more than two loci for controlling the peltate leaf form. Some F1s (chapter 3) though initially peltate produce non-peltate leaves as they mature indicating a switch in the regulation of leaf developmental genes. The fact that peltateness does not complement in *Begonia* species and behaved differently in different crosses may be indicative of the involvement of one major locus with many loci of small effect or some factors from non peltate species may be responsible for modifying peltateness.

The genetics of leaf dissection has been studied extensively by plant biologists (Table 1). These studies have shown that leaf dissection was usually dominant over entire leaves and highly dissected leaves were semi dominant over less dissected leaves. Involvement of one or two genes was reported for controlling the leaf disection. Studying the genetics of compound leaf form in *Begonia* was a target of my studies. For F1 hybrids the disection index was midway between both parent's phenotypes. In a back cross the disection indexes varies significantly and some backcross progeny were less dissected than both parents. This indicates that compoundness is a polygenic trait in *Begonia*, with species containing both loci that promote and that prevent leaf dissection.

8.3. *ARP* genes are linked to the species level leaf form variation in *Begonia conchifolia* but not in all *Begonia* species

A link between *ARP* genes and peltate leaves has been predicted based on expression analysis data in *S. actinophylla*, *P. aquatica*, *A. pentaphylla* and *O. regnellii* (Kim *et al.*, 2003). *ARP* genes are mostly expressed in the adaxial domain of simple leaved species and their expression is confined to the distal tip of the peltate leaved species (Kim *et al.*, 2003; Waites and Hudson, 1995). *ARP* gene expression also predicts the sites of leaflet placement in a range of compound leaved species (Kim *et al.*, 2003). *BARP1* genes (*ARP* orthologs in *Begonia*) are expressed on the distal tips of the developing leaf primordia in all *Begonia* species I studied. However close examination of leaf-petiole attachment points and vascular bundle patterning in petioles has revealed that all *Begonia* section Gireoudia species are peltate to some degree, so the link between *ARP* expression pattern and leaf form postulated by Kim *et al* (2003) is not broken in section Gireoudia

Association mapping studies have shown a link between *BARP1* and peltateness in *B. conchifolia*. However there was no association detected between peltateness in *B.*

nelumbiifolia and the *BARP1* locus. I have the back cross of F1 hybrid of *B. nelumbiifolia* X *B. heracleifolia* with one parent only (*B. nelumbiifolia*). The number of these back cross plants was too small to detect a link and the plants in back cross population did not show segregation for peltateness either. Further both *B. conchifolia* and *B. nelumbiifolia* have evolved peltateness independently and *BARP1* may not be a key regulator for controlling peltateness in both species.

No direct genetic link between peltate leaf form and *ARP* genes has been shown to date. On the basis of QTL analysis we have shown for the first time that peltate leaf form in *B. conchifolia* is controlled by *BARP1* locus or a gene that is linked with it.

8.4. *YABBY* genes are not linked with peltate leaf form in *Begonia*

Tropaeolum majus FILAMENTOUS FLOWER (*YABBY*) genes are suggested to control the peltate leaf form in *Tropaeolum majus* by regulating the position of lamina outgrowth (Gleissberg *et al*., 2005). *YABBY* genes are expressed at the base of leaf primordia in this species. However all the close relatives of *Tropaeolum majus* are also peltate which makes it impossible to use genetics to determine if changes at a *YABBY* locus are responsible for the leaf form in this genus?

A *FILAMENTOUS FLOWER* ortholog in *Begonia* was not linked with peltate leaf form as no association between a *Begonia YABBY* gene and peltate leaf form in *B. conchifolia* has been detected. However they are good candidates to study peltate leaf form in other *Begonia* species as peltateness has evolved independently in different *Begonia* species. This could be tested by performing association mapping studies in a fairly large population of back cross plants or F2 plants with other peltate species in *Begonia*. I presume that there would be no change in the expression pattern of *YABBY* genes between different *Gireoudia* species because the petioles of all *Begonia* of section Gireoudia are abaxialized and *YABBY* genes define the abaxial cell fate during leaf development (Bowman 2000; Golz *et al*., 2004).

8.5. *KNOX* genes are good candiate for compound leaf form

KNOX genes play a fundamental role in establishing indeterminacy by suppressing differentiation and *KNOX* genes have been reported to regulate compound leaf form in different species (Kerstetter *et al*., 1997; Long *et al*., 1996; Vollbrecht *et al*., 2000; Champagne *et al*., 2007, Chen *et al*., 1997; Harvean *et al*., 1996; Hay and Tsiantis, 2006; Kimura *et al*., 2008; Muller *et al*., 2006; Shani *et al*., 2009). *KNOX* genes have been shown to control species level variation in leaf form in *Arabidopsis* and *Cardamine hirsuta* and, *Solanum galapagense* and *Solanum cheesmaniae* (Hay & Tsiantis, 2006; Kimura *et al*., 2008). *STM* like *KNOX* genes are duplicated in *Begonia* and neither is associated with dissection index in compound leaves of *B. carolineifolia*. However the back crosses population was very small and I have back cross plants with one parent only (*B. heracleifolia*) which further restricts the possibility of revealing a robust analysis. A fairly large number of back cross population with both parents should be generated to perform association mapping studies for *KNB1* and *KNB2*.

At the expression level there was no detectable difference in expression of *KNB1* genes between different leaf forms. I can not conclude anything about *KNB2* expression patterns as the *KNB2* probe requires further optimization.

A back cross or F2 population of simple leaved *Begonia* species with each of *B. carolineifolia* and *B. thiemei* should be generated to carry out a complete survey for the evolution of compound leaf form in *Begonia* section Gireoudia. Both compound leaved species has evolved independently in this section and it would be interesting to see whether *KNOX* genes are the key regulators for complex leaf morphology in either species.

This could be done by carrying out association mapping on a fairly large population of F2 or BC1 with both *B. carolineifolia* and *B. thiemei* to infer a link of *KNB1* or *KNB2* with compound leaf form.

The *KNOX* gene family is divided into two classes. Class 1 *KNOX* gene family has four genes in *Arabidopsis*, nine in poplar, nine in maize, eight in rice, three in lycophytes and three in mosses (Hay and Tsiantis, 2010). Class 2 *KNOX* gene family has four genes in *Arabidopsis*, six in poplar, four in maize, four in rice, two in lycophytes, two in mosses and one in green algae (Hay and Tsiantis, 2010). A third class may exist in *KNOX* genes family. This class lacks the Homeodomain and has been obtained from *Arabidopsis* (*KNATM*), poplar (*KNATM*) and tomato (*PTS*) (Magnani and Hake, 2008; Kimura *et al*., 2008).

I have amplified only *STM* like *KNOX* genes from *Begonia*. It may be worth amplifying other *KNOX* like genes from *Begonia* to study the evolution of compound leaf form as *BP* like class 1 *KNOX* genes are responsible for natural variation in leaf form between *Arabidopsis* and *C. hirsuta* and over expression of *PTS* is the cause of super compoundness of *Solanum cheesmaniae* as compare to *Solanum galapagense* (Hay and Tsiantis, 2006; Kimura *et al*., 2008).

8.6. More Candidate genes for compound leaf form

A *KNOX*-independent mechanism for controlling compound leaf form exists in pea (*Pisum sativum*). *UNIFOLIATA* (*UNI*), an ortholog of the floral regulators *FLORICAULA* (*FLO*) from *Antirrhinum majus* and *LEAFY* (*LFY*) from *Arabidopsis thaliana* regulates compound leaf development in this species (Hofer *et al*., 1997; Gourlay *et al*., 2000; Hofer *et al*., 2001; Champagne *et al*., 2007).

The *NO APICAL MERISTEM/ CUP-SHAPED COTYLEDONS3* (*NAM/CUC3*) family regulates the proper expression of *KNOX/UFO* like genes during compound leaf formation in both the angiosperms and legumes. Leaflet formation in compound leaves of *Solanam lycopersicon, Cardamine hirsuta* and *Pisum sativum* is under control of a feed forward regulatory loop between *NAM/CUC3* genes and *KNOX/UNI* (Blein *et al*., 2008). Further *miR164* controls the placement of leaflets and degree of lobe or serration outgrowth in a range of species by regulating the expression patterns (Nikovics *et al*., 2006; Blein *et al*., 2008; Berger *et al*., 2009).

We isolated the sequence of one *CUC* like gene from *B. conchifolia* and *B. plebeja* through 454 sequencing. I carried out the association mapping to look for links between these *CUC* like genes and compound leaf formation in *B. carolineifolia* X *B. heracleifolia* backcross populations and there was no association detected. Phylogenetic analysis revealed that the *CUC* like genes I amplified from *B. carolineifolia* and *B. heracleifolia* are not orthologs of *CUC3* genes. Instead they are orthologs of *CUC1*/*CUC2* like genes.

Compound leaves have evolved independently mutiple times in plant history. However the development of compound leaves has been studied in very few lineages. Expanding the compound leaf development studies to include more plant lineages will shed light on the mechanisms of its development. Analyzing the expression patterns of *NAM/CUC3* or *miR164* may provide a good explanation for the evolution of compound leaf development in species where there is no variation in *KNOX* genes such as in *Begonia*.

8.7. Duplicated *ARP* and *KNOX* genes provide the possibility of novelty

Phylogenetic analysis of duplicated *ARP* like genes from Ruschioideae has shown that rapid radiation of plant form in Ruschioideae is correlated with the duplication and subsequent gene loss of *ARP* genes (Illing *et al*., 2009). Orthologs of *ARP* and *KNOX* genes from *Begonia* were analysed through phylogenetic studies. Both *ARP* and *STM*-like *KNOX* genes are duplicated in *Begonia*. The *ARP* nucleic acid and protein sequences were not grouped on the basis of final leaf morphology. Rather *BARP* genes formed a clade with *ARP* orthologs from cucumber which shows that the gene tree was congruent with the known species phylogeny. Phylogenetic analysis was not able to predict the function of *ARP* and *KNOX* genes in different species (Kim *et al*., 2003; Reiser *et al.*, 2000). Two genes carrying out the same function in two different species may not be orthologs. An example for this comes from the *SHOOTMERISTEMLESS* (*STM*) gene of *Arabidopsis* and *KN1* gene from maize (Hake *et al.*, 2004). Both *STM* and *KN1* genes are significantly similar at protein level and both are expressed specifically in meristems and are down-regulated in the P0 in *Arabidopsis* and maize respectively. However *STM* and *KN1* resolve in different clades and hence are not orthologs by descent. Simmilarly *AGAMOUS* and *PLENA* plays homologous roles for determining the carpel and stamen identity in *Arabidopsis* and *Antirrhinum* respectively. However phylogenetic analysis resolves them into different paralogous lineages (Kramer *et al.*, 2004).

Estimating the signature of selection is a useful tool for studying the evolution of different genes. *ARP* genes are transcription factors and consist of the Myb domain and the C-terminal domain. The Myb domain at the N-terminus is repeated in all *ARP*-like genes and divided into Myb domain 1 and Myb domain 2. Both Myb like domains are responsible for nucleic acid binding and the C-terminal domain for protein-protein interactions. I found evidence that the Myb domain1 and C-terminal domain in *Begonia ARP* and other species *ARP* like genes is under purifying/neutral selection whereas Myb domain 2 is under positive selection. Purifying selection eliminates deleterious mutations and favours the maintenance of protein function, whereas new valuable genetic variants sweep a population through positive selection (Nielsen *et al.*, 2005; Nielsen *et al.*, 2010).

ARP-like genes have undergone duplication in various plant species (Luo *et al.*, 2005; Illing *et al.*, 2009). Likewise *BARP* duplication is *Begonia* specific. Selective pressure on *BARP2* genes can not be estimated at the moment as I was able to get *BARP2* sequences from only two species. It would be interesting to amplify *BARP2* genes from other *Begonia* species and perform test of selection to see if it might have acted differently on different copies of the gene.

Class 1 KNOX-like proteins have KNOX1, KNOX2 (collectively known as MEIKNOX domain), ELK domain and Homeodomain. Each domain carries out a specific function. MEIKNOX domain may be involved in protein dimerization and interaction with BELL (BEL-like homeodomains) Homeodomain proteins (Bellaoui *et al*., 2001; Burglin, 1998; Nagasaki *et al*., 2001; Bhatt *et al*., 2004; Chen *et al*., 2004; Cole *et al*., 2006; Muller *et al*., 2001; Smith *et al*., 2002). The ELK domain is responsible for nuclear localization (Cole *et al*., 2006) and the Homeodomain controls DNA binding (Gehring *et al*., 1994; Treisman *et al*., 1989).

A *Begonia* specific duplication for class1 *KNOX* like genes was discovered. When I estimated the selective pressure for *KNOX* CDS (coding sequences) including 67 class 1 *KNOX* like genes, they were under purifying or neutral selection. Class 1 *KNOX* genes are key regulators for the development and maintenance of shoot apical meristem in plants. Any mutations in these genes would be deleterious for crucial gene functions and hence should be selected against. Interestingly the two copies of the *KNOX* like genes appear to be under different selection pressures; the Homeodomain and the ELK domain of *KNB1* under positive selection but the ELK domain of *KNB2* under positive selection. *KNOX* genes are required for functions other than maintaining the SAM in different species; for instance compound leaf formation and floral nectar spur development (Kerstetter *et al*., 1997; Long *et al*., 1996; Vollbrecht *et al*., 2000; Champagne *et al*., 2007, Chen *et al*., 1997; Harvean *et al*., 1996; Hay and Tsiantis, 2006; Kimura *et al*., 2008; Muller *et al*., 2006; Shani *et al*., 2009; Golz *et al*., 2002). The differential selective pressure on *KNB1* and *KNB2* indicate that theses genes may have undergone subfunctionalization or neofunctionalization in *Begonia*.

8.8. *BARP1* protein structure is conserved with *AS1* while *BARP2* has diverged

The transformation of amino acid strings into a three-dimensional shape is crucial for a protein’s biological function/functions. Proteins acquire three dimensional shapes by folding of the alpha-helices and beta-sheets into a compact globule which is driven by non specific hydrophobic interactions. This structure acquires the stability through salt bridges, hydrogen bonds, and the tight packing of side chains and disulfide bonds. However if a protein misfolds, it will loose its normal biological function which can be drastic in living bodies as this process is often irreversible. The predicted 3D structure of *BARP1* overlaps with *Asymmetric leaves 1* (*AS1*) protein of *Arabidopsis* which indicates the conservation of protein structure. While *BARP2* showed disagreements with the *AS1* protein structure in the Myb domain 2, which is under selective pressure in *ARP* genes and may be responsible for the evolution of *ARP* like proteins. *BARP2* may bind to different sequence to *BARP1* or it may not bind at all and may be acting to affect the function of *BARP1*.

8.9. *BARP1* function is conserved with *AS1* while *BARP2* has diverged

In order to see whether the structural conservation of proteins also underlies the conservation of function between *BARP1* and *AS1* and to test whether the deviation of *BARP2* protein structures from *AS1* leads it to deviate from its ancestral function, complementation tests were carried out. In these tests I introduced *BARP1* and *BARP2* under the control of 35 S promoters in *as1* mutant plants of *Arabidopsis.* Both copies of *BARP* genes have complemented the mutant phenotypes as have *RS2* from maize and *SkARP* from *Selaginella* (Harrison *et al.*, 2005; Theodoris *et al*., 2003). However when both *BARP1* and *BARP2* were over expressed in *Arabidopsis* wild type background, *BARP1* generates the expected phenotypes (Luo *et al*., 2005; Harrison *et al.*, 2005; Theodoris *et al*., 2003) while *BARP2* has acted as a dominant negative. This indicates that *BARP* genes may have undergone neofunctionalization in *Begonia*.

8.10. Expression of *ARP* and *KNOX* genes in *Begonia*

KNB1 genes are expressed at shoot apical meristems, leaf primordia, stipule margins and in vascular bundles. *BARP1* genes are expressed in young leaf primordia, stipule margins, lamina margins, and dormant axillary meristems. *BARP1* expression is not detected in active meristems in *B. mazae*. This is consistent with *ARP* genes role in regulating determinacy.

KNOX genes are expressed in SAM and turned off at leaf initiation sites (Jackson *et al*., 1994; Lincoln *et al*., 1994; Long *et al*., 1996; Sentoku *et al*., 1998). In simple leaved species *KNOX* genes stay off throughout leaf development, whereas *KNOX* expression returns as the primordia grows out in compound leaved species to create an indeterminate atmosphere for the generation of leaflets (Harevan *et al*., 1996; Hay & Tsiantis, 2006; Bharathan *et al*., 2002; Shani *et al*., 2009). *KNOX* genes expression also returns in primordia of species with secondary morphogenesis though final leaf form may be simple. For example *KNOX* genes are expressed in the complex leaf primordium of *Lepidium oleraceum* which undergoes secondary morphogenesis to form simple leaves (Bharathan *et al*., 2002).

KNOX gene expression is also reported in leaf primordia which are indeterminate such as *Welwitschia mirabilis* (Pham and Sinha, 2003). In this species leaves are continuously formed from a basal meristem through out its life span which can be up to 2000 years (Talalaj *et al*., 1991). In leaves of *Welwitschia mirabilis KNOX* genes are initially down regulated but subsequently their expression increases in the basal regions of the leaves suggests its role in the basal meristem of the leaf (Pham and Sinha, 2003).

Class1 *KNOX* genes are also expressed in the blades of *Streptocarpus* species which produce new leaves or inflorescences from groove meristems (Harrison *et al*., 2005; Nishi *et al*., 2010). Similarly in class 1 *KNOX* genes are expressed in leaves of Podostemoideae which lacks shoot apical meristems and generates new leaves from the base of older leaves (Katayama *et al*., 2010). These findings indicate that persistence of *KNOX* genes expression in leaf primordia can lead to the prolonged organogenic activity by preventing the precocious exit of tissues from the cell cycle.

ARP genes are expressed at the distal margins of peltate leaves in *S. actinophylla*, *P. aquatica*, *A. pentaphylla*, *O. regnellii* and *S. actinophylla* (Kim *et al*., 2003). *BARP1* is also expressed at the distal margins of all *Begonia* section Gireoudia species. *Begonia* leaves are peltate to some degree as all mature leaves of *Begonia* section Gireoudia species have some lamina out growth at leaf-petiole attachment points. These leaves acquire peltateness as plants mature in *B. peltata* and *B. pruniata*. *KNB1* genes are expressed in leaf primordia and may be facilitating the formation of lamina outgrowth at the sites of petiole-leaf junction. The degree of peltateness may be dependent on the level of expression of *BARP1* genes in different species of *Begonia*. This can be tested by performing RT PCR or QRT-PCR for *BARP1* genes for different *Begonia* species.

8.11. *Begonia* as a system for studying the evolution of leaf form

So far a genetic basis for interspecific variation in leaf form has been described in *Arabidopsis* and *C. hirsuta* and, *Solanum cheesmaniae* and *Solanum galapagense* (Hay and Tsiantis, 2006; Kimura *et al*., 2008). These studies have been carried out on the basis of sequence and expression analysis and transgenic work. However *Begonia* has allowed us to generate interspecific hybrids to establish a direct link between leaf form and locus underlying this morphology alongside allowing us carrying out sequence analysis, expression analysis and transgenic work to study the genetic basis of leaf form. Besides this *Begonia* section Gireoudia has a range of leaf forms such as peltate, simple, deeply lobed and compound leaved species. Using *Begonia* as a model plant for studying the evolution of leaf form we are able to show a direct association of peltate leaf form with *ARP* locus.

Studying the genetic basis of leaf form evolution in *Begonia* has enabled us to find the answers to some of the following questions

1- Is leaf form a simple genetic trait? --- No
2- Are *ARP* and *KNOX* genes genetically associated with leaf form variation between species? --- Yes to *ARP*

3- Is sequence variation in coding sequence and promoters of *ARP* and *KNOX* associated with variation in leaf form? --- No

4- Is expression variation in *ARP* and *KNOX* associated with variation in leaf form? --- Not proven

5- Is *ARP* from *Begonia* functionally equivalent to *ARP* in *Arabidopsis*? --- Yes

Studying *Begonia* as a model plant we were able to see that leaf form is a complex genetic trait. *ARP* genes are genetically linked with peltate leaf form in *Begonia*. Neither the variation in coding sequences nor the variation in expression pattern corresponds to the variation in leaf form in *Begonia*. *BARP1* is functionally equivalent to *AS1* whereas *BARP2* may have acquired new function as in over expressed lines for Modifies *AS1*, the transgenic showed dominant negative behaviour for Modifies *AS1*.

8.12. Conclusion and future plans

Genetic analysis has revealed that leaf form is a complex trait in *Begonia* section Gireoudia.

KNB1 genes are expressed at shoot apical meristem in *Begonia* which is consistent with their role of maintaining SAM in other species. Both *BARP* and *KNB* genes have undergone *Begonia* specific duplication and selection acts differently on duplicated copies which may be prerequisite for sub-functionalization or neo-functionalization of these genes. Direct transfer of duplicated copies of *BARP* genes in *Arabidopsis* has revealed that both copies are functionally equivalent to *Asymmetric leaves 1* gene of *Arabidopsis* as both complement *as1* mutant plants. Modified *AS1* genes (*BARP2* like) over expression in *Arabidopsis* have suggested that modified *AS1* is a negative dominant locus. This negative dominant behaviour of modified *AS1* may be the result of heterodimerization of *AS1* and modified *AS1*. Differences in over expressed lines' phenotypes indicate various level of homodimerization. Mutations at *ARP* (*phantastica* (*phan*)) locus of *Antirrhinum majus* have revealed that subtle changes in the level or pattern of *phan* activity can give rise to a variety of organ morphologies (Waites and Hudson, 1995). *ARP* like genes are involved in regulation of peltate leaf form in a range of species (Kim *et al.*, 2003). I have observed various degree of peltateness in *Begonia* section Gireoudia species ranging from completely peltate (Petiole attachment point almost at the centre of leaf blade such as in *B. nelumbiifolia*) to nearly non peltate leaf phenotypes (species looking like non peltate but having lamina outgrowths such as *B. mazae*). Various levels of interactions between *BARP1* and *BARP2* may have yielded various degree of peltateness in *Begonia* section Gireoudia species.

BARP1 is expressed in dormant meristems and at the tips of leaf lamina. Association mapping studies for *BARP1* gene has suggested that *BARP1* is a major locus regulating peltateness in *Begonia*. These results have been supported by QTL analysis where peltateness is linked to *BARP1* locus.

On the basis of genetic, expression and QTL analysis I propose a model for the development of peltate leaf form in *Begonia* where *ARP* genes are regulating peltate leaf form through the level of *ARP* expression rather than differential expression pattern between peltate and non peltate leaf plants. *ARP* genes may be expressing at higher level in peltate species than in non peltate. This can be tested through RT PCR. Peltateness is tightly linked with *BARP1* genes; however *NO APICAL MERISTEM* (*NAM*) genes are present at a few centimorgan distances from *BARP1*. It is tempting to carry out expression analysis for *NAM* genes to reveal if *BARP1* is controlling peltateness by regulating *NAM* like genes in *Begonia*.

Phylogenetic analysis suggests that the most basal *Begonia* species are African, from which both Asian and America *Begonias* are derived (Forrest *et al*, 2003). Parallel radiations of *Begonia* have occurred in the South East Asia and in the New World. Although the floral form is similar throughout the genus the vegetative form varies widely. It is tempting to carry out similar studies on South East Asia *Begonia* as we have done on *Begonia* section Gireoudia species to reveal if parallel evolution of leaf form in this genus has taken place through same mechanism.

References

Abdurakhmonov, I., Saha, S., Jenkins, J., Buriev, Z., Shermatov, S., Scheffler, B., Pepper, A., Yu, J., Kohel, R. and Abdukarimov, A. (2009) 'Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm', Genetica 136(3): 401-417.

Adam, H., Marguerettaz, M., Qadri, R., Adroher, B., Richaud, F., Collin, M., Thuillet, A. C., Vigouroux, Y., Laufs, P., Tregear, J. W. et al. (2011) 'Divergent expression patterns of miR164 and *CUP-SHAPED COTYLEDON* genes in palms and other monocots: implication for the evolution of meristem function in angiosperms', Mol Biol Evol 28(4): 1439-54.

Agrama, H. A., Eizenga, G. C. and Yan, W. (2007) 'Association mapping of yield and its components in rice cultivars', Molecular Breeding 19(4): 341-356.

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H. and Tasaka, M. (1997) 'Genes Involved in Organ Separation in *Arabidopsis*: An Analysis of the *cup-shaped cotyledon* Mutant', The Plant Cell Online 9(6): 841-857.

Aida, M., Vernoux, T., Furutani, M., Traas, J. and Tasaka, M. (2002) 'Roles of *PIN-FORMED1* and *MONOPTEROS* in pattern formation of the apical region of the *Arabidopsis* embryo', Development 129(17): 3965-3974.

Airoldi, C. A., S. Bergonzi and B. Davies, 2010 Single amino acid change alters the ability to specify male or female organ identity. Proceedings of the National Academy of Sciences 107: 18898-18902.

Akagi, T., Yin, D., Kawamata, N., Bartram, C. R., Hofmann, W. K., Song, J. H., Miller, C. W., den Boer, M. L. and Koeffler, H. P. (2009) 'Functional analysis of a novel DNA polymorphism of a tandem repeated sequence in the asparagine synthetase gene in acute lymphoblastic leukemia cells', Leuk Res 33(7): 991-6.

Akimenko, M. A., Ekker, M., Wegner, J., Lin, W. and Westerfield, M. (1994) 'Combinatorial expression of three zebrafish genes related to *distal-less*: part of a *Homeobox* gene code for the head', J Neurosci 14(6): 3475-86.

Allsopp, A. (1965) 'Strasburgers Textbook of Botany', ISBN-10: 9781144138347

Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D. (1990) 'Basic local alignment search tool', J Mol Biol 215(3): 403 - 410.

Andersson, S. (1991) 'Geographical variation and genetic analysis of leaf shape in *Crepis tectorum*', Plant Systematics and Evolution 178(3): 247-258.

Andersson, S. and Shaw, R. G. (1994) 'Phenotypic plasticity in *Crepis tectorum* (Asteraceae): genetic correlations across light regimens', Heredity 72(2): 113-125.

Andolfatto, P. (2005) 'Adaptive evolution of non-coding DNA in *Drosophila*', Nature 437(7062): 1149-1152.

Aranzana, M. J., Kim, S., Zhao, K., Bakker, E., Horton, M., Jakob, K., Lister, C., Molitor, J., Shindo, C., Tang, C. et al. (2005) 'Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes', Plos Genetics 1(5): e60.

Ashburner, M. (1998) '*Drosophila*. A laboratory handbook'.

Ashby, E. (1948) 'Studies in the morphogenesis of leaves. An essay on leaf shape', New Phytologist 47(2): 153-176.

Averof, M. and Patel, N. H. (1997) 'Crustacean appendage evolution associated with changes in *Hox* gene expression', Nature 388(6643): 682-686.

Barabé D, L. C., Jeune B. (2007) 'Following the Initiation and Development of Individual Leaf Primordia at the Level of the Shoot Apical Meristem: The

Case of Distichous Phyllotaxis in *Begonia*', Annals of Botany 99: 555-560.

Barkoulas, M., Hay, A., Kougioumoutzi, E. and Tsiantis, M. (2008) 'A developmental framework for dissected leaf formation in the *Arabidopsis* relative *Cardamine hirsuta*', Nature Genetics 40(9): 1136-1141.

Barton, M. K. and Poethig, R. S. (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(3): 823-831.

Barton, N. H. and Turell, M. (1989) 'Evolutionary quantitative genetics. *Annu. Rev. Genet.* **23**, 337–370

Beerling, D. J., Osborne, C. P. and Chaloner, W. G. (2001) 'Evolution of leaf-form in land plants linked to atmospheric CO2 decline in the Late Palaeozoic era', Nature 410(6826): 352-354.

Bellaoui, M., Pidkowich, M. S., Samach, A., Kushalappa, K., Kohalmi, S. E., Modrusan, Z., Crosby, W. L. and Haughn, G. W. (2001) 'The *Arabidopsis BELL1* and *KNOX* TALE homeodomain proteins interact through a domain conserved between plants and animals', Plant Cell 13(11): 2455-2470.

Beló, A., Zheng, P., Luck, S., Shen, B., Meyer, D., Li, B., Tingey, S. and Rafalski, A. (2008) 'Whole genome scan detects an allelic variant of *fad2* associated with increased oleic acid levels in maize', Molecular Genetics and Genomics 279(1): 1-10.

Berna, G., Robles, P. and Micol, J. L. (1999) 'A mutational analysis of leaf morphogenesis in *Arabidopsis thaliana*', Genetics 152(2): 729-742.

Bharathan, G., Goliber, T. E., Moore, C., Kessler, S., Pham, T. and Sinha, N. R. (2002) 'Homologies in leaf form inferred from *KNOXI* gene expression during development', Science 296(5574): 1858-1860.

Bharathan, G., Janssen, B. J., Kellogg, E. A. and Sinha, N. (1997) 'Did homeodomain proteins duplicate before the origin of angiosperms, fungi, and metazoa?', Proceedings of the National Academy of Sciences of the United States of America 94(25): 13749-13753.

Bharathan, G., Janssen, B. J., Kellogg, E. A. and Sinha, N. (1999) 'Phylogenetic relationships and evolution of the *KNOTTED* class of plant homeodomain proteins', Molecular Biology and Evolution 16(4): 553-563.

Bharathan, G. and Sinha, N. R. (2001) 'The Regulation of compound leaf development', Plant Physiology 127(4): 1533-1538.

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

Bidichandani, S. I., Ashizawa, T. and Patel, P. I. (1998) 'The GAA Triplet-Repeat Expansion in *Friedreich ataxia* Interferes with transcription and may be associated with an unusual DNA structure', The American Journal of Human Genetics 62(1): 111-121.

Billings, W. D. and Mooney, H. A. (1968) 'Ecology of arctic and alpine plants', Biological Reviews of the Cambridge Philosophical Society 43(4): 481-&.

Black, D. L. (2003) 'Mechanisms of alternative pre-messenger RNA splicing', Annual Review of Biochemistry 72: 291-336.

Blein, T., Pulido, A., Vialette-Guiraud, A., Nikovics, K., Morin, H., Hay, A., Johansen, I. E., Tsiantis, M. and Laufs, P. (2008) 'A Conserved molecular framework for compound leaf development', Science 322(5909): 1835-1839.

Blinks, L. R. (1968) 'The Biology of Aquatic Vascular Plants. C. D. Sculthorpe. St. Martin's, New York, 1967. Science 160(3824): 179.

Bodmer, R., 1993 The gene *tinman* is required for specification of the heart and visceral muscles in *Drosophila*. Development 118: 719-729.

Bommert, P., Satoh-Nagasawa, N., Jackson, D. and Hirano, H. Y. (2005) 'Genetics and evolution of inflorescence and flower development in grasses', Plant and Cell Physiology 46(1): 69-78.

Bonini, N. M., Bui, Q. T., Gray-Board, G. L. and Warrick, J. M. (1997) 'The *Drosophila eyes absent* gene directs ectopic eye formation in a pathway conserved between flies and vertebrates', Development 124(23): 4819-4826.

Bostrack, J. M. and Millington, W. F. (1962) 'On the determination of leaf form in an aquatic heterophyllous species of ranunculus', Bulletin of the Torrey Botanical Club 89(1): 1-20.

Bowers, J. E., Chapman, B. A., Rong, J. and Paterson, A. H. (2003) 'Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events', Nature 422(6930): 433-438.

Bowman, J. L. (2000) 'The *YABBY* gene family and abaxial cell fate', Current Opinion in Plant Biology 3(1): 17-22.

Bowman, J. L., Smyth, D. R. and Meyerowitz, E. M. (1991) 'Genetic Interactions among Floral Homeotic Genes of *Arabidopsis*', Development 112(1): 1-20.

Bowman, J. L. and Eshed, Y. (2000) 'Formation and maintenance of the shoot apical meristem', *Trends in Plant Science* 5(3): 110-115.

Bowman, J. L., and Y. Eshed, 2000 Formation and maintenance of the shoot apical meristem. Trends in Plant Science 5: 110-115.

Boyce, C. K. and Knoll, A. H. (2002) 'Evolution of developmental potential and the multiple independent origins of leaves in Paleozoic vascular plants', Paleobiology 28(1): 70-100.

Bradshaw, H. D., Otto, K. G., Frewen, B. E., McKay, J. K. and Schemske, D. W. (1998) 'Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*)', Genetics 149(1): 367-382.

Bradshaw, H. D. and Schemske, D. W. (2003) 'Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers', Nature 426(6963): 176-178.

Bradshaw, H. D., Wilbert, S. M., Otto, K. G. and Schemske, D. W. (1995) 'Genetic-mapping of floral traits associated with reproductive isolation in Monkeyflowers (Mimulus)', Nature 376(6543): 762-765.

Brakefield, P. M. and French, V. (1995) 'Eyespot development on butterfly wings - the epidermal response to damage', Developmental Biology 168(1): 98-111.

Brand, U., Fletcher, J. C., Hobe, M., Meyerowitz, E. M. and Simon, R. (2000) 'Dependence of stem cell fate in *Arabidopsis* and a feedback loop regulated by *CLV3* activity', Science 289(5479): 617-619.

Bright, K. L. and Rausher, M. D. (2008) 'Natural selection on a leaf-shape polymorphism in the ivyleaf morning glory (*Ipomoea Hederacea*)', Evolution 62(8): 1978-1990.

Bruce P. Dancik, B. V. B. (1974) 'Leaf diversity in yellow birch (*Betula alleghaniensis*)', Canadian Journal of Botany 52(11): 2407-2414.

Burglin, T. R. (1997) 'Analysis of TALE superclass *Homeobox* genes (MEIS, PBC, *KNOX*, Iroquois, TGIF) reveals a novel domain conserved between plants and animals', Nucleic Acids Research 25(21): 4173-4180.

Burglin, T. R. (1998) 'The PBC domain contains a MEINOX domain: Coevolution of *hox* and TALE *Homeobox* genes?', Development Genes and Evolution 208(2): 113-116.

Bylesjo, M., Segura, V., Soolanayakanahally, R., Rae, A., Trygg, J., Gustafsson, P., Jansson, S. and Street, N. (2008) 'LAMINA: a tool for rapid quantification of leaf size and shape parameters', BMC Plant Biology 8(1): 82.

Byrne, M. E., C. A. Kidner and R. A. Martienssen, 2003 Plant stem cells: divergent pathways and common themes in shoots and roots. Current Opinion in Genetics & Development 13: 551-557.

Byrne, M. E., Barley, R., Curtis, M., Arroyo, J. M., Dunham, M., Hudson, A. and Martienssen, R. A. (2000) *'Asymmetric leaves1* mediates leaf patterning and stem cell function in *Arabidopsis*', Nature 408(6815): 967-971.

Byrne, M. E., Simorowski, J. and Martienssen, R. A. (2002) *'ASYMMETRIC LEAVES1* reveals *KNOX* gene redundancy in *Arabidopsis*', Development 129(8): 1957-1965.

Candela, H., Johnston, R., Gerhold, A., Foster, T. and Hake, S. (2008) 'The *milkweed pod1* gene encodes a *KANADI* protein that is required for abaxial/adaxial patterning in maize leaves', Plant Cell 20(8): 2073-2087.

Carles, C. C. and Fletcher, J. C. (2003) 'Shoot apical meristem maintenance: the art of a dynamic balance', Trends in plant science 8(8): 394-401.

Carroll, S. B. (2005) Endless Forms Most Beautiful: The New

Science of Evo Devo: W. W. Norton.

Carroll, S. B., 1995 Homeotic genes and the evolution of arthropods and chordates. Nature 376: 479-485.

Carroll, S. B. (2008) 'Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution', Cell 134(1): 25-36.

Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., Franks, C. D. and Kresovich, S. (2008) 'Community resources and strategies for association mapping in sorghum', Crop Science 48(1): 30-40.

Causier, B., Castillo, R., Zhou, J., Ingram, R., Xue, Y., Schwarz-Sommer, Z. and Davies, B. (2005) 'Evolution in action: Following function in duplicated floral homeotic genes', Current biology 15(16): 1508-1512.

Champagne, C. and Sinha, N. (2004) 'Compound leaves: equal to the sum of their parts?', Development 131(18): 4401-4412.

Champagne, C. E., Goliber, T. E., Wojciechowski, M. F., Mei, R. W., Townsley, B. T., Wang, K., Paz, M. M., Geeta, R. and Sinha, N. R. (2007) 'Compound leaf development and evolution in the legumes', Plant Cell 19(11): 3369-78.

Champagne, C. E. M. and Ashton, N. W. (2001) 'Ancestry of *KNOX* genes revealed by bryophyte (*Physcomitrella patens*) homologs', New Phytologist 150(1): 23-36.

Chen, H., Rosin, F. M., Prat, S. and Hannapel, D. J. (2003) 'Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation', Plant Physiology 132(3): 1391-1404.

Chen, J. J., Janssen, B. J., Williams, A. and Sinha, N. (1997) 'A gene fusion at a *Homeobox* locus: Alterations in leaf shape and implications for morphological evolution', Plant Cell 9(8): 1289-1304.

Chinwalla AT, LL Cook, KD Delehaunty, GA Fewell, LA Fulton, RS Fulton *et al*., (2002) 'Initial sequencing and comparative analysis of the mouse genome', Nature 420(6915): 520-562.

Chitwood, D. H., Guo, M., Nogueira, F. T. S. and Timmermans, M. C. P. (2007) 'Establishing leaf polarity: the role of small RNAs and positional signals in the shoot apex', Development 134(5): 813-823.

Chuck, G., Lincoln, C. and Hake, S. (1996) '*KNAT1* induces lobed leaves with ectopic meristems when overexpressed in *Arabidopsis*', Plant Cell 8(8): 1277-1289.

Civetta, A. and Singh, R. (1995) 'High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species', J Mol Evol 41: 1085 - 1095.

Clark, A. and Hollingsworth, M. (2006) 'Molecular biology for phylogenetic and population genetic studies' Royal Botanic Gardens Edinburgh.

Clark, S. E., Jacobsen, S. E., Levin, J. Z. and Meyerowitz, E. M. (1996) 'The *CLAVATA* and *SHOOT MERISTEMLESS* loci competitively regulate meristem activity in *Arabidopsis*', Development 122(5): 1567-75.

Clark, S. E., Running, M. P. and Meyerowitz, E. M. (1993) '*CLAVATA1*, a regulator of meristem and flower development in *Arabidopsis*', Development 119(2): 397-418.

Clark, S. E., Running, M. P. and Meyerowitz, E. M. (1995) '*CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*', Development 121(7): 2057-2067.

Clark, S. E., Williams, R. W. and Meyerowitz, E. M. (1997) 'The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*', Cell 89(4): 575-585.

Clausen, J., Keck, D. D. and Hiesey, W. M. (1947) 'Heredity of geographically and ecologically isolated races', American Naturalist 81(797): 114-133.

Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B. A. and Johnston, M. (2003) 'Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting', Science 301(5629): 71-76.

Clough, S. J., and A. F. Bent, 1998 Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. The Plant Journal 16: 735-743.

Coen, E. S. and Meyerowitz, E. M. (1991) 'The war of the whorls - genetic interactions controlling flower development', Nature 353(6339): 31-37.

Cohen, S. M., Bronner, G., Kuttner, F., Jurgens, G. and Jackle, H. (1989) '*Distal-less* encodes a homoeodomain protein required for limb development in *Drosophila*', Nature 338(6214): 432-434.

Colasanti, J., Tyers, M. and Sundaresan, V. (1991) 'Isolation and characterization of cdna clones encoding a functional P34cdc2 homolog from *Zea mays'*, Proceedings of the national academy of sciences of the USA 88(8): 3377-3381.

Cole, M., Nolte, C. and Werr, W. (2006) 'Nuclear import of the transcription factor *SHOOT MERISTEMLESS* depends on heterodimerization with BLH proteins expressed in discrete sub-domains of the shoot apical meristem of *Arabidopsis thaliana*', Nucleic Acids Research 34(4): 1281-1292.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, M. Dickson *et al*., 2005 Widespread parallel evolution in *Sticklebacks* by repeated fixation of *Ectodysplasin* Alleles. Science 307: 1928-1933.

Conner, J. K. (2002) 'Genetic mechanisms of floral trait correlations in a natural population', Nature 420(6914): 407-410.

Coyne, J. A. (2005) 'Switching on evolution', Nature 435(7045): 1029-1030.

Coyne, J. A. and Lande, R. (1985) 'The Genetic basis of species differences in plants', The American Naturalist 126(1): 141-145.

Crawford, D. L., Segal, J. A. and Barnett, J. L. (1999) 'Evolutionary analysis of TATA-less proximal promoter function', Molecular Biology and Evolution 16(2): 194-207.

Cronk, Q. C. B. (2001) 'Plant evolution and development in a post-genomic context', Nature Reviews Genetics 2(8): 607-619.

D'hoop, B., Paulo, M., Mank, R., van Eck, H. and van Eeuwijk, F. (2008) 'Association mapping of quality traits in potato (*Solanum tuberosum* L.)', Euphytica 161(1): 47-60.

Daborn, P. J., Yen, J. L., Bogwitz, M. R., Le Goff, G., Feil, E., Jeffers, S., Tijet, N., Perry, T., Heckel, D., Batterham, P. et al. (2002) 'A single P450 allele associated with insecticide resistance in *Drosophila*', Science 297(5590): 2253-2256.

David-Schwartz, R., Runo, S., Townsley, B., Machuka, J. and Sinha, N. (2008) 'Long-distance transport of mRNA via parenchyma cells and phloem across the host-parasite junction in *Cuscuta*', New Phytologist 179(4): 1133-1141.

David-Schwartz, R. and Sinha, N. (2007) 'Evolution and development in plants: Bridging the gap', International Journal of Plant Sciences 168(1): 49-59.

Davidson, E. H. (2006) The regulatory genome: Gene regulatory networks in development and evolution: Academic Press.

Davidson, E. H., Peterson, K. J. and Cameron, R. A. (1995) 'Origin of bilaterian body plans - evolution of developmental regulatory mechanisms', Science 270(5240): 1319-1325.

Davis, S. L. (2001) 'Phenotypic and genetic correlations among floral traits in two species of *Thalictrum*', Journal of Heredity 92(4): 361-366.

Day, O. L. A. S., 2003 Mechanisms in plant development. Blackwell Science Ltd.

de Meaux, J., Goebel, U., Pop, A. and Mitchell-Olds, T. (2005) 'Allele-specific assay reveals functional variation in the chalcone synthase promoter of *Arabidopsis thaliana* that is compatible with neutral evolution', Plant Cell 17(3): 676-690.

Dermitzakis, E. T., Kirkness, E., Schwarz, S., Birney, E., Reymond, A. and Antonarakis, S. E. (2004) 'Comparison of human chromosome 21 conserved nongenic sequences (CNGs) with the mouse and dog genomes shows that their selective constraint is independent of their genic environment', Genome Research 14(5): 852-859.

Dickinson, T. A. (1986) 'Topodeme differentiation in ontario taxa of *Crataegus* (*Rosaceae maloideae*) - leaf morphometric evidence', Canadian Journal of Botany-Revue Canadienne De Botanique 64(11): 2738-2747.

Dickinson, T. A., Parker, W. H. and Strauss, R. E. (1987) 'Another approach to leaf shape comparisons', Taxon 36(1): 1-20.

Dickinson, T. A. and Phipps, J. B. (1984) 'Studies in crataegus (*Rosaceae, Maloideae*) .9. short-shoot leaf heteroblasty in crataegus-crus-galli-sensu-lato', Canadian Journal of Botany-Revue Canadienne De Botanique 62(9): 1775-1780.

Dong, Z. C., Zhao, Z., Liu, C. W., Luo, J. H., Yang, J., Huang, W. H., Hu, X. H., Wang, T. L. and Luo, D. (2005) 'Floral patterning in *Lotus japonicus'*, Plant Physiology 137(4): 1272-1282.

Drea, S. C., Lao, N. T., Wolfe, K. H. and Kavanagh, T. A. (2006) 'Gene duplication, exon gain and neofunctionalization of OEP16-related genes in land plants', The Plant Journal 46(5): 723-735.

Dudareva, N., Cseke, L., Blanc, V. M. and Pichersky, E. (1996) 'Evolution of floral scent in clarkia: novel patterns of *S-linalool synthase* gene expression in the *C-breweri* flower', Plant Cell 8(7): 1137-1148.

Durbin, M. L., McCaig, B. and Clegg, M. T. (2000) 'Molecular evolution of the chalcone synthase multigene family in the morning glory genome', Plant Molecular Biology 42(1): 79-92.

Durst, C. E. (1929) 'Inheritance in lettuce', Science 69(New Series): 2.

Dutcher, S. (2001) 'The tubulin fraternity: alpha to eta', Curr Opin Cell Biol 13: 49 - 54.

Eckardt, N. A. (2002) 'Alternative splicing and the control of flowering time', Plant Cell 14(4): 743-747.

Ehrenreich, I., Hanzawa, Y., Chou, L., Roe, J., Kover, P. and Purugganan, M. (2009) 'Candidate Gene Association Mapping of *Arabidopsis* Flowering Time', Genetics.

El-Gazzar, A. (1980) 'The taxonomic significance of leaf morphology in Crataegus (Rosaceae)', Bot. Jahrb. Syst. 101: 457-469.

Elmore, C. D. (1986) 'Mode of Reproduction and Inheritance of Leaf Shape in Ipomoea-Hederacea', Weed Science 34(3): 391-395.

EL-DIN EL-ASSAL, S., C. ALONSO-BLANCO, A. J. M. PEETERS, V. RAZ and M. KOORNNEEF, 2001 A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. Nat Genet 29: 435-440.

Endrizzi, K., Moussian, B., Haecker, A., Levin, J. Z. and Laux, T. (1996) 'The *SHOOT MERISTEMLESS* gene is required for maintenance of undifferentiated cells in *Arabidopsis* shoot and floral meristems and acts at a different regulatory level than the meristem genes *WUSCHEL* and *ZWILLE*', The Plant Journal 10(6): 967-979.

Engstrom, E. M., Izhaki, A. and Bowman, J. L. (2004) 'Promoter bashing, microRNAs and *KNOX* genes, new insights, regulators and targets of regulation in the establishment of lateral organ polarity in *Arabidopsis*', Plant Physiology 135(2): 685-694.

Eshed, Y., Izhaki, A., Baum, S. F., Floyd, S. K. and Bowman, J. L. (2004) '*Asymmetric leaf* development and blade expansion in *Arabidopsis* are mediated by *KANADI* and *YABBY* activities', Development 131(12): 2997-3006.

Evans, S. M., Yan, W., Murillo, M. P., Ponce, J. and Papalopulu, N. (1995) '*tinman*, a *Drosophila Homeobox* gene required for heart and visceral mesoderm specification, may be represented by a family of genes in vertebrates: XNkx-2.3, a second vertebrate homologue of *tinman*', Development 121(11): 3889-3899.

Evenhuis, N. (1994) 'Catalogue of the fossil flies of the world (Insecta: Diptera)'.

Falconer, D. S., and T. F. C. MACKAY. (1996) Introduction to quantitative

genetics, Harlow, UK.

Fang, G., Bhardwaj, N., Robilotto, R. and Gerstein, M. B. (2010) 'Getting started in gene orthology and functional analysis', Plos Computational Biology 6(3): -.

Flatt, T. (2005) 'The evolutionary genetics of canalization', Quarterly Review of Biology 80(3): 287-316.

Fleming, A. J., McQueen-Mason, S., Mandel, T. and Kuhlemeier, C. (1997) 'Induction of leaf primordia by the cell wall protein Expansin', Science 276(5317): 1415-1418.

Fletcher, J. C., Brand, U., Running, M. P., Simon, R. and Meyerowitz, E. M. (1999) 'Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems', Science 283(5409): 1911-4.

FLINT, J., and R. MOTT, 2001 Finding the molecular basis of quatitative traits: successes and pitfalls. Nat Rev Genet 2: 437-445.

Floyd, S. K. and Bowman, J. L. (2006) 'Distinct developmental mechanisms reflect the independent origins of leaves in vascular plants', Current Biology 16(19): 1911-1917.

Force, A., Lynch, M., F Pickett, F., Amores, A., Yan, Y. and Postlethwait, J. (1999) 'Preservation of duplicate genes by complementary, degenerative and speciation: a new paradigm', Genetics 151: 1531 - 1545.

Forrest LL. (2000) 'A phylogeny of the *Begoniaceae.* Bercht. & J. Presl. Ph.D Thesis, University of Glasgow.

Forrest, L. L., and P. M. Hollingsworth, 2003 A recircumscription of *Begonia* based on nuclear ribosomal sequences. Plant Systematics and Evolution 241: 193-211.

Fourest-Lieuvin, A., Peris, L., Gache, V., Garcia-Saez, I., Juillan-Binard, C., Lantez, V. and Job, D. (2006) 'Microtubule regulation in mitosis: tubulin phosphorylation by the *cyclin-dependent kinase Cdk1*', Mol Biol Cell 17(3): 1041 - 50.

Fox-Walsh, K., Dou, Y., Lam, B., Hung, S., Baldi, P. and Hertel, K. (2005) 'The architecture of pre-mRNAs affects mechanisms of splice-site pairing', Proc Natl Acad Sci USA 102(45): 16176 - 81.

Freeling, M., Bertrandgarcia, R. and Sinha, N. (1992) 'Maize mutants and variants altering developmental time and their heterochronic interactions', Bioessays 14(4): 227-236.

Frugis, G., Giannino, D., Mele, G., Nicolodi, C., Chiappetta, A., Bitonti, M. B., Innocenti, A. M., Dewitte, W., Van Onckelen, H. and Mariotti, D. (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(4): 1370-1380.

Furner I, J. and Pumfrey, J. E. (1992) 'Cell fate in the shoot apical meristem of *Arabidopsis thaliana*', Development 115(3): 755-764.

Furutani, M., Vernoux, T., Traas, J., Kato, T., Tasaka, M. and Aida, M. (2004) 'PIN-FORMED1 and PINOID regulate boundary formation and cotyledon development in Arabidopsis embryogenesis', Development 131(20): 5021-5030.

Fyrberg, E., Fyrberg, C., Biggs, J., Saville, D., Beall, C. and Ketchum, A. (1998) 'Functional nonequivalence of *Drosophila* actin isoforms', Biochem Genet 36: 271 - 287.

Gasch, A., Hinz, U. and Renkawitz-Pohl, R. (1989) 'Intron and upstream sequences regulate expression of the *Drosophila* f83-tubulin gene in the visceral and somatic musculature, respectively', Proc Natl Acad Sci USA 86: 3215 - 3218.

Gaunt, M. and Miles, M. (2002) 'An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks', Mol Biol Evol 19: 748 - 761.

Gehring, W. J., Affolter, M. and Burglin, T. (1994) 'Homeodomain Proteins', Annual Review of Biochemistry 63: 487-526.

Gesteland, R. F., Cech, T. R. & Atkins, J. F. (1999) 'The RNA World, Cold Spring Harbor Laboratory Press, New York.

Givnish, T., 1986 On the economy of plant form and function. Cambridge university press.

Giania, S., Morelloa, L., Bardinib, M. and Breviarioa, D. (2003) 'Tubulin intron sequences: multi-functional tools', Cell Biology International 27: 203 - 205.

Glaser, T., Walton, D. S. and Maas, R. L. (1992) 'Genomic structure, evolutionary conservation and aniridia mutations in the human *PAX6* gene', Nature Genetics 2(3): 232-239.

Gleissberg, S. (2004) 'Comparative analysis of leaf shape development in *Eschscholzia californica* and other Papaveraceae-Eschscholzioideae. American Journal of Botany 91: 306-312.

Gleissberg, S., Groot, E. P., Schmalz, M., Eichert, M., Kolsch, A. and Hutter, S. (2005) 'Developmental events leading to peltate leaf structure in *Tropaeolum majus* (Tropaeolaceae) are associated with expression domain changes of a *YABBY* gene', Development Genes and Evolution 215(6): 313-319.

Goliber, T., Kessler, S., Chen, J. J., Bharathan, G. and Sinha, N. (1999) 'Genetic, molecular, and morphological analysis of compound leaf development', Current Topics in Developmental Biology, Vol 43 43: 259-290.

Golz, J. F., Roccaro, M., Kuzoff, R. and Hudson, A. (2004) '*GRAMINIFOLIA* promotes growth and polarity of *Antirrhinum* leaves', Development 131(15): 3661-3670.

Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A. and Carroll, S. B. (2005) 'Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*', Nature 433(7025): 481-487.

González-Martínez, S. C., Ersoz, E., Brown, G. R., Wheeler, N. C. and Neale, D. B. (2006) 'DNA sequence variation and selection of tag single nucleotide polymorphisms at candidate genes for drought stress response in *Pinus taeda* L', Genetics 172(3): 1915-1926.

Gonzalez-Martinez, S. C., Wheeler, N. C., Ersoz, E., Nelson, C. D. and Neale, D. B. (2007) 'Association genetics in *Pinus taeda* L. I. Wood property traits', Genetics 175(1): 399-409.

Gordon, S. P., Chickarmane, V. S., Ohno, C. and Meyerowitz, E. M. (2009) 'Multiple feedback loops through cytokinin signaling control stem cell number within the *Arabidopsis* shoot meristem', Proc Natl Acad Sci U S A 106(38): 16529-34.

Gottlieb, L. D. (1986) 'The Genetic-Basis of Plant Form', Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 313(1159): 197-208.

Gourlay, C. W., Hofer, J. M. I. and Ellis, T. H. N. (2000) 'Pea Compound Leaf Architecture Is Regulated by Interactions among the Genes *UNIFOLIATA, COCHLEATA, AFILA,* and *TENDRIL-LESS'*, The Plant Cell Online 12(8): 1279-1294.

Grabowski, P. J. and Black, D. L. (2001) 'Alternative RNA splicing in the nervous system', Progress in Neurobiology 65(3): 289-308.

Grant-Downton, R. T. and Dickinson, H. G. (2006) 'Epigenetics and its Implications for Plant Biology 2. The ‘Epigenetic Epiphany’: Epigenetics, Evolution and Beyond', Annals of Botany 97(1): 11-27.

Greyson, R. I., Walden, D. B. and Smith, W. J. (1982) 'Leaf and Stem Heteroblasty in *Zea*', Botanical Gazette 143(1): 73-78.

Groot, E. P., Sinha, N. and Gleissberg, S. (2005) 'Expression patterns of *STM*-like *KNOX* and *Histone H4* genes in shoot development of the dissected-leaved basal eudicot plants *Chelidonium majus* and *Eschscholzia californica* (Papaveraceae)', Plant Molecular Biology 58(3): 317-331.

Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W. and Li, W.-H. (2003) 'Role of duplicate genes in genetic robustness against null mutations', Nature 421(6918): 63-66.

Guo, M., Thomas, J., Collins, G. and Timmermans, M. C. (2008) 'Direct repression of *KNOX* loci by the *ASYMMETRIC LEAVES1* complex of *Arabidopsis*', Plant Cell 20(1): 48-58.

Gurevitch, J. (1988) 'Variation in leaf dissection and leaf energy budgets among populations of *Achillea* from an altitudinal gradient', American Journal of Botany 75(9): 1298-1306.

Haecker, A., Gross-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrmann, M. and Laux, T. (2004) 'Expression dynamics of *WOX* genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*', Development 131(3): 657-68.

Hahn, M. W., Rockman, M. V., Soranzo, N., Goldstein, D. B. and Wray, G. A. (2004) 'Population genetic and phylogenetic evidence for positive selection on regulatory mutations at the Factor VII locus in humans', Genetics 167(2): 867-877.

Hake, S., Sinha, N., Veit, B., Vollbrecht, E. and Walko, R. (1991) 'Mutations of *Knotted* Alter Cell-Interactions in the Developing Maize Leaf', Plant Molecular Biology 2 212: 555-562.

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

Halder, G., Callaerts, P. and Gehring, W. J. (1995) 'Induction of Ectopic Eyes by Targeted Expression of the *Eyeless* Gene in *Drosophila*', Science 267(5205): 1788-1792.

Hammond, D. (1941) 'The expression of genes for leaf shape in *Gossypium hirsutum* L and *Gossypium arboreum* L. American Journal of Botany 28(2): 124-138.

Hareven, D., Gutfinger, T., Parnis, A., Eshed, Y. and Lifschitz, E. (1996) 'The making of a compound leaf: Genetic manipulation of leaf architecture in tomato', Cell 84(5): 735-744.

Harris, W., Beever, R. E. and Heenan, P. B. (1998) 'Phenotypic variation of leaves and stems of wild stands of *Cordyline australis* (Lomandraceae)', New Zealand Journal of Botany 36(4): 593-604.

Harrison, C. J., Corley, S. B., Moylan, E. C., Alexander, D. L., Scotland, R. W. and Langdale, J. A. (2005) 'Independent recruitment of a conserved developmental mechanism during leaf evolution', Nature 434(7032): 509-514.

Harrison, C. J., Rezvani, M. and Langdale, J. A. (2007) 'Growth from two transient apical initials in the meristem of *Selaginella kraussiana'*, Development 134(5): 881-9.

Hartmann, H. D. I. H. E. K. (1982) Leaf surfaces in Mesembryanthemaceae, Academic Press, London.

Harvey, R. P. (1996) *'NK-2 Homeobox* genes and heart development', Developmental Biology 178(2): 203-216.

Hay, A., Barkoulas, M. and Tsiantis, M. (2006) *'ASYMMETRIC LEAVES1* and auxin activities converge to repress *BREVIPEDICELLUS* expression and promote leaf development in *Arabidopsis*', Development 133(20): 3955-3961.

Hay, A. and Tsiantis, M. (2006) 'The genetic basis for differences in leaf form between *Arabidopsis thaliana* and its wild relative *Cardamine hirsuta'*, Nature Genetics 38(8): 942-947.

Hay, A. and Tsiantis, M. (2009) 'A *KNOX* family TALE', Current Opinion in Plant Biology 12(5): 593-598.

Hay, A. and Tsiantis, M. (2010) '*KNOX* genes: versatile regulators of plant development and diversity', Development 137(19): 3153-3165.

Heisler, M. G., Ohno, C., Das, P., Sieber, P., Reddy, G. V., Long, J. A. and Meyerowitz, E. M. (2005) 'Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the *Arabidopsis* inflorescence meristem', Current Biology 15(21): 1899-1911.

Henikoff, S. and Henikoff, J. (1992) 'Amino acid substitution matrices from protein blocks', Proc Natl Acad Sci USA 89: 10915 - 10919.

Hernandez, M. L., Passas, H. J. and Smith, L. G. (1999) 'Clonal analysis of epidermal patterning during maize leaf development', Developmental Biology 216(2): 646-658.

Hertel, K. (2008) 'Combinatorial control of exon recognition', Journal of Bio. Chem. 283(3): 1211 - 1215.

Hibara, K.-i., Karim, M. R., Takada, S., Taoka, K.-i., Furutani, M., Aida, M. and Tasaka, M. (2006) '*Arabidopsis CUP-SHAPED COTYLEDON3* regulates postembryonic shoot meristem and organ boundary formation', The Plant Cell 18(11): 2946-2957.

Higgins, D., Thompson, J., Gibson, T., Thompson, J., Higgins, D. and Gibson, T. (1998) ' Sequencher sequence analysis software, Gene Codes Corporation, Ann Arbor, MI USA.

Julie, D. T, Desmond G.H. and Toby J.G., 1994' CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice', Nuleic Acid Research, 22: 4673 - 4680.

Hileman, D. R., Huluka, G., Kenjige, P. K., Sinha, N., Bhattacharya, N. C., Biswas, P. K., Lewin, K. F., Nagy, J. and Hendrey, G. R. (1994) 'Canopy photosynthesis and transpiration of field grown cotton exposed to free air CO2 enrichment (face) and differential irrigation', Agricultural and Forest Meteorology 70(1-4): 189-207.

Hill, R. S. (1980) 'A Numerical taxonomic approach to the study of angiosperm leaves', Botanical Gazette 141(2): 213-229.

Hilu, K. W. (1983) 'The Role of single gene mutations in the evolution of flowering plants', Evolutionary Biology 16: 97-128.

Hodges, S. A., Whittall, J. B., Fulton, M. and Yang, J. Y. (2002) 'Genetics of floral traits influencing reproductive isolation between *Aquilegia formosa* and *Aquilegia pubescens'*, American Naturalist 159: S51-S60.

Hofer, J., Gourlay, C., Michael, A. and Ellis, T. H. N. (2001) 'Expression of a class 1 *knotted1*-like *Homeobox* gene is down-regulated in pea compound leaf primordia', Plant Molecular Biology 45(4): 387-398.

Hofer, J., Turner, L., Hellens, R., Ambrose, M., Matthews, P., Michael, A. and Ellis, N. (1997) '*UNIFOLIATA* regulates leaf and flower morphogenesis in pea', Current biology : CB 7(8): 581-587.

Hogan, B. L. M., Horsburgh, G., Cohen, J., Hetherington, C. M., Fisher, G. and Lyon, M. F. (1986) *'Small eyes* (*Sey*): a homozygous lethal mutation on chromosome 2 which affects the differentiation of both lens and nasal placodes in the mouse', Journal of Embryology and Experimental Morphology 97(1): 95-110.

Holland, N. D., Panganiban, G., Henyey, E. L. and Holland, L. Z. (1996) 'Sequence and developmental expression of *AmphiDll*, an amphioxus *Distal-less* gene transcribed in the ectoderm, epidermis and nervous system: insights into evolution of craniate forebrain and neural crest', Development 122(9): 2911-20.

Horiguchi, G., Mollá-Morales, A., Pérez-Pérez, J. M., Kojima, K., Robles, P., Ponce, M. R., Micol, J. L. and Tsukaya, H. (2011) 'Differential contributions of ribosomal protein genes to *Arabidopsis thaliana* leaf development', The Plant Journal 65(5): 724-736.

Horiguchi, G., Nakayama, H., Ishikawa, N., Kubo, M., Demura, T., Fukuda, H. and Tsukaya, H. (2011) '*ANGUSTIFOLIA3* Plays Roles in

Adaxial/Abaxial Patterning and Growth in Leaf Morphogenesis', Plant and Cell Physiology 52(1): 112-124.

Hoyle, H. and Raff, E. (1990) 'St. Mivart. On the genesis of species (London 1871)

Hoyle, H., Turner, F. and Raff, E. (2008) 'Axoneme-dependent tubulin modifications in singlet microtubules of the *Drosophila* sperm tail', Cell Motil Cytoskeleton 65(4): 295 - 313.

Hughes, A. (2005) 'Gene duplication and the origin of novel proteins', Proc Natl Acad Sci USA 102(25): 8933 - 8.

Hutchens, J., Hoyle, H., Turner, F. and Raff, E. (1997) 'Structurally similar *Drosophila* alpha-tubulins are functionally distinct in vivo', Mol Biol Cell 8: 481 - 500.

Ihlenfeldt, H. D., 1994 Diversification in an arid world: The mesembryanthemaceae. Annual Review of Ecology and Systematics 25: 521-546.

Illing, N., Klak, C., Johnson, C., Brito, D., Negrao, N., Baine, F., van Kets, V., Ramchurn, K., Seoighe, C. and Roden, L. (2009) 'Duplication of the *Asymmetric Leaves1*; Rough Sheath 2 / *PHANTASTICA* gene precedes the explosive radiation of the Ruschioideae', Development Genes and Evolution 219(6): 331-338.

Imrei, B. C. and Knowles, P. F. (1970) 'Inheritance studies in interspecific hybrids between *Carthamus flavescens* and *C. tinctorius'*, Crop Science 10(4): 349-352

Irish, V. F. and Sussex, I. M. (1992) 'A fate map of the *Arabidopsis* embryonic shoot apical meristem', Development 115(3): 745-753.

Jackson, D., Greene, B., Kerstetter, R., Sinha, N., Smith, L., Veit, B. and Hake, S. (1992) 'The Role of *Homeobox* genes in maize development', Molecular Biology of the Cell 3: A99-A99.

Jackson, D., Veit, B. and 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(2): 405-413.

Janssen, B. J., Lund, L. and Sinha, N. (1998) 'Overexpression of a *Homeobox* gene, *LeT6,* reveals indeterminate features in the tomato compound leaf', Plant Physiology 117(3): 771-786.

Janssen, B. J., Williams, A., Chen, J. J., Mathern, J., Hake, S. and Sinha, N. (1998) 'Isolation and characterization of two *knotted*-like *Homeobox* genes from tomato', Plant Molecular Biology 36(3): 417-425.

Ji, J. B., Strable, J., Shimizu, R., Koenig, D., Sinha, N. and Scanlon, M. J. (2010) '*WOX4* promotes procambial development', Plant Physiology 152(3): 1346-1356.

Jordan, T., Hanson, I., Zaletayev, D., Hodgson, S., Prosser, J., Seawright, A., Hastie, N. and van Heyningen, V. (1992) 'The human *PAX6* gene is mutated in two patients with *aniridia*', Nature Genetics 1(5): 328-332.

Jouannic, S., Collin, M., Vidal, B., Verdeil, J. L. and 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(3): 551-568.

Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A. and Timmermans, M. C. P. (2004) 'microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity', Nature 428(6978): 84-88.

Juarez, M. T., Twigg, R. W. and Timmermans, M. C. P. (2004) 'Specification of adaxial cell fate during maize leaf development', Development 131(18): 4533-4544.

Jun, T.-H., Van, K., Kim, M., Lee, S.-H. and Walker, D. (2008) 'Association analysis using SSR markers to find QTL for seed protein content in soybean', Euphytica 162(2): 179-191.

Kalisz, S. and Purugganan, M. D. (2004) 'Epialleles via DNA methylation: consequences for plant evolution', Trends in ecology & evolution (Personal edition) 19(6): 309-314.

Kaplan, D. R. (1970) 'Comparative foliar histogenesis in *Acorus calamus* and its bearing on phyllode theory of monocotyledonous leaves', American Journal of Botany 57(3): 331-&.

Karimi, M., Inze, D. and Depicker, A. (2002) 'GATEWAY vectors for *Agrobacterium*-mediated plant transformation', Trends Plant Sci 7(5): 193-5.

Katayama, N., Koi, S. and Kato, M. (2010) 'Expression of *SHOOT MERISTEMLESS*, WUSCHEL, and *ASYMMETRIC LEAVES1* homologs in the shoots of podostemaceae: implications for the evolution of novel shoot Organogenesis', The Plant Cell Online.

Kawasaki, H., Sugaya, K., Quan, G., Nohata, J. and Mita, K. (2003) 'Analysis of alpha- and beta-tubulin genes of Bombyx mori using an EST database', Insect Biochem Mol Biol 33: 131 - 7.

Kelley, L. A. and Sternberg, M. J. E. (2009) 'Protein structure prediction on the Web: a case study using the Phyre server', Nat. Protocols 4(3): 363-371.

Kenrick, P. and Crane, P. R. (1997) 'The origin and early evolution of plants on land', Nature 389(6646): 33-39.

Kerstetter, R., Vollbrecht, E., Lowe, B., Veit, B., Yamaguchi, J. and Hake, S. (1994) 'Sequence analysis and expression patterns divide the maize *knotted*1-like *Homeobox* genes into two classes', The Plant Cell 6(12): 1877-1887.

Kerstetter, R. A., Bollman, K., Taylor, R. A., Bomblies, K. and Poethig, R. S. (2001) '*KANADI* regulates organ polarity in *Arabidopsis*', Nature 411(6838): 706-709.

Kerstetter, R. A., LaudenciaChingcuanco, D., Smith, L. G. and Hake, S. (1997) 'Loss-of-function mutations in the maize *Homeobox* gene, *knotted*1, are defective in shoot meristem maintenance', Development 124(16): 3045-3054.

Kessler, S., Kim, M., Pham, T., Weber, N. and Sinha, N. (2001) 'Mutations altering leaf morphology in tomato', International Journal of Plant Sciences 162(3): 475-492.

Kessler, S., Seiki, S. and Sinha, N. (2002) '*Xcl1* causes delayed oblique periclinal cell divisions in developing maize leaves, leading to cellular differentiation by lineage instead of position', Development 129(8): 1859-1869.

Kessler, S. and Sinha, N. (2004) 'Shaping up: the genetic control of leaf shape', Current Opinion in Plant Biology 7(1): 65-72.

Kessler, S., Townsley, B. and Sinha, N. (2006) 'L1 division and differentiation patterns influence shoot apical meristem maintenance', Plant Physiology 141(4): 1349-1362.

Khaitovich, P., Weiss, G., Lachmann, M., Hellmann, I., Enard, W., Muetzel, B., Wirkner, U., Ansorge, W. and Paabo, S. (2004) 'A neutral model of transcriptome evolution', Plos Biology 2(5): 682-689.

Kidner, C. A. and Martienssen, R. A. (2004) 'Spatially restricted microRNA directs leaf polarity through *ARGONAUTE1*', Nature 428(6978): 81-84.

Kidner, C. A. and Timmermans, M. C. P. (2007) 'Mixing and matching pathways in leaf polarity', Current Opinion in Plant Biology 10(1): 13-20.

Kidner, C. A. and Umbreen, S. (2010) 'Why is leaf shape so variable?', International Journal of Plant Developmental Biology. 4: 64-75.

Kim, M., Canio, W., Kessler, S. and Sinha, N. (2001) 'Developmental changes due to long-distance movement of a *Homeobox* fusion transcript in tomato', Science 293(5528): 287-289.

Kim, M., McCormick, S., Timmermans, M. and Sinha, N. (2003) 'The expression domain of *PHANTASTICA* determines leaflet placement in compound leaves, Nature 425(6953): 102-102.

Kim, M., Pham, T., Hamidi, A., McCormick, S., Kuzoff, R. K. and Sinha, N. (2003) 'Reduced leaf complexity in tomato wiry mutants suggests a role for *PHAN* and *KNOX* genes in generating compound leaves', Development 130(18): 4405-4415.

Kim, M. and Sinha, N. (2003) 'Regulating shapes and sizes', Developmental Cell 4(4): 441-442.

Kim, Y. S., Kim, S. G., Lee, M., Lee, I., Park, H. Y., Seo, P. J., Jung, J. H., Kwon, E. J., Suh, S. W., Paek, K. H. (2008) *'HDZIP* III activity is modulated by competitive inhibitors via a feedback loop in *Arabidopsis* shoot apical meristem development', The Plant Cell 20(4): 920-933.

Kimura, S., Koenig, D., Kang, J., Yoong, F. Y. and Sinha, N. (2008) 'Natural variation in leaf morphology results from mutation of a novel *KNOX* gene', Current Biology 18(9): 672-677.

Klak, C., Khunou, A., Reeves, G. and Hedderson, T. (2003) 'A phylogenetic hypothesis for the Aizoaceae (Caryophyllales) based on four plastid DNA regions', American Journal of Botany 90(10): 1433-1445.

Kliebenstein, D. J., West, M. A. L., van Leeuwen, H., Kim, K., Doerge, R. W., Michelmore, R. W. and St Clair, D. A. (2006) 'Genomic survey of gene expression diversity in *Arabidopsis thaliana*', Genetics 172(2): 1179-1189.

Koch, M. A., Weisshaar, B., Kroymann, J., Haubold, B. and Mitchell-Olds, T. (2001) 'Comparative genomics and regulatory evolution: Conservation and function of the *CHS* and *Apetala3* promoters', Molecular Biology and Evolution 18(10): 1882-1891.

Koenig, D., Bayer, E., Kang, J., Kuhlemeier, C. and Sinha, N. (2009) 'Auxin patterns *Solanum lycopersicum* leaf morphogenesis', Development 136(17): 2997-3006.

Koenig, D. and Sinha, N. (2010) Evolution Of leaf shape: A pattern emerges. Current Topics in Developmental Biology, vol. Volume 91: Academic Press.

Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T. and Yano, M. (2002) 'Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions', *Plant Cell Physiol* 43(10): 1096-105.

Koltai, H. and Bird, D. M. (2000) ’Epistatic repression of *PHANTASTICA* and class I *KNOTTED* genes is uncoupled in tomato’, Plant Journal 22: 455-459.

Komuro, I. and Izumo, S. (1993) '*Csx* - a Murine *Homeobox* containing gene specifically expressed in the developing heart', Proceedings of the National Academy of Sciences of the United States of America 90(17): 8145-8149.

Konieczny, A. and Ausubel, F. M. (1993) 'A procedure for mapping *Arabidopsis* mutations using codominant ecotype specific PCR based markers', Plant Journal 4(2): 403-410.

Koonin, E. (2006) 'The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate?', Biology Direct 1: 22.

Koonin, E. V. (2005) 'Orthologs, paralogs, and evolutionary genomics', Annual Review of Genetics 39: 309-338.

Kopperund, C. and Einset, J. W. (1995) 'DNA isolation from *Begonia* leaves' Plant Molecular Biology Reporter 13: 129-130.

Kraakman, A. T. W., Martinez, F., Mussiraliev, B., van Eeuwijk, F. A. and Niks, R. E. (2006) 'Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars', Molecular Breeding 17(1): 41-58.

Kuhlemeier, C. and Sinha, N. (2007) 'Growth and development - The diversity of plant development', Current Opinion in Plant Biology 10(1): 1-3.

Kumar, S., Tamura, K. and Nei, M. (2004) 'MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment', Briefings in Bioinformatics 5: 150 - 163.

Landon, E. M. and Blum, K. E. (1978) 'Leaf Variation in Five Species of Acaulescent Blue Violets (*Viola*). 1', Castanea 43(2): 95-107.

Lenhard, M. and Laux, T. (2003) 'Stem cell homeostasis in the *Arabidopsis* shoot meristem is regulated by intercellular movement of *CLAVATA3* and its sequestration by *CLAVATA1*', Development 130(14): 3163-73.

Lerman, D. N., Michalak, P., Helin, A. B., Bettencourt, B. R. and Feder, M. E. (2003) 'Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements', Molecular Biology and Evolution 20(1): 135-144.

Lewontin, R. and Hubby, J. (1966) 'A molecular approach to the study of genic heterozygosity in natural populations. II. of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura'*, Genetics 54: 595 - 609.

Li, W. H. (1997) Molecular Evolution, Sinauer Associates. Sunderland, MA.

Li, X. L. and Noll, M. (1994) 'Compatibility between enhancers and promoters determines the transcriptional specificity of *Gooseberry* and *Gooseberry Neuro* in the *Drosophila* Embryo', Embo Journal 13(2): 400-406.

Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K. and Hake, S. (1994) 'A *Knotted*1-like *Homeobox* gene in *Arabidopsis* is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants', Plant Cell 6(12): 1859-1876.

Lints, T. J., Parsons, L. M., Hartley, L., Lyons, I. and Harvey, R. P. (1993) *'Nkx-2.5* - a novel murine *Homeobox* gene expressed in early heart progenitor cells and their myogenic descendants (Vol 119, Pg 419, 1993)', Development 119(3): 969-969.

Long, J. A., Moan, E. I., Medford, J. I. and Barton, M. K. (1996) 'A member of the *KNOTTED* class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*', Nature 379(6560): 66-69.

Long, M. and Langley, C. (1993) 'Natural selection and the origin of *Jingwei,* a chimeric processed functional gene in *Drosophila*', Science 260(5104): 91-95.

Loudet, O., Chaillou, S., Krapp, A. and Daniel, F. (2003) 'Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in *Arabidopsis thaliana*', Genetics 163(2): 711-722.

Luo, J. H., Yan, J., Weng, L., Yang, J., Zhao, Z., Chen, J. H., Hu, X. H. and Luo, D. (2005) 'Different expression patterns of duplicated *PHANTASTICA*-like genes in *Lotus japonicus* suggest their divergent functions during compound leaf development', Cell Res 15(8): 665-677.

Lutz, B., Lu, H. C., Eichele, G., Miller, D. and Kaufman, T. C. (1996) 'Rescue of *Drosophila labial* null mutant by the chicken ortholog *Hoxb-1* demonstrates that the function of *Hox* genes is phylogenetically conserved', Genes & Development 10(2): 176-184.

Lynch, M. and Richardson, A. O. (2002) 'The evolution of spliceosomal introns', Current Opinion in Genetics & Development 12(6): 701-710.

Lyndon, R. F. (1998) The Shoot Apical Meristem, Cambridge University Press, Cambridge.

Maizel, A., Busch, M. A., Tanahashi, T., Perkovic, J., Kato, M., Hasebe, M. and Weigel, D. (2005) 'The floral regulator *LEAFY* evolves by substitutions in the DNA binding domain', Science 308(5719): 260-263.

Malosetti, M., van der Linden, C. G., Vosman, B. and van Eeuwijk, F. A. (2007) 'A mixed-model approach to association mapping using pedigree information with an illustration of resistance to *Phytophthora infestans* in Potato', Genetics 175(2): 879-889.

Marais, G., Nouvellet, P., Keightley, P. and Charlesworth, B. (2005) 'Intron size and exon evolution in *Drosophila*', Genetics 170(1): 481 - 5.

Martin, C. and PazAres, J. (1997) 'MYB transcription factors in plants', Trends in Genetics 13(2): 67-73.

Matthews, K., Miller, D. and Kaufman, T. (1989) 'Developmental distribution of RNA and protein products of the *Drosophila alpha-tubulin* gene family', Developmental Biology 132: 45 - 61.

McConnell, J. R. and Barton, M. K. (1998) 'Leaf polarity and meristem formation in *Arabidopsis*', Development 125(15): 2935-2942.

McConnell, J. R., Emery, J., Eshed, Y., Bao, N., Bowman, J. and Barton, M. K. (2001) 'Role of *PHABULOSA* and *PHAVOLUTA* in determining radial patterning in shoots', Nature 411(6838): 709-713.

McHale, N. A. and Koning, R. E. (2004) '*PHANTASTICA* Regulates development of the adaxial mesophyll in *Nicotiana* leaves', The Plant Cell Online 16(5): 1251-1262.

McKay, J. K., Richards, J. H. and Mitchell-Olds, T. (2003) 'Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits', Molecular Ecology 12(5): 1137-1151.

McKean, P., Vaughan, S. and Gull, K. (2001) 'The extended tubulin superfamily', J Cell Sci 114: 2723 - 2733.

Mclellan, T. (1990) 'Development of differences in leaf shape in *Begonia dregei* (Begoniaceae)', American Journal of Botany 77(3): 323-337.

McLellan, T. (2000) 'Geographic variation and plasticity of leaf shape and size in *Begonia dregei* and *B. homonyma* (Begoniaceae)', Botanical Journal of the Linnean Society 132(1): 79-95.

Meagher, R. and Hightower, R. (1986) 'The Molecular evolution of actin', Genetics 114: 315 - 332.

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

Medford, J. I., Behringer, F. J., Callos, J. D. and Feldmann, K. A. (1992) 'Normal and abnormal development in the *Arabidopsis* vegetative shoot apex', Plant Cell 4(6): 631-643.

Melville, R. (1960) 'A metrical study of leaf shape in hybrids: I. The leaf shape of some F1 Hybrids and their parents', Kew Bulletin 14(1): 88-102.

Mencarelli, C., Caroti, D., Bre, M., Levilliers, N. and Dallai, R. (2005) 'Tubulin glycylation and glutamylation deficiencies in unconventional insect axonemes', Cell Motil Cytoskeleton 61(4): 226 - 36.

Merrill, E. K. (1979) 'Comparison of ontogeny of 3 types of leaf architecture in *Sorbus* L (Rosaceae)', Botanical Gazette 140(3): 328-337.

Michaels, S. D. and Amasino, R. M. (1998) 'A robust method for detecting single-nucleotide changes as polymorphic markers by PCR', Plant Journal 14(3): 381-385.

Michalak, P., Minkov, I., Helin, A., Lerman, D. N., Bettencourt, B. R., Feder, M. E., Korol, A. B. and Nevo, E. (2001) 'Genetic evidence for adaptation driven incipient speciation of *Drosophila melanogaster* along a microclimatic contrast in "Evolution Canyon," Israel', Proceedings of the National Academy of Sciences of the USA 98(23): 13195-13200.

Michalis, A., (2010) 'Evolution of new appendage types by gradual changes in Hox gene expression the case of crustacean maxillipeds', Palaeodiversity 3: 141-145

Micol, J. L. (2009) 'Leaf development: time to turn over a new leaf?', Current Opinion in Plant Biology 12(1): 9-16.

Mikkelsen TS, Hillier LD, Eichler EE, et al. (11 co-authors) (2005) 'Initial sequence of the chimpanzee genome and comparison with the human genome', Nature 437(7055): 69-87.

Miller, R. B. (1981) 'Hawkmoths and the Geographic Patterns of Floral Variation in Aquilegia-Caerulea', Evolution 35(4): 763-774.

Modrek, B. and Lee, C. (2002) 'A genomic view of alternative splicing', Nature Genetics 30(1): 13-19.

Müller, J., Wang, Y., Franzen, R., Santi, L., Salamini, F. and Rohde, W. (2001) 'In vitro interactions between barley TALE homeodomain proteins suggest a role for protein–protein associations in the regulation of *KNOX* gene function', The Plant Journal 27(1): 13-23.

Mukherjee, K., Brocchieri, L. and Burglin, T. R. (2009) 'A comprehensive classification and evolutionary analysis of plant *Homeobox* genes', Mol Biol Evol 26(12): 2775-94.

Muller, K. J., He, X., Fischer, R. and Prufer, D. (2006) 'Constitutive *KNOX*1 gene expression in dandelion (*Taraxacum officinale*, Web.) changes leaf morphology from simple to compound', Planta 224(5): 1023-7.

Muller, K. J., Romano, N., Gerstner, O., Garciamaroto, F., Pozzi, C., Salamini, F. and Rohde, W. (1995) 'The Barley Hooded Mutation Caused by a Duplication in a *Homeobox* Gene Intron', Nature 374(6524): 727-730.

Nagasaki, H., Sakamoto, T., Sato, Y. and Matsuoka, M. (2001) 'Functional analysis of the conserved domains of a rice *KNOX* homeodomain protein, OSH15', Plant Cell 13(9): 2085-2098.

Nagato, K. H. a. Y. (2006) '*OsNAM* and *OsCUC3* are expressed specically in organ boundaries', Rice Genetics Newsletter 23: 96–97.

Natzel, J. and McCarthy, B. (1984) 'Regulation of *Drosophila* alpha- and beta-tubulin genes during development', Developmental Biology 104: 187 - 98.

Nikovics, K., T. Blein, A. Peaucelle, T. Ishida, H. Morin *et al*., 2006 The balance between the Mir164A and *CUC2* genes controls leaf margin serration in *Arabidopsis*. The Plant Cell Online 18: 2929-2945.

Nielsen, M., Caserta, J., Kidd, S. and Phillips, C. (2006) 'Functional constraint underlies 60 million year stasis of Dipteran testis-specific beta-tubulin', Evolution & Development 8: 23 - 29.

Nielsen, M., Gadagkar, S. and Gutzwiller, L. (2010) 'Tubulin evolution in insects: gene duplication and subfunctionalization provide specialized isoforms in a functionally constrained gene family', BMC Evolutionary Biology 10(1): 113.

Nielsen, M. and Raff, E. (2002) 'The best of all worlds or the best possible world? Developmental constraint in the evolution of beta tubulin and the sperm tail axoneme', Evolution & Development 4: 303 - 315.

Nielsen, M., Turner, F., Hutchens, J. and Raff, E. (2001) 'Axoneme-specific beta-tubulin specialization a conserved C-terminal motif specifies the central pair', Current Biology 11: 529 - 533.

Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., Fledel-Alon, A., Tanenbaum, D. M., Civello, D., White, T. J. et al. (2005) 'A scan for positively selected genes in the genomes of humans and chimpanzees', PLoS Biol 3(6): e170.

Nogales, E., Whittaker, M., Milligan, R. and Downing, K. (1999) 'High-resolution model of the microtubule', Cell 96(1): 79 - 88.

Nogales, E., Wolf, S. and Downing, K. (1998) 'Structure of the alpha beta tubulin dimer by electron crystallography', Nature 391: 199 - 203.

Nogueira, F. T. S., S. Madi, D. H. Chitwood, M. T. Juarez and M. C. P. Timmermans, 2007 Two small regulatory RNAs establish opposing fates of a developmental axis. Genes & Development 21: 750-755.

Nurmberg, P. L., *KNOX*, K. A., Yun, B.-W., Morris, P. C., Shafiei, R., Hudson, A. and Loake, G. J. (2007) 'The developmental selector *AS1* is an evolutionarily conserved regulator of the plant immune response', Proceedings of the National Academy of Sciences 104(47): 18795-18800.

Ohno, S. (1970) Evolution by gene duplication: Springer Verlag, New York.

Oleksiak, M. F., Churchill, G. A. and Crawford, D. L. (2002) 'Variation in gene expression within and among natural populations', Nature Genetics 32(2): 261-266.

Olsen, K. M., Halldorsdottir, S. S., Stinchcombe, J. R., Weinig, C., Schmitt, J. and Purugganan, M. D. (2004) 'Linkage disequilibrium mapping of *Arabidopsis CRY2* flowering time alleles', Genetics 167(3): 1361-1369.

Ori, N., Eshed, Y., Chuck, G., Bowman, J. L. and Hake, S. (2000) 'Mechanisms that control *KNOX* gene expression in the *Arabidopsis* shoot', Development 127(24): 5523-5532.

Ottoline Leyser, H. M. and Furner, I. J. (1992) 'Characterisation of three shoot apical meristem mutants of *Arabidopsis thaliana*', Development 116(2): 397-403.

Oyama, K. E. N. (1996) 'Quantitative variation within and among populations of *Arabis serrata* Fr. & Sav. (Brassicaceae)', Botanical Journal of the Linnean Society 120(3): 243-256.

Palaisa, K. A., Morgante, M., Williams, M. and Rafalski, A. (2003) 'Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci', Plant Cell 15(8): 1795-1806.

Palazzo1, A., Ackerman, B. and Gundersen, G. (2003) 'Tubulin acetylation and cell motility', Nature 421: 230.

Parker, W. H., Bradfield, G. E., Maze, J. and Lin, S. C. (1979) 'Analysis of Variation in Leaf and Twig Characters of *Abies lasiocarpa* and *Abies amabilis* from North-Coastal British-Columbia', Canadian Journal of Botany-Revue Canadienne De Botanique 57(12): 1354-1366.

Parker, W. H., J. Maze and G. E. Bradfield, 1981 Implications of morphological and anatomical variation in *Abies balsamea* and *A. lasiocarpa* (Pinaceae) from Western Canada. American Journal of Botany 68: 843-854.

Parker, W. H. and Maze, J. (1984) 'Intraspecific Variation in *Abies lasiocarpa* from British-Columbia and Washington', American Journal of Botany 71(8): 1051-1059.

Paterson, A. H., Bowers, J. E. and Chapman, B. A. (2004) 'Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics', Proceedings of the National Academy of Sciences of the USA 101(26): 9903-9908.

Paxman, G. J. (1956) 'Differentiation and stability in the development of *Nicotiana rustica*', Annals of Botany 20(2): 331-347.

Phipps, J. B., 1983 Biogeographic, taxonomic, and cladistic relationships between east asiatic and north American *Crataegus*. Annals of the Missouri Botanical Garden 70: 667-700.

Pérez-Pérez, J. M., Serrano-Cartagena, J. and Micol, J. L. (2002) 'Genetic analysis of natural variations in the architecture of *Arabidopsis thaliana* vegetative leaves', Genetics 162(2): 893-915.

Peris, L., Thery, M., Faure, J., Saoudi, Y., Lafanechere, L., Chilton, J., Gordon-Weeks, P., Galjart, N., Bornens, M., Wordeman, L. et al. (2006) 'Tubulin tyrosination is a major factor affecting the recruitment of CAP-Gly proteins at microtubule plus ends', J Cell Biol 174(6): 839 - 49.

Peter, C. and Prins, B. (2008) '*Vitis* shoots show reversible change in leaf shape along the shoot axis', Proceedings of the 2nd annual national viticulture research conference: 17-18.

Pham, T. and Sinha, N. (2003) 'Role of *KNOX* genes in shoot development of *Welwitschia mirabilis'*, International Journal of Plant Sciences 164(3): 333-343.

Phelps-Durr, T. L., Thomas, J., Vahab, P. and Timmermans, M. C. P. (2005) 'Maize *rough sheath2* and its *Arabidopsis* orthologue *ASYMMETRIC LEAVES1* interact with HIRA, a predicted histone chaperone, to maintain *KNOX* gene silencing and determinacy during organogenesis', The Plant Cell 17(2): 2886-2898.

Phillips, R. B. (1983) 'Shape characters in numerical taxonomy and problems with ratios', Taxon 32(4): 535-544.

Piatigorsky, J. (2007) Gene sharing and evolution: the diversity of protein functions, Cambridge, Harvard Univ. Press Massachusetts.

Poethig, R. S. (1988) 'Heterochronic mutations affecting shoot development in maize', Genetics 119(4): 959-973.

Ponomarenko, A. (1995) 'The geological history of beetles', biology, phylogeny, and classification of *Coleoptera*: papers celebrating the 80th birthday of Roy A. Crowson: 155 - 171.

Prigge, M. J., Otsuga, D., Alonso, J. M., Ecker, J. R., Drews, G. N. and Clark, S. E. (2005) 'Class III *Homeodomain-Leucine Zipper* gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development', The Plant Cell 17(1): 61-76.

Prince, V. E. and Pickett, F. B. (2002) 'Splitting pairs: The diverging fates of duplicated genes', Nature Reviews Genetics 3(11): 827-837.

Provencher, L. M., Miao, L., Sinha, N. and Lucas, W. J. (2001) 'Sucrose export defective1 encodes a novel protein implicated in chloroplast-to-nucleus signaling', Plant Cell 13(5): 1127-1141.

Quentin Cronk, R. B., Julie A Hawkins and (2002)' Developmental denetics and plant evolution, UK. Taylor and Francis, London.

Quiring, R., Walldorf, U., Kloter, U. and Gehring, W. J. (1994) 'Homology of the *Eyeless* gene of *Drosophila* to the *Small Eye* gene in mice and *Aniridia* in humans', Science 265(5173): 785-789.

Raff, E., Fackenthal, J., Hutchens, J., Hoyle, H. and Turner, F. (1997) 'Microtubule architecture specified by a beta-tubulin isoform', Science 275: 70 - 73.

Raff, E., Hoyle, H., Popodi, E. and Turner, F. (2008) 'Axoneme beta-tubulin sequence determines attachment of outer dynein arms', Current Biology 18: 911 - 914.

Raff, E., Hutchens, J., Hoyle, H., Nielsen, M. and Turner, F. (2000) 'Conserved axoneme symmetry altered by a component beta-tubulin', Current Biology 10: 1391 - 4.

Rausher, K. L. B. a. M. D. (2008) 'Natural selection on a leaf shape polymorphism in the ivyleaf morning glory (*Ipomoea hederacea*)', Evolution 62(8): 1978–1990.

Redei, G. P. (1965) 'Non-mendelian mega gametogenesis in *Arabidopsis*', Genetics 51(6): 857.

Reinhardt, D., Wittwer, F., Mandel, T. and Kuhlemeier, C. (1998) 'Localized up regulation of a new *Expansin* gene predicts the site of leaf formation in the Tomato Meristem', The Plant Cell 10(9): 1427-1438.

Reiser, L., Sanchez-Baracaldo, P. and Hake, S. (2000) 'Knots in the family tree: evolutionary relationships and functions of *KNOX Homeobox* genes', Plant Molecular Biology 42(1): 151-166.

Rockman, M. V., Hahn, M. W., Soranzo, N., Goldstein, D. B. and Wray, G. A. (2003) 'Positive selection on a human-specific transcription factor binding site regulating *IL4* expression', Current Biology 13(23): 2118-2123.

Rockman, M. V., Hahn, M. W., Soranzo, N., Loisel, D. A., Goldstein, D. B. and Wray, G. A. (2004) 'Positive selection on *MMP3* regulation has shaped heart disease risk', Current Biology 14(17): 1531-1539.

Rodriguez-Trelles, F., Tarrio, R. and Ayala, F. (2006) 'Origins and evolution of spliceosomal introns', Annual Review of Genetics 40: 47 - 76.

Ronquist, F., John, P., Huelsenbeck, J. and Mark, P. (2005) 'Mr. Bayes v. 3.1.

Roy, S. (2004) 'The origin of recent introns: transposons?' Genome Biology 5: 251.

Sachs, T. (1972) 'Induction fiber differentiation in peas', Annals of Botany 36(144): 189.

Saïdou, A.-A., Mariac, C., Luong, V., Pham, J.-L., Bezançon, G. and Vigouroux, Y. (2009) 'Association studies identify natural variation at *PHYC* linked to flowering time and morphological variation in *Pearl millet'*, Genetics 182(3): 899-910.

Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahori, S. and Matsuoka, M. (2001) '*KNOX* homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem', Genes Dev 15(5): 581-90.

Sano, R., Juarez, C. M., Hass, B., Sakakibara, K., Ito, M., Banks, J. A. and Hasebe, M. (2005) '*KNOX Homeobox* genes potentially have similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems', Evolution & Development 7(1): 69-78.

Santel, A., Kaufmann, J., Hyland, R. and Renkawitz-Pohl, R. (2000) 'The initiator element of the *Drosophila beta2 tubulin* gene core promoter contributes to gene expression in vivo but is not required for male germ-cell specific expression', Nucleic Acids Research 28: 1439 - 46.

Sato, T., Yamanishi, Y., Horimoto, K., Kanehisa, M. and Hiroyuki, T. (2006) 'Partial correlation coefficient between distance matrices as a new indicator of protein-protein interactions', Bioinformatics 22(20): 2488 - 2492.

Savard, J., Tautz, D., Richards, S., Weinstock, G., Gibbs, R., Werren, J., Tettelin, H. and Lercher, M. (2006) 'Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects', Genome Res 16(11): 1334 - 8.

Schneeberger, R., Tsiantis, M., Freeling, M. and Langdale, J. A. (1998) 'The *rough sheath2* gene negatively regulates *Homeobox* gene expression during maize leaf development', Development 125(15): 2857-2865.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F. X., Jurgens, G. and Laux, T. (2000) 'The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes', Cell 100(6): 635-644.

Schulte, P. M. (2001) 'Environmental adaptations as windows on molecular evolution', comparative biochemistry and physiology. Biochemistry & Molecular Biology 128(3): 597-611.

Schulte, P. M., Glemet, H. C., Fiebig, A. A. and Powers, D. A. (2000) 'Adaptive variation in lactate dehydrogenase-B gene expression: Role of a stress-responsive regulatory element', Proceedings of the National Academy of Sciences of the USA 97(12): 6597-6602.

Schulte, P. M., GomezChiarri, M. and Powers, D. A. (1997) 'Structural and functional differences in the promoter and 5' flanking region of *Ldh-B* within and between populations of the teleost *Fundulus heteroclitus'*, Genetics 145(3): 759-769.

Sculthorpe, C. D., 1967. The Biology of aquatic vascular plants. Edward Arnold, London.

Semiarti, E., Ueno, Y., Tsukaya, H., Iwakawa, H., Machida, C. and Machida, Y. (2001) 'The *asymmetric leaves2* gene of *Arabidopsis thaliana* regulates formation of a symmetric lamina, establishment of venation and repression of meristem-related *Homeobox* genes in leaves', Development 128(10): 1771-1783.

Sentoku, N., Tamaoki, M., Nishimura, A. and Matsuoka, M. (1998) 'The *Homeobox* gene *NTH23* of tobacco is expressed in the basal region of leaf primordia', Biochimica 1399(2-3): 203-208.

Serikawa, K. A. and Mandoli, D. F. (1999) '*AaKNOX1*, a *kn1*-like *Homeobox* gene in *Acetabularia acetabulum*, undergoes developmentally regulated subcellular localization', Plant Molecular Biology 41(6): 785-793.

Serikawa, K. A., Martinez-Laborda, A. and Zambryski, P. (1996) 'Three *knotted*1-like *Homeobox* genes in *Arabidopsis*', Plant Molecular Biology 32(4): 673-683.

Shani, E., Burko, Y., Ben-Yaakov, L., Berger, Y., Amsellem, Z., Goldshmidt, A., Sharon, E. and Ori, N. (2009) 'Stage-specific regulation of *Solanum lycopersicum* leaf maturation by class 1 *KNOTTED*1-LIKE *HOMEOBOX* proteins', Plant Cell 21(10): 3078-92.

Sheldon, C. C., Rouse, D. T., Finnegan, E. J., Peacock, W. J. and Dennis, E. S. (2000) 'The molecular basis of vernalization: The central role of *FLOWERING LOCUS C* (*FLC*)', Proceedings of the National Academy of Sciences of the United States of America 97(7): 3753-3758.

Shull, G. H. (1909) *Bursa Bursa-pastoris* and *Bursa Heegeri*: biotypes and hybrids, Washington: Carnegie institution of Washington.

Siegfried, K. R., Eshed, Y., Baum, S. F., Otsuga, D., Drews, G. N. and Bowman, J. L. (1999) 'Members of the *YABBY* gene family specify abaxial cell fate in *Arabidopsis*', Development 126(18): 4117-4128.

Simillion, C., Vandepoele, K., Van Montagu, M. C. E., Zabeau, M. and Van de Peer, Y. (2002) 'The hidden duplication past of *Arabidopsis thaliana*', Proceedings of the National Academy of Sciences of the USA 99(21): 13627-13632.

Singer, S. D. and Ashton, N. W. (2007) 'Revelation of ancestral roles of *KNOX* genes by a functional analysis of *Physcomitrella* homologues', Plant Cell Reports 26(12): 2039-2054.

Sinha, N. (1997) 'Simple and compound leaves: reduction or multiplication?', Trends in Plant Science 2(10): 396-402.

Sinha, N. (1998) 'Organ and cell fusions in the *adherent1* mutant in maize', International Journal of Plant Sciences 159(5): 702-715.

Sinha, N. (1999) 'Leaf development in angiosperms', Annual Review of Plant Physiology and Plant Molecular Biology 50: 419-446.

Sinha, N. (2000) 'The response of epidermal cells to contact', Trends in plant science 5(6): 233-234.

Sinha, N. (2005) 'The role of *KNOX* genes in shoot morphogenesis and compound leaf development', Mechanisms of Development 122: S13-S13.

Sinha, N. and Hake, S. (1990) 'Mutant characters of *Knotted* maize leaves are determined in the innermost tissue layers', Developmental Biology 141(1): 203-210.

Sinha, N. and Hake, S. (1992) 'Perturbations in leaf development caused by the dominant *Knotted* mutation in maize', Progress in Plant Growth Regulation 13: 360-370.

Sinha, N. and Hake, S. (1993) 'Cloning, mapping, and characterization of *Homeobox* containing sequences in tomato', Journal of Cellular Biochemistry: 36-36.

Sinha, N. and Hake, S. (1994) 'The *Knotted* leaf blade is a mosaic of blade, sheath, and auricle identities', Developmental Genetics 15(5): 401-414.

Sinha, N., Hake, S. and Freeling, M. (1993) 'Genetic and molecular analysis of leaf development', Current Topics in Developmental Biology, Vol 28 28: 47-80.

Sinha, N. R. and Kellogg, E. A. (1996) 'Parallelism and diversity in multiple origins of C4 photosynthesis in the grass family', American Journal of Botany 83(11): 1458-1470.

Sinha, N. R., Williams, R. E. and Hake, S. (1993) 'Overexpression of the maize *homeobox* gene, *KNOTTED*-1, causes a switch from determinate to indeterminate cell fates', Genes Dev 7(5): 787-95.

Skøt, L., Humphreys, M. O., Armstead, I., Heywood, S., Skøt, K. P., Sanderson, R., Thomas, I. D., Chorlton, K. H. and Hamilton, N. R. S. (2005) 'An association mapping approach to identify flowering time genes in natural populations of *Lolium perenne* (L.)', Molecular Breeding 15(3): 233-245.

Sliwinski, M. K., Bosch, J. A., Yoon, H. S., von Balthazar, M. and Baum, D. A. (2007) 'The role of two *LEAFY* paralogs from *Idahoa scapigera* (Brassicaceae) in the evolution of a derived plant architecture', Plant Journal 51(2): 211-219.

Smith, H. M. S., Boschke, I. and Hake, S. (2002) 'Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity', Proceedings of the National Academy of Sciences of the USA 99(14): 9579-9584.

Smith, L. B., D. H. Wasshausen, J.Golding, and C. E. Karegeannes 1986) Begoniaceae Part I: Illustrated key, Part II: Annotated species list: Smithsonian Contr. Bot.

Soltis, D. E., Buggs, R. J., Barbazuk, W. B., Schnable, P. S. and Soltis, P. S. (2009) 'On the origins of species: does evolution repeat itself in polyploid populations of independent origin?', Cold Spring Harb Symp Quant Biol 74: 215-23.

Sosinski, B., Gannavarapu, M., Hager, L. D., Beck, L. E., King, G. J., Ryder, C. D., Rajapakse, S., Baird, W. V., Ballard, R. E. and Abbott, A. G. (2000) 'Characterization of microsatellite markers in peach (*Prunus persica* (L.) Batsch)', TAG Theoretical and Applied Genetics 101(3): 421-428.

Souer, E., vanHouwelingen, A., Kloos, D., Mol, J. and Koes, R. (1996) *'The no apical meristem* gene of *petunia* is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries', Cell 85(2): 159-170.

Stephens, S. G. (1945) 'Canalization of gene action in the *Gossypium* leaf shape system and its bearing on certain evolutionary mechanisms', Journal of Genetics 46(2-3): 345-357.

Stern, D. L. and Orgogozo, V. (2008) 'The loci of evolution: How predictable is genetic evolution?', Evolution 62(9): 2155-2177.

Stich, B., Melchinger, A., Heckenberger, M., Möhring, J., Schechert, A. and Piepho, H.-P. (2008) 'Association mapping in multiple segregating populations of sugar beet *Beta vulgaris*; L.)', Theoretical and Applied Genetics 117(7): 1167-1179.

Stich, B., Piepho, H.-P., Schulz, B. and Melchinger, A. (2008) 'Multi-trait association mapping in sugar beet (*Beta vulgaris*; L.)', Theoretical and Applied Genetics 117(6): 947-954.

Stracke, S., Haseneyer, G., Veyrieras, J.-B., Geiger, H., Sauer, S., Graner, A. and Piepho, H.-P. (2009) 'Association mapping reveals gene action and interactions in the determination of flowering time in barley', Theoretical and Applied Genetics 118(2): 259-273.

Sucena, E., Delon, I., Jones, I., Payre, F. and Stern, D. L. (2003) 'Regulatory evolution of *shavenbaby*/*ovo* underlies multiple cases of morphological parallelism', Nature 424(6951): 935-938.

Sundas-Larsson, A., Svenson, M., Liao, H. and Engstrom, P. (1998) 'A *Homeobox* gene with potential developmental control function in the meristem of the conifer *Picea abies'*, Proceedings of the National Academy of Sciences of the USA 95(25): 15118-15122.

Sverdlov, A., Babenko, V., Rogozin, I. and Koonin, E. (2004) 'Preferential loss and gain of introns in 3' portions of genes suggests a reverse-transcription mechanism of intron insertion', Gene 338(1): 85 - 91.

Sverdlov, A., Rogozin, I., Babenko, V. and Koonin, E. (2003) 'Evidence of splice signal migration from exon to intron during intron evolution', Current Biology 13(24): 2170 - 4.

Swarup, K., Alonso-Blanco, C., Lynn, J. R., Michaels, S. D., Amasino, R. M., Koornneef, M. and Millar, A. J. (1999) 'Natural allelic variation identifies new genes in the *Arabidopsis* circadian system', Plant Journal 20(1): 67-77.

Tajima, F. (1989) 'Statistical method for testing the neutral mutation hypothesis by DNA polymorphism', Genetics 123: 585 - 595.

Takada, S., Hibara, K., Ishida, T. and Tasaka, M. (2001) 'The *CUP-SHAPED COTYLEDON1* gene of *Arabidopsis* regulates shoot apical meristem formation', Development 128(7): 1127-1135.

Talalaj, S. T. D. T. J. (1991) The strangest plants in the world, Hill of Content Melbourne.

Talbert, P. B., Adler, H. T., Parks, D. W. and Comai, L. (1995) 'The *REVOLUTA* gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of *Arabidopsis thaliana*', Development 121(9): 2723-2735.

Tattersall, A. D., Turner, L., *KNOX*, M. R., Ambrose, M. J., Ellis, T. H. N. and Hofer, J. M. I. (2005) 'The mutant *crispa* reveals multiple roles for *PHANTASTICA* in pea compound leaf development', The Plant Cell 17: 1046-1060.

Tautz, D. and Schlötterer, C. (1994) 'Simple sequences', Current Opinion in Genetics & Development 4(6): 832-837.

Taylor, J. S. and Raes, J. (2004) 'Duplication and divergence: the evolution of new genes and old ideas', Annual Review of Genetics 38: 615-43.

Theodoris, G., Inada, N. and Freeling, M. (2003) 'Conservation and molecular dissection of *ROUGH SHEATH2* and *ASYMMETRIC LEAVES1* function in leaf development', Proc Natl Acad Sci USA 100(11): 6837-42.

Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D. and Buckler, E. S. (2001) '*Dwarf8* polymorphisms associate with variation in flowering time', Nature Genetics 28(3): 286-289.

Thornton, J. W. and DeSalle, R. (2000) 'Gene family evolution and homology: Genomics meets phylogenetics', Annual Review of Genomics and Human Genetics 1(1): 41-73.

Timmermans, C. A. K. a. M. C. P. (2010) 'Signaling sides: adaxial–abaxial patterning in leaves', Current Topics in Developmental Biology 91: 141-168.

Timmermans,M. C., N. P. Schultes, J. P. Jankovsky and T. Nelson, 1998 *Leafbladeless1* is required for dorsoventrality of lateral organs in maize. Development 125: 2813-2823.

Timmermans, M. C. P., Hudson, A., Becraft, P. W. and Nelson, T. (1999) *'ROUGH SHEATH2*: A Myb protein that represses *KNOX Homeobox* genes in maize lateral organ primordia', Science 284(5411): 151-153.

Tomescu, A. M. F. (2009) 'Megaphylls, microphylls and the evolution of leaf development', Trends in plant science 14(1): 5-12.

Torgerson, D. and Singh, R. (2004) 'Rapid evolution through gene duplication and subfunctionalization of the testes specific alpha4 proteasome subunits in *Drosophila*', Genetics 168: 1421 - 1432.

Triesman, J., and C. Desplan, 1989 The products of the *Drosophila* gap genes hunchback and Kruppel bind to the hunchback promoters. Nature 341: 335-337.

Trelsman, J., Gönczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989) 'A single amino acid can determine the DNA binding specificity of homeodomain proteins', Cell 59(3): 553-562.

Trotochaud, A. E., Jeong, S. and Clark, S. E. (2000) '*CLAVATA3*, a multimeric ligand for the *CLAVATA1* receptor-kinase', Science 289(5479): 613-617.

Tsiantis, M., Schneeberger, R., Golz, J. F., Freeling, M. and Langdale, J. A. (1999) 'The maize *rough sheath2* gene and leaf development programs in monocot and dicot plants', Science 284(5411): 154-156.

Tuszynski, J., Carpenter, E., Huzil, J., Malinski, W., Luchko, T. and Luduena, R. (2006) 'The evolution of the structure of tubulin and its potential consequences for the role and function of microtubules in cells and embryos', Int J Dev Biol 50: 341 - 358.

Uchida, N., Kimura, S., Koenig, D. and Sinha, N. (2010) 'Coordination of leaf development via regulation of *KNOX*1 genes', Journal of Plant Research 123(1): 7-14.

Uchida, N., Townsley, B., Chung, K. H. and Sinha, N. (2007) 'Regulation of *SHOOT MERISTEMLESS* genes via an upstream-conserved noncoding sequence coordinates leaf development', Proc Natl Acad Sci U S A 104(40): 15953-8.

Ungerer, M. C., Halldorsdottir, S. S., Modliszewski, J. L., Mackay, T. F. C. and Purugganan, M. D. (2002) 'Quantitative trait loci for inflorescence development in *Arabidopsis thaliana*', Genetics 160(3): 1133-1151.

Veit, B., Greene, B., Lowe, B., Mathern, J., Sinha, N., Vollbrecht, E., Walko, R. and Hake, S. (1991) 'Genetic approaches to inflorescence and leaf development in maize', Development: 105-111.

Vent, J., Wyatt, T., Smith, D., Banerjee, A., Luduena, R., Sisson, J. and Hallworth, R. (2005) 'Direct involvement of the isotype-specific Cterminus of beta tubulin in ciliary beating', J Cell Sci 118(Pt 19): 4333 - 41.

Verhey, K. and Gaertig, J. (2007) 'The Tubulin Code', Cell Cycle 6: 2152 - 2160.

Vernoux, T., Besnard, F. and Traas, J. (2010) 'Auxin at the shoot apical meristem', Cold Spring Harb Perspect Biol 2(4): 24.

Vogel, S. (2009) 'Leaves in the lowest and highest winds: temperature, force and shape', New Phytologist 183(1): 13-26.

Vollbrecht, E., Veit, B., Sinha, N. and Hake, S. (1991) 'The developmental gene *Knotted*-1 Is a member of a maize *Homeobox* gene family', Nature 350(6315): 241-243.

Vollbrecht, E., L. Reiser and S. Hake, 2000 Shoot meristem size is dependent on inbred background and presence of the maize homeobox gene, *knotted1*. Development 127: 3161-3172.

Waites, R. and Hudson, A. (1995) '*PHANTASTICA* - a gene required for dosoventrarlity of leaves in *Antirrhinum majus'*, Development 121(7): 2143-2154.

Waites, R., Selvadurai, H. R. N., Oliver, I. R. and Hudson, A. (1998) 'The *PHANTASTICA* gene encodes a MYB transcription factor involved in growth and dosoventrarlity of lateral organs in *Antirrhinum*', Cell 93(5): 779-789.

Walsh, J. B. (1995) 'How often do duplicated genes evolve new functions', Genetics 139(1): 421-428.

Walther, C. and Gruss, P. (1991) *'Pax-6*, a murine paired box gene, is expressed in the developing CNS', Development 113(4): 1435-49.

Wang, H., Chen, J., Wen, J., Tadege, M., Li, G., Liu, Y., Mysore, K. S., Ratet, P. and Chen, R. (2008) 'Control of compound leaf development by

FLORICAULA/*LEAFY* ortholog *SINGLE LEAFLET1* in *Medicago truncatula'*, Plant Physiology 146(4): 1759-72.

Wang, J., McClean, P., Lee, R., Goos, R. and Helms, T. (2008) 'Association mapping of iron deficiency chlorosis loci in soybean (*Glycine max* L. Merr.) advanced breeding lines', TAG Theoretical and Applied Genetics 116(6): 777-787.

Wei, X. M., Jackson, P. A., McIntyre, C. L., Aitken, K. S. and Croft, B. (2006) 'Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure', Theoretical and Applied Genetics 114(1): 155-164.

Weil, J., Min, T. P., Yang, C., Wang, S. R., Sutherland, C., Sinha, N. and Kang, C. H. (1999) 'Stabilization of the i-motif by intra molecular adenine-adenine-thymine base triple in the structure of d(ACCCT)', Acta Crystallographica Section Biological Crystallography 55: 422-429.

Weir, I., Lu, J. P., Cook, H., Causier, B., Schwarz-Sommer, Z. and Davies, B. (2004) 'CUPULIFORMIS establishes lateral organ boundaries in *Antirrhinum*', Development 131(4): 915-922.

Wenkel, S., J. Emery, B.-H. Hou, M. M. S. Evans and M. K. Barton, 2007 A feedback regulatory module formed by *LITTLE ZIPPER* and *HD-ZIPIII* genes. The Plant Cell Online 19: 3379-3390.

West, J. G. and Noble, I. R. (1984) 'Analyses of digitized leaf images of the *Dodonaea viscosa* complex in Australia', Taxon 33(4): 595-613.

Whaley, W. G. and Whaley, C. Y. (1942) 'A developmental analysis of inherited leap patterns in *Tropaeolum*', American Journal of Botany 29(3): 195-200.

Wheat, C., Watt, W., Pollock, D. and Schulte, P. (2006) 'From DNA to fitness differences: sequences and structures of adaptive variants of Colias phosphoglucose isomerase (PGI)', Mol Biol Evol 3: 499 - 512.

Wheeler, D. A., Kyriacou, C. P., Greenacre, M. L., Yu, Q., Rutila, J. E., Rosbash, M. and Hall, J. C. (1991) 'Molecular transfer of a species specific behaviour from *Drosophila simulans* to *Drosophila melanogaster'*, Science 251(4997): 1082-1085.

Whitaker, T. W. (1944) 'The Inheritance of certain characters in a cross of two American species of *Lactuca*', Bulletin of the Torrey Botanical Club 71(4): 347-355.

White, R. and Turner, M. (1995) 'Anatomy and development of the fern sporophyte'.

Whiting, M. (2002) 'Phylogeny of holometabolus insect orders: molelcular evidence', Zoologica Scripta 31: 3 - 17.

Whitt, S. R., Wilson, L. M., Tenaillon, M. I., Gaut, B. S. and Buckler, E. S. (2002) 'Genetic diversity and selection in the maize starch pathway', Proceedings of the National Academy of Sciences of the USA 99(20): 12959-12962.

Wilson, L. M., Whitt, S. R., Ibanez, A. M., Rocheford, T. R., Goodman, M. M. and Buckler, E. S. (2004) 'Dissection of maize kernel composition and starch production by candidate gene association', Plant Cell 16(10): 2719-2733.

Wittkopp, P. J., Vaccaro, K. and Carroll, S. B. (2002) 'Evolution of *yellow* gene regulation and pigmentation in *Drosophila*', Current Biology 12(18): 1547-1556.

Wray, G. A., Hahn, M. W., Abouheif, E., Balhoff, J. P., Pizer, M., Rockman, M. V. and Romano, L. A. (2003) 'The Evolution of Transcriptional Regulation in Eukaryotes', Molecular Biology and Evolution 20(9): 1377-1419.

Wu, G., Lin, W. C., Huang, T. B., Poethig, R. S., Springer, P. S. and Kerstetter, R. A. (2008) '*KANADI*1 regulates adaxial-abaxial polarity in *Arabidopsis* by directly repressing the transcription of *ASYMMETRIC LEAVES2'*, Proceedings of the National Academy of Sciences of the USA 105(42): 16392-16397.

Wyatt, R. and Antonovics, J. (1981) 'Butterflyweed re-revisited spatial and temporal patterns of leaf shape variation in *Asclepias tuberosa'*, Evolution 35(3): 529-542.

Xia, L., Hai, B., Gao, Y., Burnette, D., Thazhath, R., Duan, J., Bre, M.-H., Levilliers, N., Gorovsky, M. and Gaertig, J. (2000) 'Polyglycylation of tubulin is essential and affects cell motility and division in *Tetrahymena thermophila'*, J Cell Biol 149: 1097 - 1106.

Xu, R. and Li, Q. (2008) 'Protocol: Streamline cloning of genes into binary vectors in Agrobacterium via the Gateway(R) TOPO vector system', Plant Methods 4(1): 4.

Yalamanchili, K., Sukhija, R., Aronow, W. S., Sinha, N., Fleisher, A. G. and Lehrman, S. G. (2004) 'Prevalence of increased cardiac troponin I levels in patients with and without acute pulmonary embolism and relation of increased cardiac troponin I levels with in-hospital mortality in patients with acute pulmonary embolism', American Journal of Cardiology 93(2): 263-264.

Yamaguchi, T. and Tsukaya, H. (2010) 'Evolutionary and developmental studies of unifacial leaves in monocots: *Juncus* as a model system', J Plant Res 123(1): 35-41.

Yamaguchi, T., Yano, S. and Tsukaya, H. (2010) 'Genetic framework for flattened leaf blade formation in unifacial leaves of *Juncus prismatocarpus'*, Plant Cell 22(7): 2141-55.

Yang, Z. (2007) 'PAML 4: a program package for phylogenetic analysis by maximum likelihood', Molecular Biology and Evolution 24: 1586 - 1591.

Yano, M., Harushima, Y., Nagamura, Y., Kurata, N., Minobe, Y. and Sasaki, T. (1997) 'Identification of quantitative trait loci controlling heading date in rice using a high-density linkage map', *TAG Theoretical and Applied Genetics* 95(7): 1025-1032.

Yao, J., Wang, L., Liu, L., Zhao, C. and Zheng, Y. (2009) 'Association mapping of agronomic traits on chromosome 2A of wheat', Genetica 137(1): 67-75.

Yoon, H. S. and Baum, D. A. (2004) 'Transgenic study of parallelism in plant morphological evolution', Proc Natl Acad Sci U S A 101(17): 6524-9.

Zambito, A. and Wolff, J. (1997) 'Palmitoylation of tubulin', J Biochem Biophys Res Commun 239(3): 650 - 4.

Zhang, G. H., Xu, Q., Zhu, X. D., Qian, Q. and Xue, H. W. (2009) *'SHALLOT-LIKE1* is a *KANADI* transcription factor that modulates rice leaf rolling by regulating leaf abaxial cell development', Plant Cell 21(3): 719-735.

Zhao, K. Y., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C. L., Toomajian, C., Zheng, H. G., Dean, C., Marjoram, P. et al. (2007) 'An *Arabidopsis* example of association mapping in structured samples', Plos Genetics 3(1)

Zhang, J., 2003 Evolution by gene duplication: an update. Trends in ecology & evolution', 18: 292-298.

Zhong, R., Taylor, J. J. and Ye, Z. H. (1997) 'Disruption of inter fascicular fibre differentiation in an *Arabidopsis* mutant', The Plant Cell 9(12): 2159-2170.

Appendix A

List of tables

Table XXVII Detail of morphological characteristics of *Begonia* section Gireoudia species

Trait	*B. nelumbiifolia*	*B. conchifolia*	*B. peltata*	*B. thiemei*	*B. carolineifolia*	*B. heracleifolia*	*B. plebeja*	*B. lindleyana*	*B. sericoneura*
Native to	Mexico	Costa Rica and Panama	Mexico and Northern C. America	Mexico to northern Honduras	Mexico	Mexican states of Durango and Tamaulipas south to Honduras	Central Mexico to Central Panama	Guatemala	Mexico, Central America to Colombia
Habitat	Moist forests	lower portions of tree trunks as epiphyte and steep rock faces	semi-desert shrub	steep forested slopes, rocky areas, road banks	semi-desert shrub	epiphytically or grows in exposed rocks	epiphytically or grows in exposed rocks	wet tropical, lower mountain	wet tropical, lower mountain
Rhizome type	Creeping	Repent	Creeping to 3cm	Repent, 8-30 cm	Erect, (root-like) stem	Creeping to 60cm	Repent	Repent	Repent
Leaf form	Peltate	Peltate	Peltate	Palmately compound	Palmately compound	Palmate lobed	Simple	Peltate or non peltate	Simple
Petiole	Abaxialized	Abaxialized	Abaxialized	Abaxialized	Abaxialized	Abaxialized	Abaxialized	Abaxialized	Abaxialized
Flower Time	Late winter to Spring	Winter through spring		March-May	Early Spring, Spring, Late Winter	Mid spring	October-January	November-March, July	December to April
Chromosome #	2n=28	2n=28	2n=28	2n=28	2n=28	2n=28	2n=28	2n=28	2n=28

Adapted from Burt-Utley (1985) in “A revision of Central American species of *Begonia* section Gireoudia (Begoniaceae).

Table XXVIII List of *ARP* Primers

Sr. #	Primer	Sequence 5´- 3´	Tm	Location on *B. nelumbiifolia* with respect to start codon (bp)
1	*SkARP*F	ATGMRIGARMGICARMGITGG	55	1
2	*SkARP*R	IACYTCCCACCAYTTNCC	54	288
3	*BARP*265F	CCAGGTCGCACAGCTAAGAG	61	434
6	*BARP*88F	GTCTCACTTCGCATGAACAC	56	254
7	*BARP*111F	CCTCAATAGGGATGCAAAGT	56	280
8	*BARP*nelA3'F1	CCAACTCTGCTTCCTTCGTG	61	662
9	*BARP*nelA3'F2	CCTTCATGTGGAGAAAATCC	57	857
10	*BARP*nelA5'F1	CTGATGACGGTTGATTTTGC	59	38
11	*BARP*nelA5'F3	TGAAGAATTTAAGGATGCTCCA	59	119
12	*BARP*nelA5'F2	TTTGGAGATCCCTACAGATACAATC	60	87
13	*BARP*1245R	TCTCACCGGCCATTAGAATC	56	1248
14	*BARP*1239R	CGGCCATTAGAATCAACGAC	56	1242
15	*BARP*1265R	TTGTTTCTGTACGGTATCCATCTC	58	1268
16	*BARP*-120F	CGGTTGATTTTGCGGAATAC	60	-120
17	*BARP*1271R	GGGATCGGTCAGCTAATACG	60	1441
18	*BARP*7F	AGGCAGAGGTGGAGAGCTAA	60	179
19	*BARP*1008R	AGTGTTTTGCTGCCCATTGT	60	1177
20	*BARP*-36F	GGATGCTCCAATGGACAGTT	60	131
21	*BARP*980R	CTTCTGCTCCTTTGCCTCTG	60	1151
22	*BARP*T1	CTTCCACTTGTTGCCGTGTT	61	415
23	*BARP*T2	AGTTCCTCCACCTTTCTAAGC	57	323
24	*BARP*T3	GCGAAGTGAGACAAGATTCC	57	249
25	MYBTAIL3 F	GGCTAAACACGGCAACAAGTG	60	221

26	MYBTAIL1 F	GGTGGAGAGCTAATGAGGACT	60	16
27	MYB TAIL2 F	CACTTCGCATGAACACACCC	60	91
28	New *BARP* T1	GGGGAGATTTTCAAACTCCCAC	59	309
29	New *BARP* T2	CGTGCCATTTCTAATACATCTC	56	402
30	New *BARP* T3	CAGYAGTCATCTCATGTCTAATC	58	481
31	-500*BARP* RT F	CCTGGTAAGTTTGAGGTTCATC	57	-500
32	-400*BARP* RT F	GAGATGTATTAGAAATGGCACG	56	-400
33	-328 *BARP* RT F	GTGGGAGTTYGAAAATCTCYCC	59	-328
34	-300 *BARP* RT F	GGAGTCCTTGATACGTAG	52	-300
35	LUX 237*BARP*2F	CAAGTCGACAGAAAATTGCACAG	58	237
36	LUX115*BARP*2F	GGGATAGTCTTGCTTAGAAAGG	57	115
37	-450*BARP* RT	GATTAGACATGAGATGACTTCTG	56	-450
38	-992*BARP* F	GAAACAGAAGTAACGTATCACTG	56	-992
39	-980 *BARP* F	GGAGGGGTTCTGACTGTTTC	58	-980
40	*BARP*2 T2	CCTATGGTTTGTACTCTCCTC	57	371
41	*BARP*2 T1	CCTCCGTTGGAAGATGTCATGA	59	451
42	LUX237*BARP*2 R	CTGTGCAATTTTCTGTCGACTTG	58	237
43	*BARP*2F deleted frag	CAGAAAATTGCACAGCTAAAGACT	57	235
44	T1 Thiemei	GGGAGATCCTCCCTAATCTG	58	-1422
45	T2 Thiemei	CGTCCTTTCCCGTCGAATC	58	-1455
46	T3 Thiemei	CCTCGTTCATCGTCTCCAAC	58	-1482
47	*BARP* 190R	CTTGTTGCCGTGTTTAGCCT	56	109

Table XXIX List of *KNOX* Primers

Sr. #	Primer	Sequence 5´- 3´	Tm	Location on *B. nelumbiifolia* with respect to start codon (bp)
1	*KNOX*Beg1	ATGGCTCATCCTCACTAYCATCGYC	61	312
2	*KNOX* Beg2	CTGTACCTKCGIAGIAGCTRACCCTT	61	1764
3	*KNOX*HOX1	TGGAGCCGCCACTACAAATG	58	1881
6	*KNOX* HOX2	TGAACCAGTTGTTGATYTGCTT	53	2026
7	*KNOX*BegBF	CAArATCATGGCTCATCCTCAC	57	308
8	*KNOX*BegBR	GAiGGgTAAGGCCATTTrTArTG	54	1890
9	girK_16F	TCCTCACTACCATCGCCTCT	58	320
10	girK1_1087R	GCGCCTCCTGTTTCAAATAA	54	1966
11	girK_28F	TCGCCTCTTAGCTGCGTATC	58	332
12	girK_1124R	TGATCCAGACCAGTTGACTCC	59	2004
13	girK_9F	TGGCTCATCCTCACTACCATC	59	313
14	gir*KNB2*_1147R	TGAACCAGCTGTTGATCTGC	56	2028
15	gir*KNB2*_1106R	TCGCCTCTTAGCTGCGTAT	55	1987
16	gir*KNB2*_1072R	CGCAATCGGGTCATTATTT	51	1953
17	*KNB1*exon2_3F	TCACTGGTGCCATTAGGTTG	56	889
18	*KNB1*exon3R	TCTGGCTTCTTTTGGCAACT	54	1843
19	*KNB2*exon1F	TGCTGACCAAATACGAGCAA	54	724
20	*KNB2*exon2F	AGGGAGCTTGTTTCCCTTGT	56	986
21	*KNB2*exon4R	TGTTGCTTCTGATGCCAGTC	56	1260
22	*KNB2*exon2_3F	TCTTCAAATATGTATGCAGGTTGC	57	779
23	*KNB2*bexon2_3F	TTCAAAcATcTATGCAGGTTGC	56	777
24	*KNB2*exon3R	ACTGGATTGGTGGAGCAGAC	58	1790
25	*KNB2*exon3Rb	GGATTGGTGGAGtAGACATC	56	1794

26	*KNB*T6	AGTGAGGATGAGCCATGATT	54	308
27	*KNB1*T5	CGCTTTTTCCTACCTTTTGA	52	388
28	*KNB1*T4	TCATGGCAGAAACATCAGTC	54	550
29	*KNB*T3	ATCTTCCCCGAGACATCCTG	58	636
30	*KNB*T2	TGAGCTCTTGCTCGTATTTG	54	700
31	*KNB*T1	AAAAAGCATGGCTTCCTTGAA	53	731
32	Tail1 *KNB* caro	CTCCGCGTTTGAGTTCAGAT	56	-193
33	Tail2 *KNB* caro	CCCACCCTCCATTGTCTCT	58	-321
34	Tail3 *KNB* caro	TGGGTGGGATTCTGCTATATG	57	-391
35	Tail4 *KNB* caro	ATGAAGAGGCTCCCGAGATT	56	-485
36	TAIL3 *KNB* herac	GGGATTCTGCTATATGGTTGTC	57	-398
37	RB Core F	GAGTTTTCATCGGGAAGGTAAAG	58	-430
38	Degenerate RB core F	GAGTTTTCATYGGGAAGGTAAAR	53	-430
39	KBOX core R	GCTCTCCCGGGTTTTATG	55	-13
40	-321*KNB* F	AGAGACAATGGAGGGTGGG	58	240
41	-440*KNB* F	GGPTTGATTACGTATGAAAGAG	53	119
42	360*KNB1*nel R	GCTTCCATGAACTGGTCCAG	58	676
43	28*KNB1* nel R	GATACGCAGCTAAGAGGCGA	58	322
44	*KNB*3 1F	GGTCCAGCCCGGTCGACTCAG	67	Not known
45	*KNB*3 F	GTTCTTCAGCTCTCGGTCTT	56	Not known
46	*KNB*3 R	AAGAGATTACGAGGCCAATC	54	Not known
47	*KNB*3ex2F	CGTCCAGGGATTTATCCAAA	54	Not known
48	*KNB*3ex4R1	AGACTTTTGCCTGGCCTCTT	56	Not known
49	*KNB*3ex4R2	TGTAATGCAACTCCCACCAG	56	Not known
50	*KNB*3ex5R	CGCCACCTTCTCCGTCTA	57	Not known
51	DQFM F1	CGGACCCGGAGCTGGAYSARTTYATGG	65	Not known
52	QINN R1	CCGCTGGTTGATGAACCARTTRTT	57	Not known

Table XXX List of plasmids

Plasmid	F Primer	R Primer	Template	Vector	Cells
pSU1	*BARP*36F	*BARP*1265R	*B.cardiocARPa*	pGEM T easy	TOP10
pSU2	*BARP*36F	*BARP*1265R	*B.lyman-smithi*	pGEM T easy	TOP10
pSU3	*BARP*36F	*BARP*1265R	*B.multinervia*	pGEM T easy	TOP10
pSU4	*BARP*36F	*BARP*1265R	*B.plebja*	pGEM T easy	TOP10
pSU5	*BARP*36F	*BARP*1265R	*B.sarcophylla*	pGEM T easy	TOP10
pSU6	*BARP*36F	*BARP*1265R	*B.sericonaeura*	pGEM T easy	TOP10
pSU7	*BARP*36F	*BARP*1265R	*B.stigmosa*	pGEM T easy	TOP10
pSU8	AD2	Tail3	*B.carolinifolia*	pGEM T easy	TOP10
pSU9	AD4	Tail3	*B.carolinifolia*	pGEM T easy	TOP10
pSU10	AD4	Tail3	*B.carolinifolia*	pGEM T easy	TOP10
pSU11	AD3	Tail3	*B.carolinifolia*	pGEM T easy	TOP10
pSU12	AD5	Tail3	*B.carolinifolia*	pGEM T easy	TOP10
pSU13	AD3	Tail3	*B.theimei*	pGEM T easy	TOP10
pSU14	AD3	Tail3	*B.theimei*	pGEM T easy	TOP10
pSU15	AD6	Tail3	*B.theimei*	pGEM T easy	TOP10
pSU17	LFY1	LFY2	*B.carolinifolia*	pGEM T easy	TOP10
pSU18	LFY1	LFY2	*B. conchifolia*	pGEM T easy	TOP10
pSU19	*BARP*111F	*BARP*1265R	*B.nelumbiifolia*	pGEM T easy	TOP10
pSU20	*BARP*111F	*BARP*1265R	*B.theimei*	pGEM T easy	TOP10
pSU21	SmallKBOX1F	KBOXexon2R	*B.theimei*	pGEM T easy	TOP10
pSU22	SmallKBOX1F	KBOXexon2R	*B.carolinifolia*	pGEM T easy	TOP10
pSU23	88F	1008R	*B.kellemanii*	pGEM T easy	TOP10
pSU24	BnelA3'F1	980R	*B. peltata*	pGEM T easy	TOP10
pSU25	88F	1008R	*B. theimei*	pGEM T easy	TOP10
pSU26			*pTRV2*	pGEM T easy	TOP10

pSU27	AD6	Tail3	*B.heracleifolia*	pGEM T easy	TOP10
pSU28	AD1	Tail3	*B.peltata*	pGEM T easy	TOP10
pSU29	AD5	Tail3	*B.peltata*	pGEM T easy	TOP10
pSU30	*BARP*120F	*BARP*1265R	*B.nelumbiifolia*	pGEM T easy	TOP10
pSU31	*BARP*111F	*BARP*1245R	*B.nelumbiifolia*	pGEM T easy	TOP10
pSU32	*BARP*111F	*BARP*1245R	*B.kellemanii*	pGEM T easy	TOP10
pSU33	*BARP*111F	*BARP*1245R	*B.peltata*	pGEM T easy	TOP10
pSU34	*BARP*111F	*BARP*1245R	*B.stigmosa*	pGEM T easy	TOP10
pSU35	*BARP*111F	*BARP*1245R	*B.cardiocARPa*	pGEM T easy	TOP10
pSU36	*BARP*120F	*BARP*1265R	*B.cardiocARPa*	pGEM T easy	TOP10
pSU37	*BARP*7F	*BARP*1265R	*B.nelumbiifolia*	pGEM T easy	TOP10
pSU38	*BARP*7F	*BARP*1265R	*B.theimei*	pGEM T easy	TOP10
pSU39	*BARP*120F	*BARP*1245R	*B.kellemanii*	pGEM T easy	TOP10
pSU40	QINF1	DQR	*B.heracleifolia*	pGEM T easy	TOP10
pSU41	QINF2	DQR	*B.carolinifolia*	pGEM T easy	TOP10
pSU42	LFY1	LFY2	*B.carolinifolia*	pGEM T easy	TOP10
pSU43	*BARP*111F	*BARP*1265R	*B.heracleifolia*	pGEM T easy	TOP10
pSU44	LFY1	LFY2	*B.carolinifolia*	pGEM T easy	TOP10
pSU45	*BARP*36F	*BARP*1265R	*B.conchifolia*	pGEM T easy	TOP10
pSU46	*BARP*36F	*BARP*1265R	*B.plebeja*	pGEM T easy	TOP10
pSU47	*BARP*36F	*BARP*1265R	*B.carolinifolia*	pGEM T easy	TOP10
pSU48	*BARP*36F	*BARP*1265R	*B.heracleifolia*	pGEM T easy	TOP10
pSU49	*BARP*84F	*BARP*1008R	*B.carolinifolia*	pGEM T easy	TOP10
pSU60	*KNB*31F	*KNB*3R	*B.carolinifolia*	pGEM T easy	DH5ALPHA
pSU61	*KNB*31F	*KNB*3R	*B.conchifolia*	pGEM T easy	DH5ALPHA
pSU62	*KNB*31F	*KNB*3R	*B.plebeja*	pGEM T easy	DH5ALPHA
pSU63	*KNB*31F	*KNB*3R	*B.nelumbiifolia*	pGEM T easy	DH5ALPHA
pSU64	TOPO *AS1* F	TOPO *AS1* R	*Arabidopsis*	pENTR D TOPO	TOP10

pSU65	TOPO*BARP*1 F	TOPO *BARP*1 R	*B.nelumbiifolia*	pENTR D TOPO	TOP10
pSU66	SYNTHETIC GENE		*AS1 MODIFIED*	pJ201	DH5ALPHA
pSU67	T3 PROMOTER	T7 PROMOTER	*Arabidopsis*	pGEM T easy	DH5ALPHA
pSU68	TOPO *AS1* F	TOPO *AS1* R	*Arabidopsis*	pGEM T easy	DH5ALPHA
pSU69	TOPO *AS1* F	TOPO *AS1* R	*Arabidopsis*	GATEWAY	TOP10
pSU70	TOPO*BARP*1 F	TOPO *BARP*1 R	*B.nelumbiifolia*	GATEWAY	TOP10
Psu80	Infusion F	Infusion R	*Arabidopsis*	pGEM T easy	TOP10
pSU81	Infusion80.4F	Infusion80.4R	*pSU64.1*	pENTR D TOPO	TOP10
pSU82	Infusion80.4F	Infusion80.4R	*pSU64.8*	pENTR D TOPO	TOP10
pSU83	*BARP*-c-ter F	*BARP*-c-ter R	*B.heracleifolia*	PCR2.1 TOPO	TOP10
pSU84	*BARP*-c-ter F	*BARP*-c-ter R	*B.carolinifolia*	PCR2.1 TOPO	TOP10
pSU85	*BARP*-c-ter F	*BARP*-c-ter R	*B.peltata*	PCR2.1 TOPO	TOP10
pSU86	*KNB1* EXON3F	*KNB1* EXON3R	*B.heracleifolia*	PCR2.1 TOPO	TOP10
pSU87	*KNB1* EXON3F	*KNB1* EXON3R	*B.carolinifolia*	PCR2.1 TOPO	TOP10
pSU88	*KNB1* EXON3F	*KNB1* EXON3R	*B.peltata*	PCR2.1 TOPO	TOP10
pSU89	*KNB2* EXON3F	*KNB2* EXON3R	*B.heracleifolia*	PCR2.1 TOPO	TOP10
pSU90	*KNB2* EXON3F	*KNB2* EXON3R	*B.carolinifolia*	PCR2.1 TOPO	TOP10
pSU91	*KNB2* EXON3F	*KNB2* EXON3R	*B.peltata*	PCR2.1 TOPO	TOP10
pSU92	PDSF1	PDSR1	*B.heracleifolia*	PCR8 GW TOPO	TOP10
pSU93	PDSF5	PDSR5	*B.heracleifolia*	PCR8 GW TOPO	TOP10
pSU94	PDSF6	PDSR6	*B.heracleifolia*	PCR8 GW TOPO	TOP10
pSU95	TOPO *AS1* F	TOPO *AS1* R	*pSU81.1*	pENTRE D TOPO	TOP10
pSU96	TOPO *AS1* F	TOPO *AS1* R	*pSU69.3*	GW	GV3101
pSU97	TOPO*BARP*1 F	TOPO *BARP*1 R	*pSU70.2*	GW	GV3101
pSU98	TOPO *AS1* F	TOPO *AS1* R	*pSU72.1*	GW	GV3101

Table XXXI List of Species used in *KNOX* genes sequence analysis

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	NM180620
Cardamine hirsuta	*ChSTM*	DQ512732
Brassica napus	*BnHD1*	Z29073
	BnSTM1	GU480584
Brassica oleracea	*BoSTM1*	AF527947
Brassica rapa	*BrSTM1*	GU480585
Glycine max	*KNT1*	DQ857339
Helianthus	*Htknot1*	AJ519674
Ipomoea nil	*STM*	EU672819
	InPKn1	AB015999
	InPKn2	AB016000
	InPKn3	AB016002
Ipomoea batatas	*IBKN1*	AB283027
Solanum lycopersicum	*TKN1*	U32247
	TKN2	U76407
	TKN3	U76408
	TKN4	AF375968
	LeT6	AF000141
	LET12	AF000142
	THox2	U76410
Malus X domestica	*MdKN11*	Z71978
	MdKNAP3	Z71980

	MtKn1	AF308454
Medicago truncatula	*NTH1*	AB025573
Nicotiana tabacum	*NTH9*	AB025713
	NTH15	AB004785
	NTH20	AB025714
	NTH22	AB025715
	NTH23	AB004797
	PhSTM1	AY112704
Petunia x hybrida	*Hop1*	AF063307
Petunia x hybrida	*HERMIT*	GQ409545
Pisnum sativum	*ARBORKNOX1*	AY755413
Populus alba x Populus tremula	*PtdKn2*	AY684937
Populus balsamifera x Populus deltoides	*PtdKn3*	AY684938
Populus tomentosa	*POTH1*	U65648
Solanum tuberosum	*SdSTM1*	AY655752
Streptocarpus dunnii	*SrSTM1*	AY655753
Streptocarpus rexii	*SsSTM2*	AY655754
Streptocarpus saxorum		
Monocot		
Hordeum vulgare	*Hvkn1*	AF544045
Oryza sativa	*OSH1*	AC145380
	OSH6	AB028883
Triticum aestivum	*TaKnox1b*	AF224499
Zea mays	*KN3*	FU959496
	liguless3	AF100455
	liguless4a	AF457118
	liguless4b	AF457119
Gymnosperm		
Picea	*PaKn*	AF063248
Pinus taeda	*PtKN1*	AY680402
	PtKN2	AY680403
	PtKN3	AY680404
Pteridophyte		

Ceratopteris richardii	*CRKNOX1* *CRKNOX2* *CRKNOX3*	AB043954 AB043956 AB043957
Lycophyte		
Selaginella kraussiana	*SkKNOX1* *SkKNOX2* *SkKNOX3*	AY667449 AY667450 AY667451
Selaginella moellendorffii	*SmKNOX3*	XM-002993541
Bryophyte		
Physcomitrella patens	*MKN2* *MKN4* *MKN1-3* *MKN6* *MKN5*	AF285147 AF284817 AF285148 XM-001765523 AB266747
Green algae		
Acetabularia acetabulum	*AaKNOX1*	AF170172

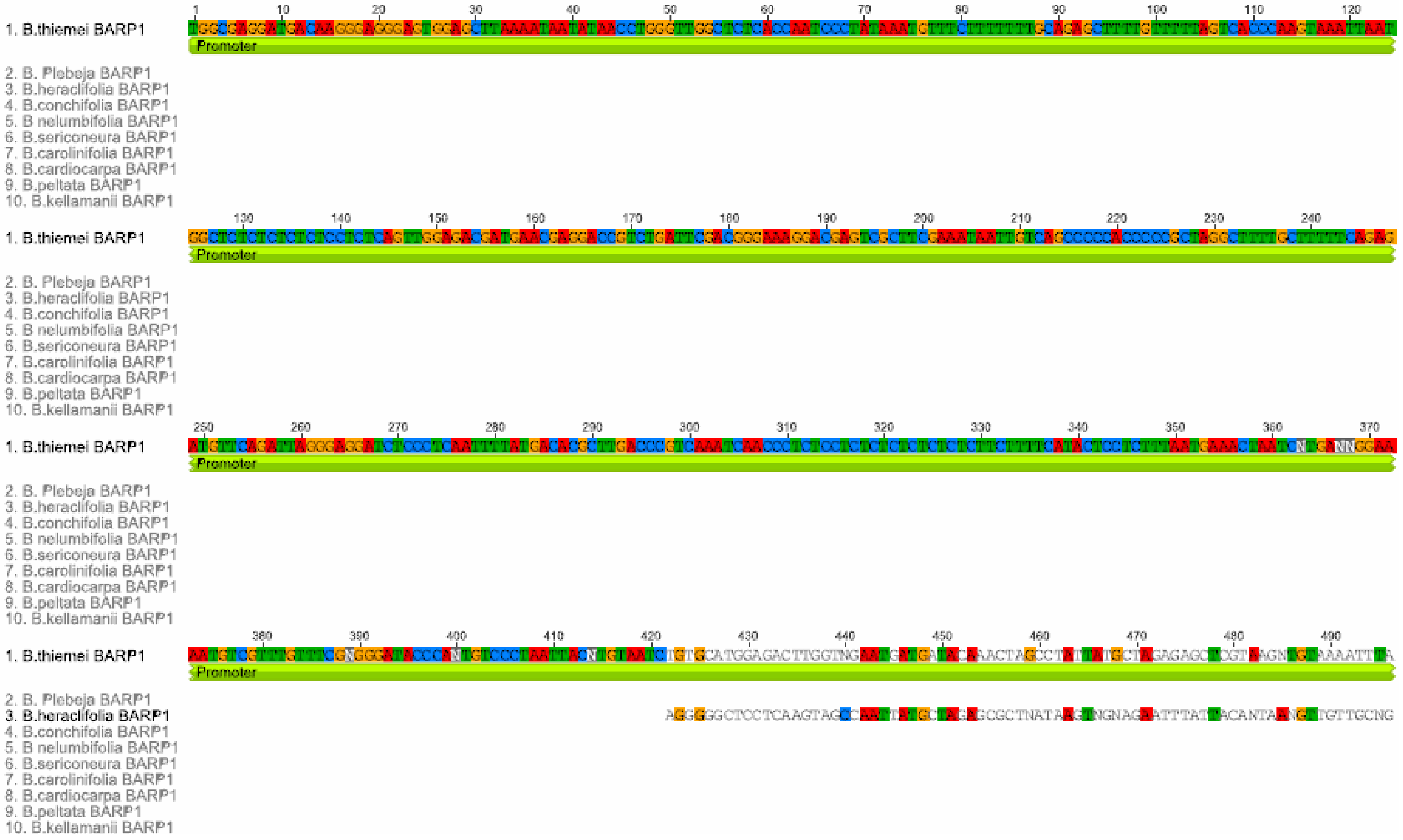
Table XXXII List of Species used in *ARP* genes sequence analysis

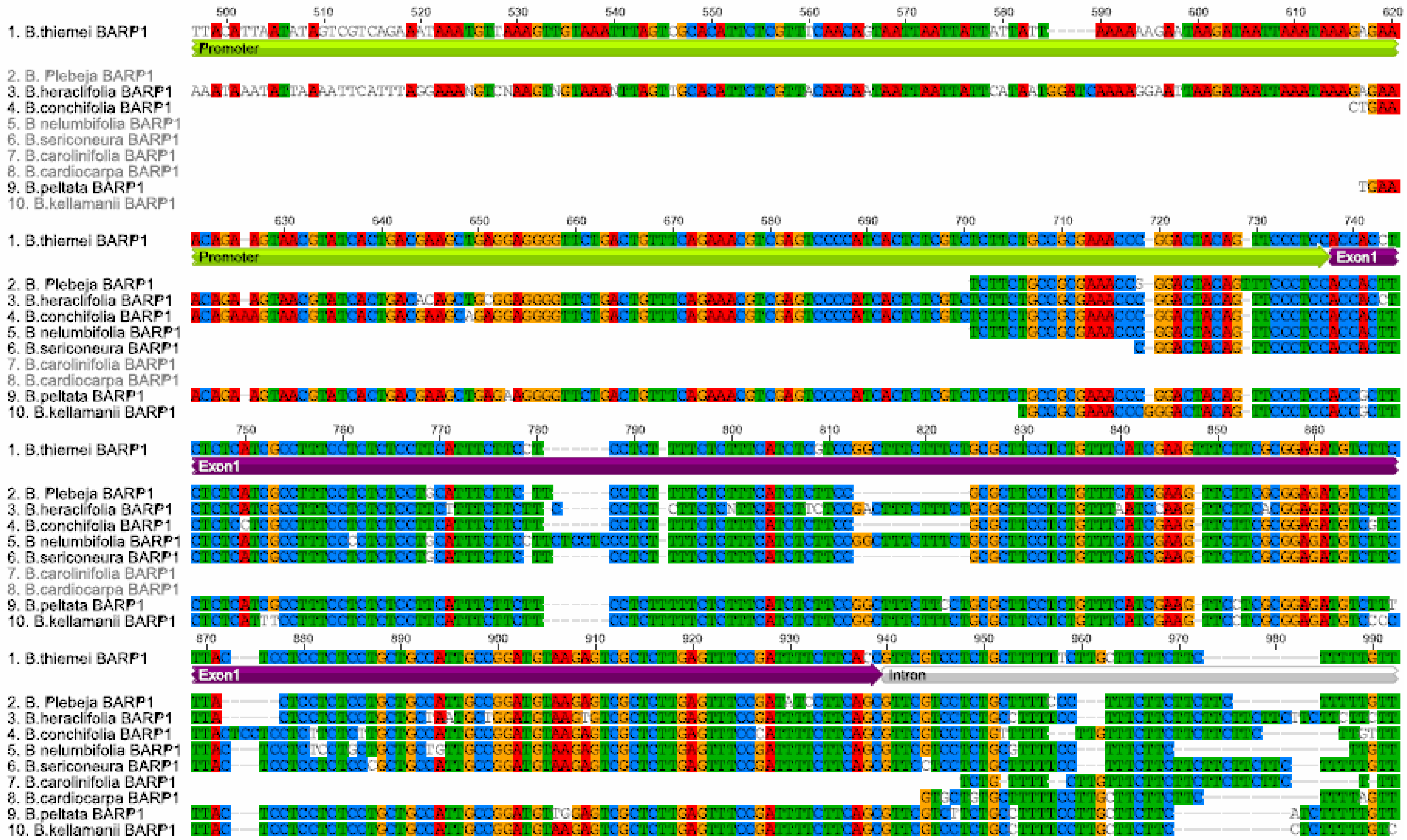
Species	**Gene name**	**Accession # / Locus name**
Selaginella kraussiana	*PHAN*	AY667452
Zea mays	*RS2*	AF143447
Antirrhinum majus	*PHAN*	AJ005586
Carica papaya	*ARP1*	evm.TU.supercontig_333.1
Carica papaya	*ARP2*	evm.TU.supercontig_333.1
Arabidopsis lyrata	*AS1*	XM_002881476
Arabidopsis thaliana	*AS1*	AF175996
Manihot esculenta	*ARP*	cassava4.1_010610m.g
Ricinus communis	*ARP1*	28603.t000008
Ricinus communis	*ARP2*	**29767.t000001**
Vitis vinifera	*ARP*	VV78X006745
Populus trichocarpa	*ARP 1*	POPTR_0006s08610
Medicago truncatula	*ARP*	AF308453
Glycine max	*PHAN1*	AY790252
Glycine max	*PHAN2*	AY790253
Lotus japonicus	*PHANa*	AY790244
Lotus japonicus	*PHANb*	AY790245
Scopelogena bruynsii	*ARPa*	FJ571376
Scopelogena bruynsii	*ARPb*	FJ571391
Carruanthus ringens	*ARPa*	FJ571373
Carruanthus ringens	*ARPb*	FJ571390
Pleiospilos simulans	*ARPa*	FJ571378
Pleiospilos simulans	*ARPb*	FJ571394
Cucumis sativus	*ARP*	csa.16957326

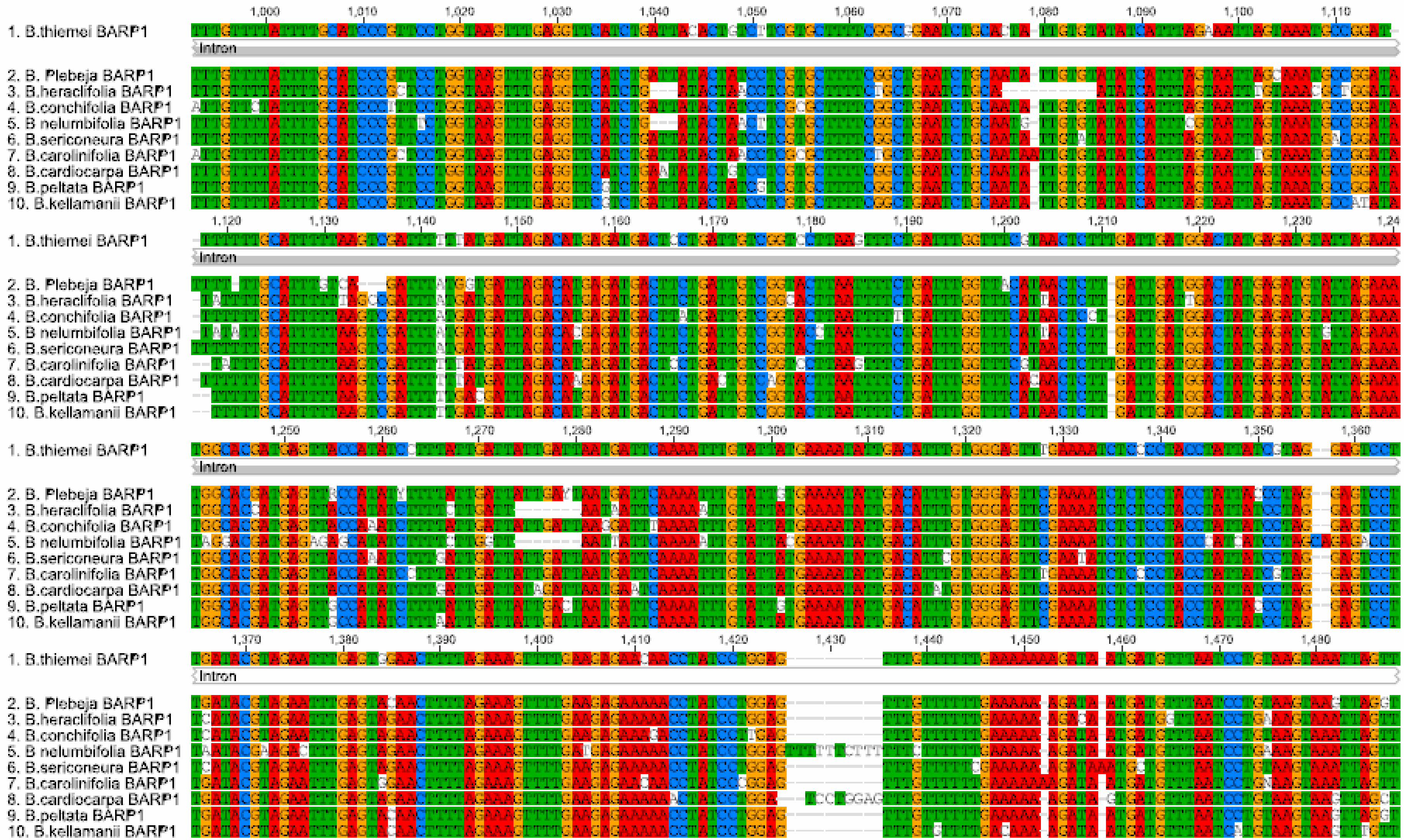
Appendix B

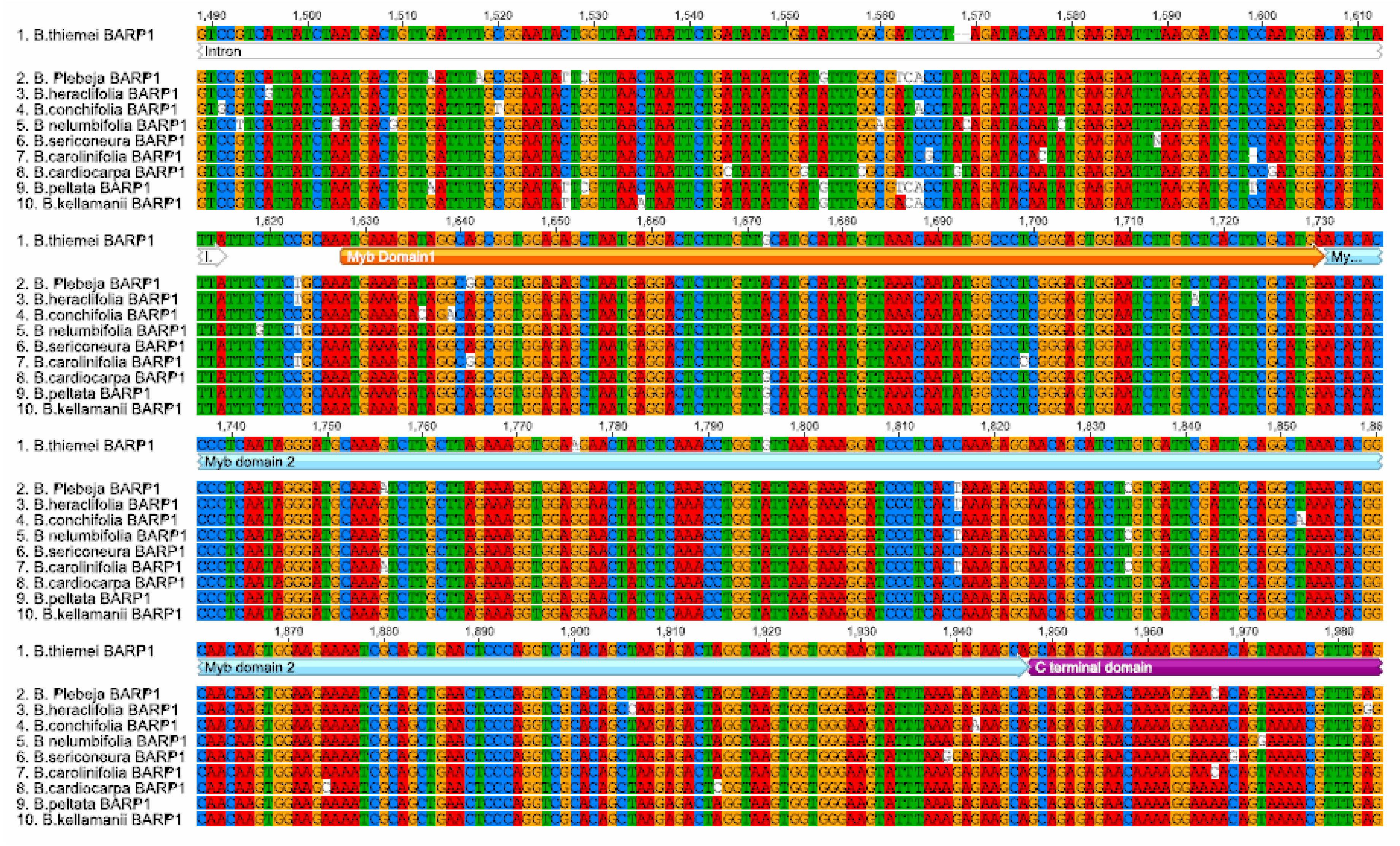
List of figures

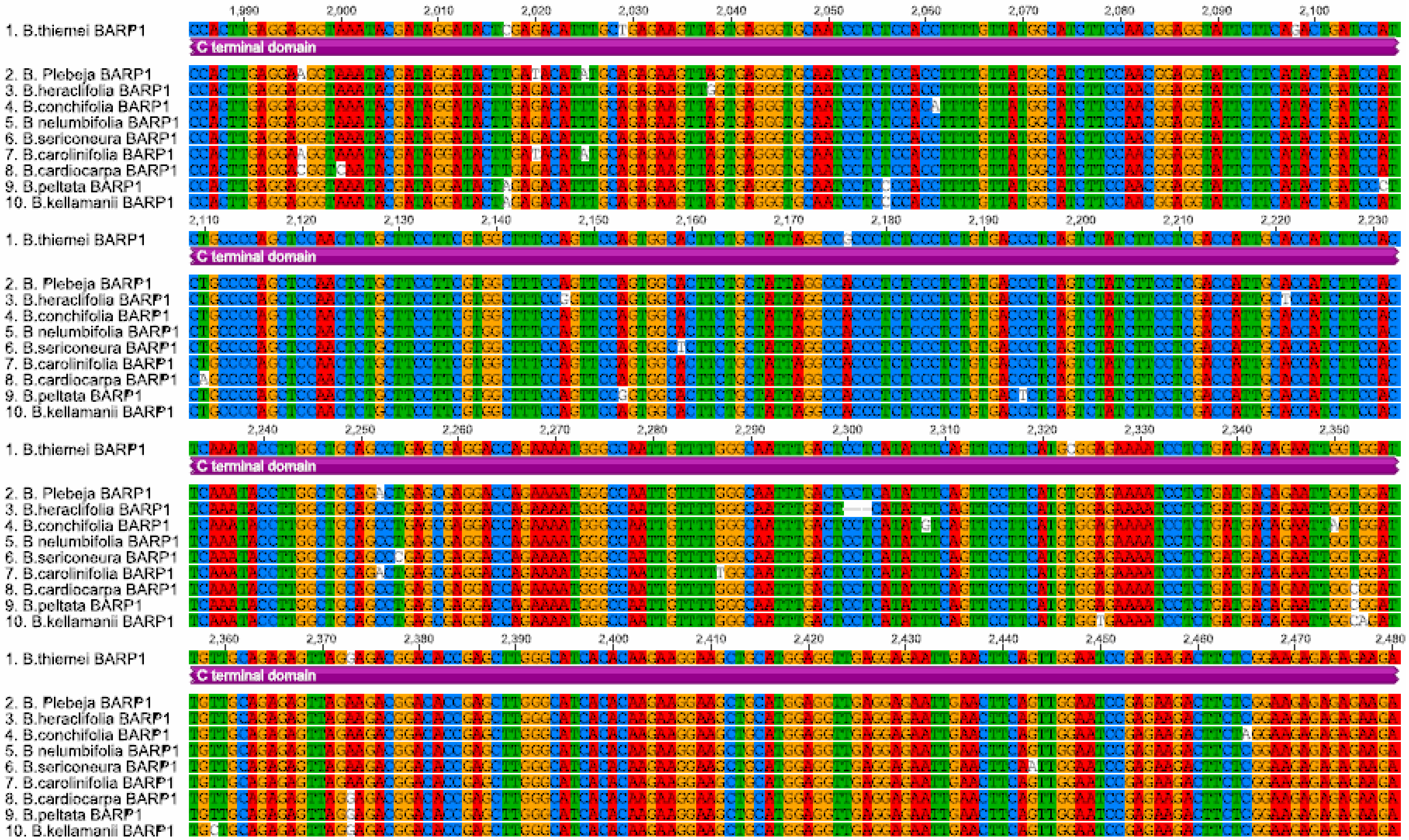
Figure CXIII *ARP* alignments.
***BARP* genes showed a high degree of similarity to each other in coding regions as well as in non coding parts of the gene indicating that they are orthologs.**




1. B.thiemei BARP1
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1
Promoter
Exon1
Intron
500
510
520
530
540
550
560
570
580
590
600
610
620
630
640
650
660
670
680
690
700
710
720
730
740
750
760
770
780
790
800
810
820
830
840
850
860
870
880
890
900
910
920
930
940
950
960
970
980
990


1. B.thiemei BARP1
Intron
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1
1,000
1,010
1,020
1,030
1,040
1,050
1,060
1,070
1,080
1,090
1,100
1,110
1,120
1,130
1,140
1,150
1,160
1,170
1,180
1,190
1,200
1,210
1,220
1,230
1,24
1,250
1,260
1,270
1,280
1,290
1,300
1,310
1,320
1,330
1,340
1,350
1,360
1,370
1,380
1,390
1,400
1,410
1,420
1,430
1,440
1,450
1,460
1,470
1,480


1,490
1,500
1,510
1,520
1,530
1,540
1,550
1,560
1,570
1,580
1,590
1,600
1,610
1. B.thiemei BARP1
Intron
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1
1,620
1,630
1,640
1,650
1,660
1,670
1,680
1,690
1,700
1,710
1,720
1,730
1. B.thiemei BARP1
I.
Myb Domain1
My...
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1
1,740
1,750
1,760
1,770
1,780
1,790
1,800
1,810
1,820
1,830
1,840
1,850
1,86
1. B.thiemei BARP1
Myb domain 2
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1
1,870
1,880
1,890
1,900
1,910
1,920
1,930
1,940
1,950
1,960
1,970
1,980
1. B.thiemei BARP1
Myb domain 2
C terminal domain
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1


1. B.thiemei BARP1
C terminal domain
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1
1,990
2,000
2,010
2,020
2,030
2,040
2,050
2,060
2,070
2,080
2,090
2,100
2,110
2,120
2,130
2,140
2,150
2,160
2,170
2,180
2,190
2,200
2,210
2,220
2,230
2,240
2,250
2,260
2,270
2,280
2,290
2,300
2,310
2,320
2,330
2,340
2,350
2,360
2,370
2,380
2,390
2,400
2,410
2,420
2,430
2,440
2,450
2,460
2,470
2,480

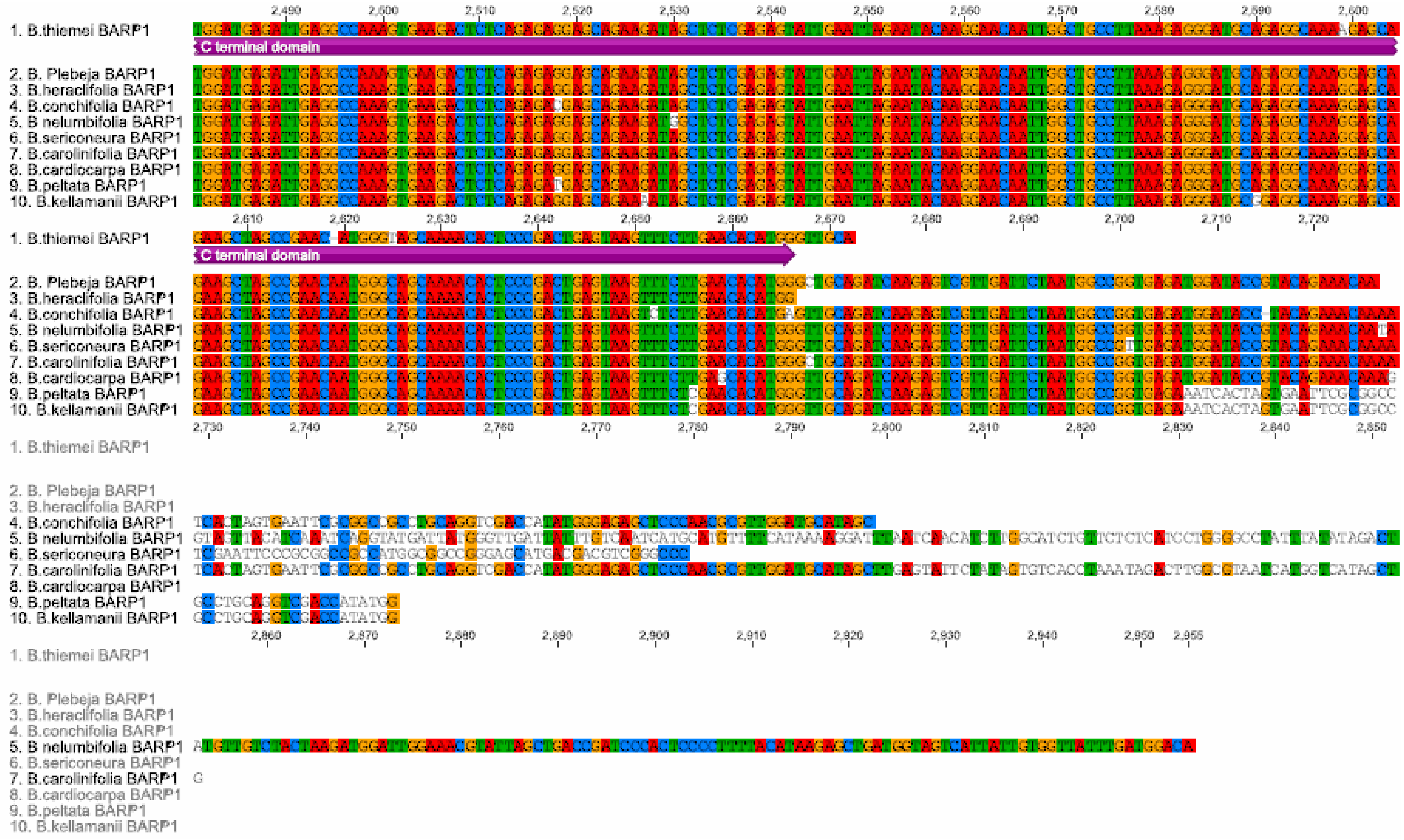
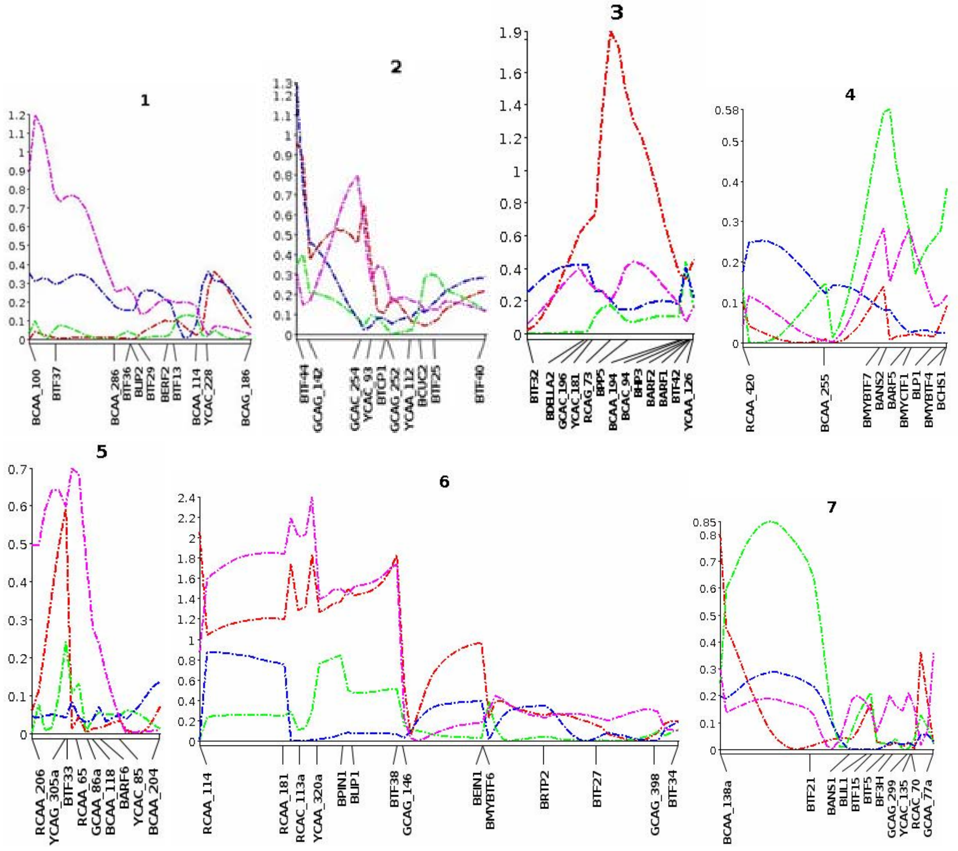
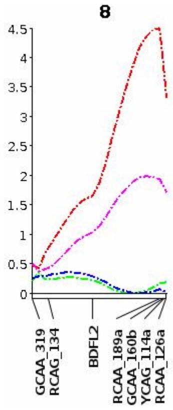
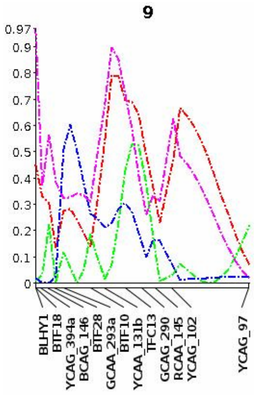
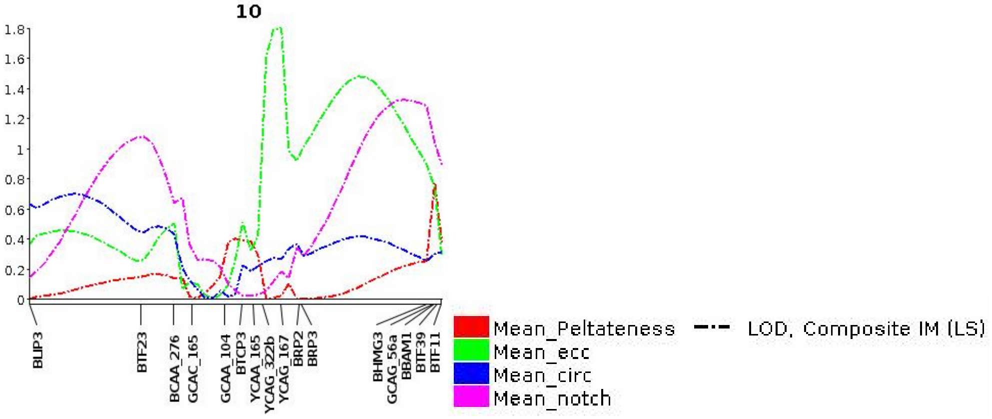

2,490
2,500
2,510
2,520
2,530
2,540
2,550
2,560
2,570
2,580
2,590
2,600
1. B.thiemei BARP1
C terminal domain
2. B. Plebeja BARP1
3. B.heraclifolia BARP1
4. B.conchifolia BARP1
5. B nelumbifolia BARP1
6. B.sericoneura BARP1
7. B.carolinifolia BARP1
8. B.cardiocarpa BARP1
9. B.peltata BARP1
10. B.kellamanii BARP1
2,610
2,620
2,630
2,640
2,650
2,660
2,670
2,680
2,690
2,700
2,710
2,720
2,730
2,740
2,750
2,760
2,770
2,780
2,790
2,800
2,810
2,820
2,830
2,840
2,850
2,860
2,870
2,880
2,890
2,900
2,910
2,920
2,930
2,940
2,950
2,955

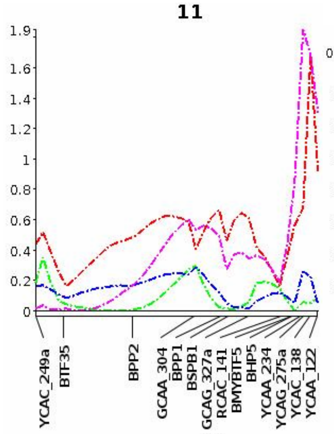
Figure CXIV CIM for ARB312 for leaf peltateness, eccentricity, notchiness and circularity.
The graph indicates that peltateness and notchiness boundary may be controlled by *BARP1* locus or locus/loci controlling these traits is very close to it on chromosome 13.

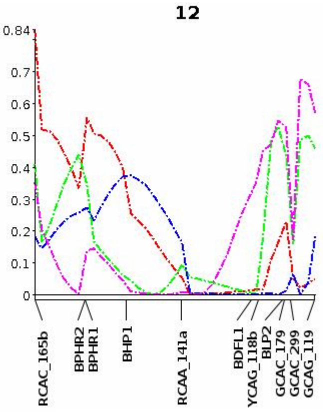


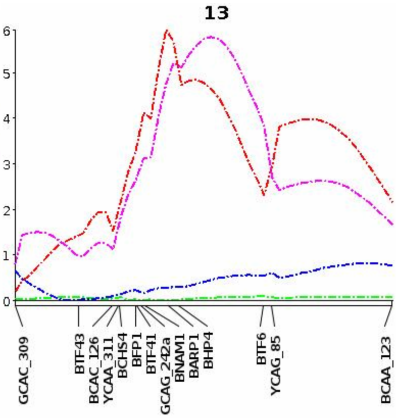

8
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
GCAA_319
RCAG_134
BDFL2
RCAA_189a
GCAA_160b
YCAG_114a
RCAA_126a

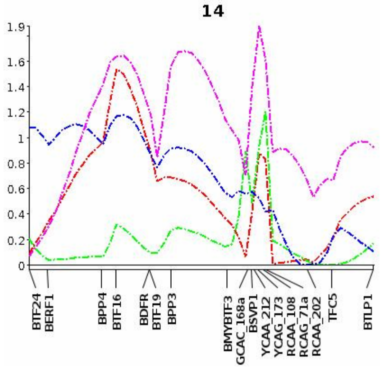

9
0.97
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
BLHY1
BTF18
YCAG_394a
BCAG_146
BTF28
GCAA_293a
BTF10
YCAA_131b
TFC13
GCAG_290
RCAA_145
YCAG_102
YCAG_97


10
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
BLIP3
BTF23
BCAA_276
GCAC_165
GCAA_104
BTCP3
YCAA_165
YCAG_322b
YCAG_167
BRP2
BRP3
BHMG3
GCAG_56a
BBAM1
BTF39
BTF11
Mean_Peltateness
Mean_ecc
Mean_circ
Mean_notch
LOD, Composite IM (LS)


11
1.9
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
YCAC_249a
BTF35
BPP2
GCAA_304
BPP1
BSPB1
GCAG_327a
RCAC_141
BMYBTF5
BHP5
YCAA_234
YCAG_275a
YCAC_138
YCAA_122

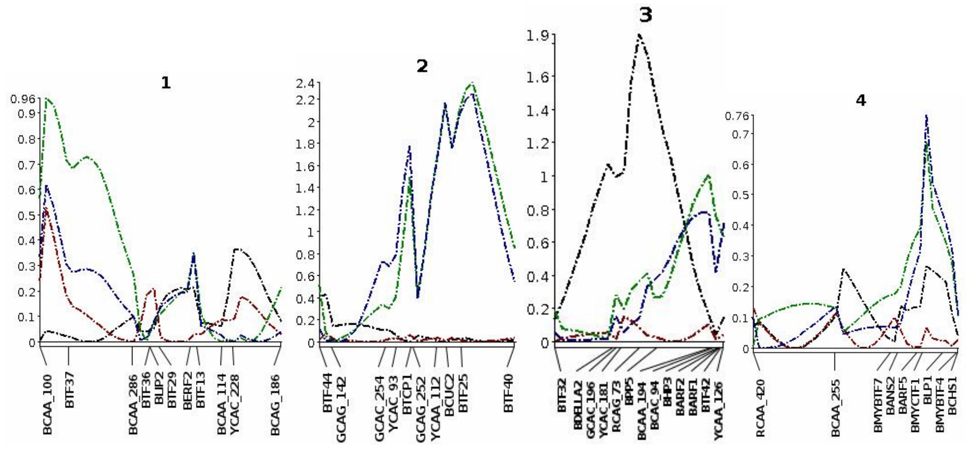

12
0.84
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
RCAC_165b
BPHR2
BPHR1
BHP1
RCAA_141a
BDFL1
YCAG_118b
BLP2
GCAC_179
GCAC_299
GCAG_119

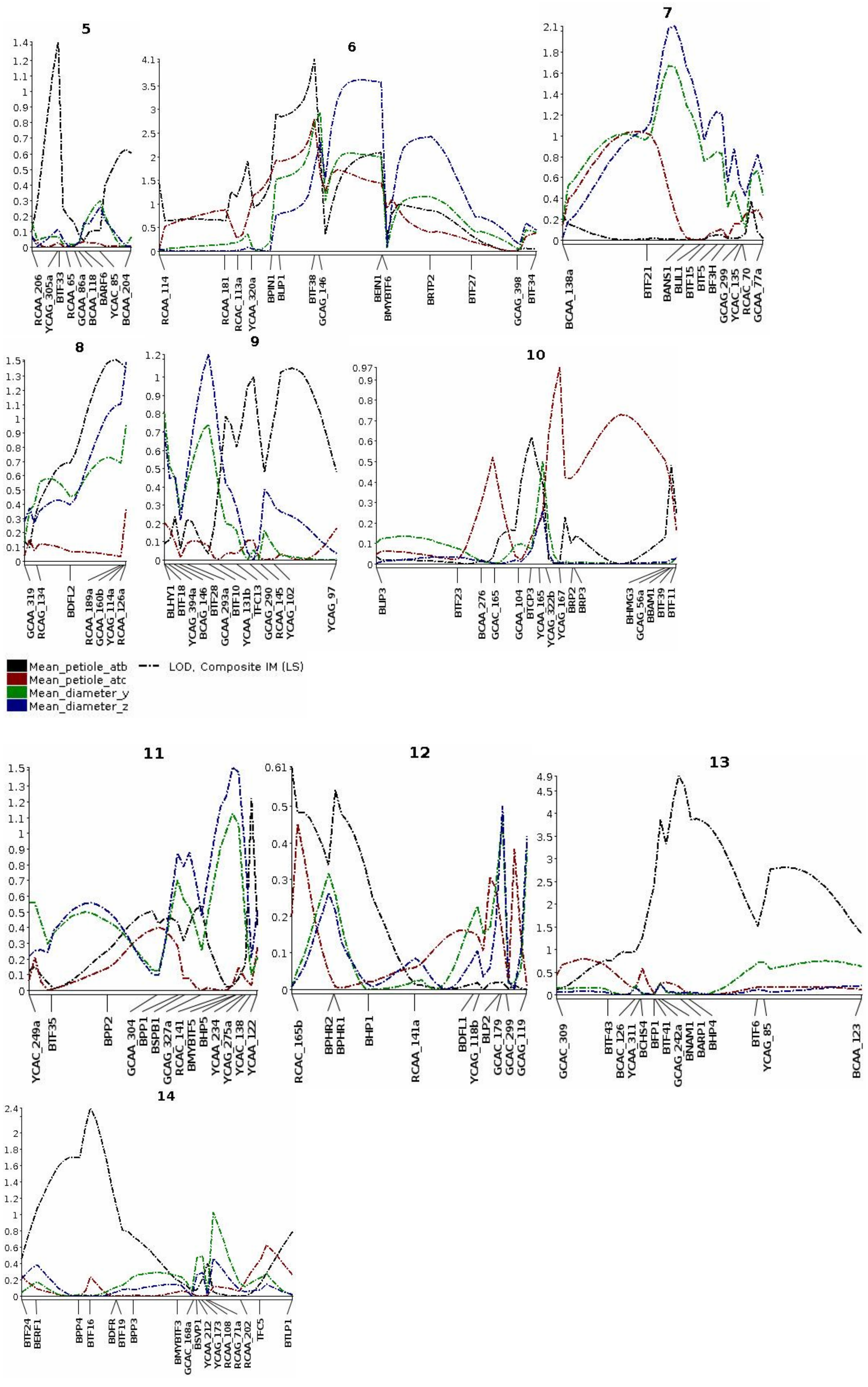

13
6
5
4
3
2
1
0
GCAC_309
BTF43
BCAC_126
YCAA_311
BCHS4
BFP1
BTF41
GCAG_242a
BNAM1
BARP1
BHP4
BTF6
YCAG_85
BCAA_123


14
1.9
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
BTF24
BERF1
BPP4
BTF16
BDFR
BTF19
BPP3
BMYBTF3
GCAC_168a
BSVP1
YCAA_212
YCAG_173
RCAA_108
RCAG_71a
RCAA_202
TFC5
BTLP1

🟥 Mean_Peltateness -·- LOD, Composite IM (LS)
🟩 Mean_ecc
🟦 Mean_circ
🟪 Mean_notch

Figure CXV CIM for ARB312 for distance from petiole attachment point to leaf boundary, petiole attachment point to centre, leaf length and leaf width.
The graph indicates that distance from petiole attachment point to leaf boundary may be under control of *BARP1* locus or locus controlling these traits is very close to it on chromosome 13.




5
RCAA_206
YCAG_305a
BTF33
RCAA_65
GCAA_86a
BCAA_118
BARF6
YCAC_85
BCAA_204
6
RCAA_114
RCAA_181
RCAC_113a
YCAA_320a
BPIN1
BLIP1
BTF38
GCAG_146
BEIN1
BMYBTF6
BRTP2
BTF27
GCAG_398
BTF34
7
BCAA_138a
BTF21
BANS1
BLIL1
BTF15
BTF5
BF3H
GCAG_299
YCAC_135
RCAC_70
GCAA_77a
8
GCAA_319
RCAG_134
BDFL2
RCAA_189a
GCAA_160b
YCAG_114a
RCAA_126a
9
BLHY1
BTF18
YCAG_394a
BCAG_146
BTF28
GCAA_293a
BTF10
YCAA_131b
TFC13
GCAG_290
RCAA_145
YCAG_102
YCAG_97
10
BLIP3
BTF23
BCAA_276
GCAC_165
GCAA_104
BTCP3
YCAA_165
YCAG_322b
YCAG_167
BRP2
BRP3
BHMG3
GCAG_56a
BBAM1
BTF39
BTF11
Mean_petiole_atb
Mean_petiole_atc
Mean_diameter_y
Mean_diameter_z
LOD, Composite IM (LS)
11
YCAC_249a
BTF35
BPP2
GCAA_304
BPP1
BSPB1
GCAG_327a
RCAC_141
BMYBTF5
BHP5
YCAA_234
YCAG_275a
YCAC_138
YCAA_122
12
RCAC_165b
BPHR2
BPHR1
BHP1
RCAA_141a
BDFL1
YCAG_118b
BLP2
GCAC_179
GCAC_299
GCAG_119
13
GCAC_309
BTF43
BCAC_126
YCAA_311
BCHS4
BFP1
BTF41
GCAG_242a
BNAM1
BARP1
BHP4
BTF6
YCAG_85
BCAA_123
14
BTF24
BERF1
BPP4
BTF16
BDFR
BTF19
BPP3
BMYBTF3
GCAC_168a
BSVP1
YCAA_212
YCAG_173
RCAA_108
RCAG_71a
RCAA_202
TFC5
BTLP1

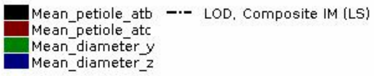
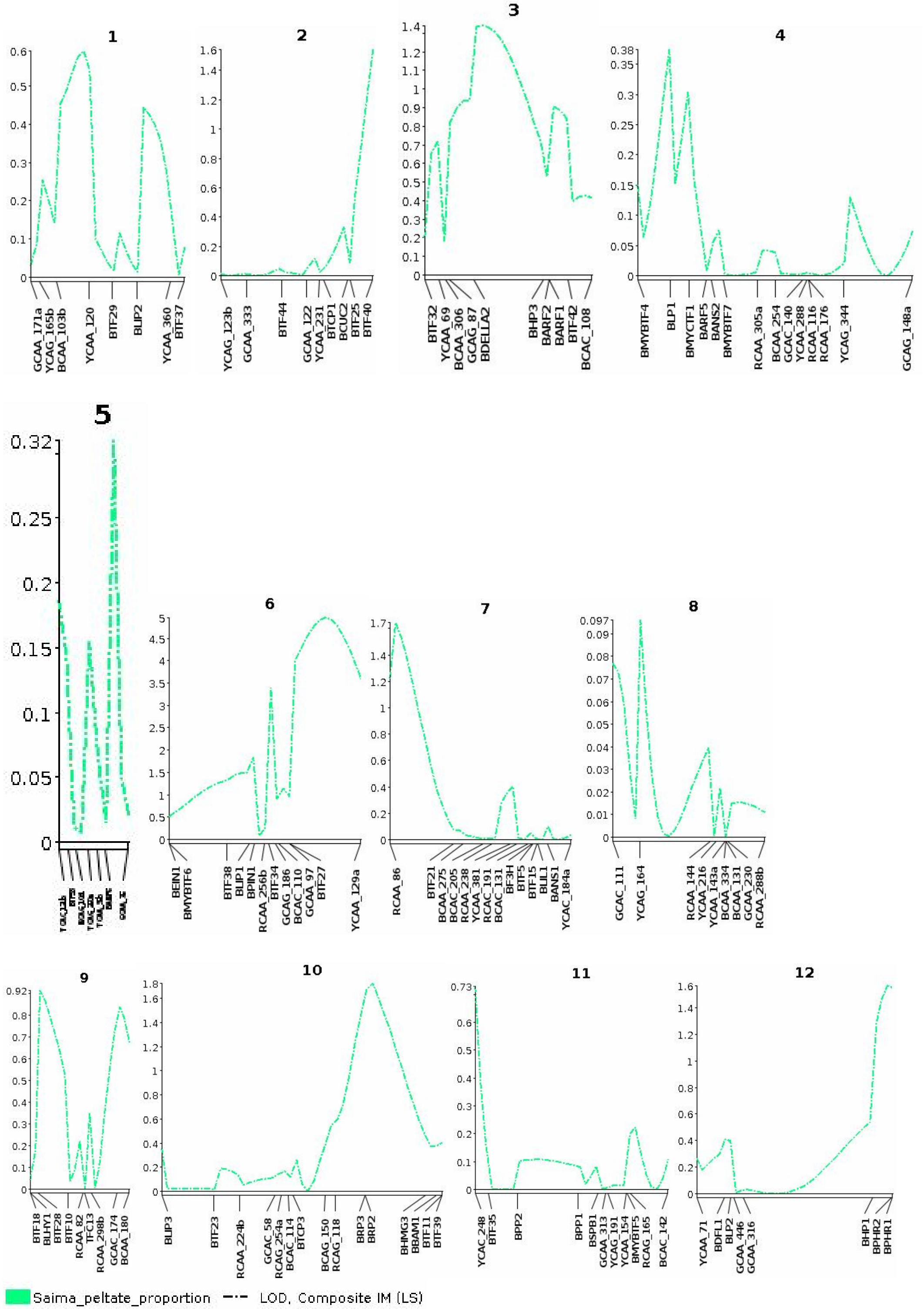

Mean_petiole_atb
Mean_petiole_atc
Mean_diameter_y
Mean_diameter_z
LOD, Composite IM (LS)

Figure CXVI CIM for BOB 360 for peltateness.
The graph indicates that peltateness is not linked to *BARP1* in this population.



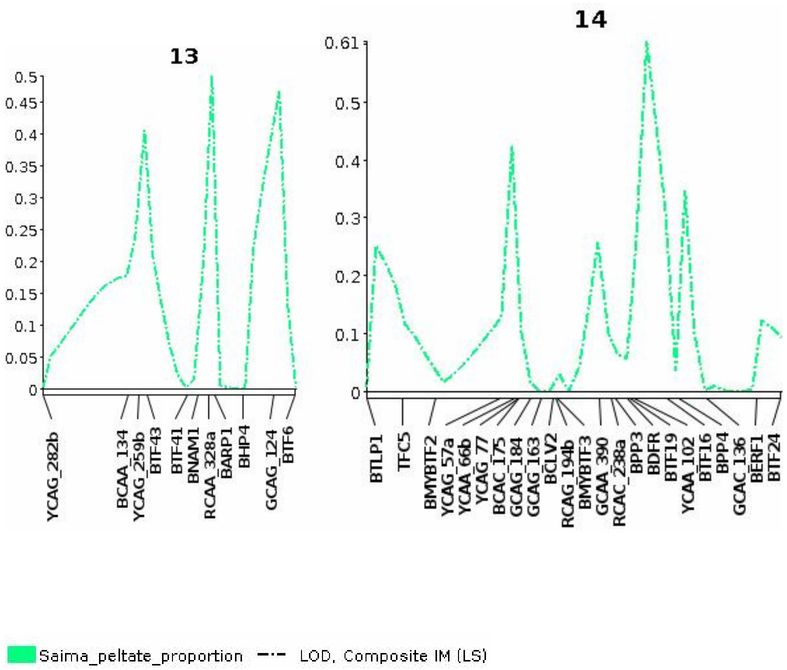
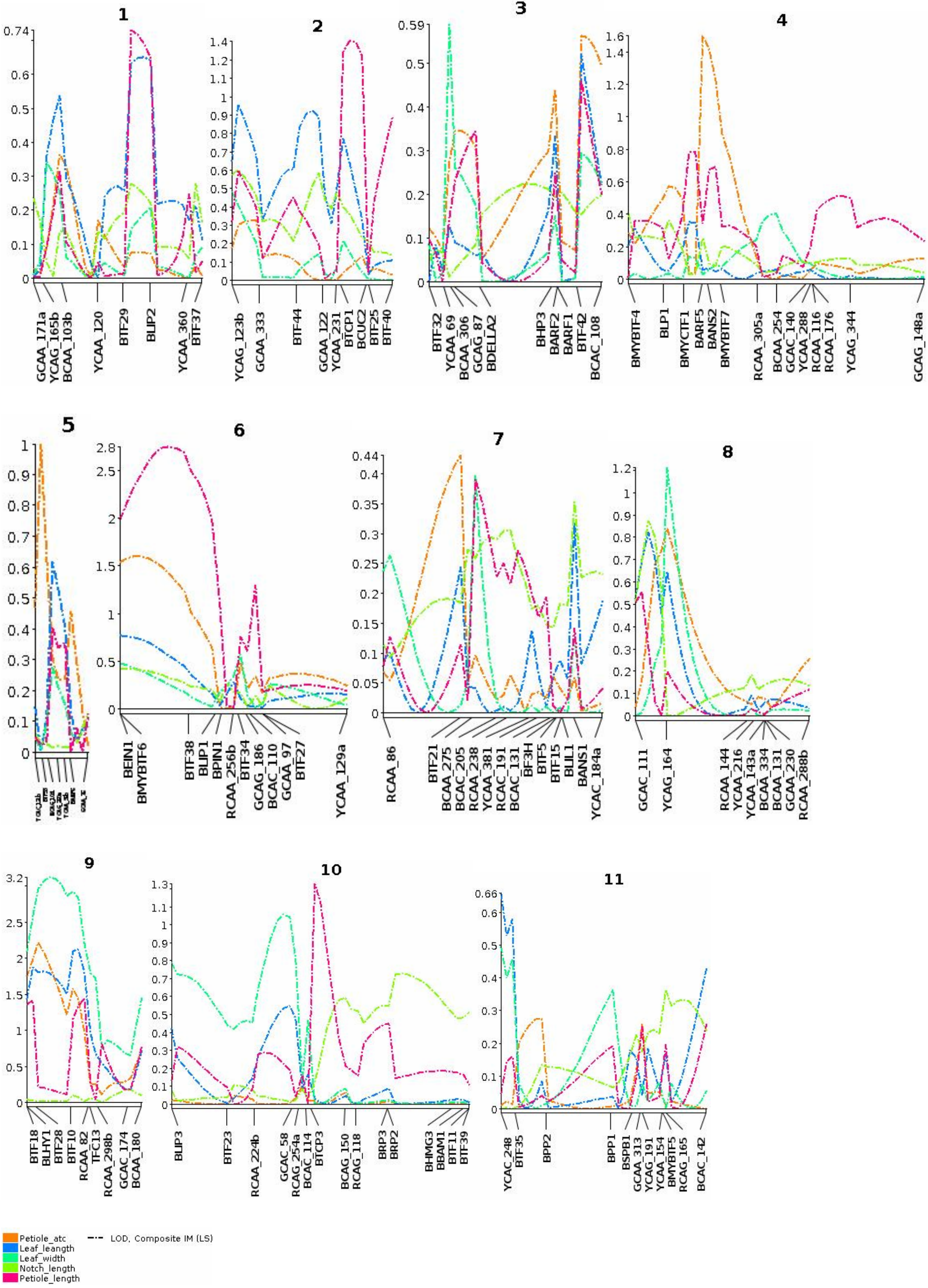

13
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
YCAG_282b
BCAA_134
YCAG_259b
BTF43
BTF41
BNAM1
RCAA_328a
BARP1
BHP4
GCAG_124
BTF6
14
0.61
0.5
0.4
0.3
0.2
0.1
0
BTLP1
TFC5
BMYBTF2
YCAG_57a
YCAA_66b
YCAG_77
BCAC_175
GCAG_184
GCAG_163
BCLV2
RCAG_194b
BMYBTF3
GCAA_390
RCAC_238a
BPP3
BDFR
BTF19
YCAA_102
BTF16
BPP4
GCAC_136
BERF1
BTF24
🟩 Saima_peltate_proportion
-·- LOD, Composite IM (LS)

Figure CXVII CIM for BOB 360 for distance from petiole attachment point to leaf boundary, petiole attachment point to centre, leaf length, leaf width and petiole length. The graph indicates that none of these traits are controlled by *BARP1* locus or locus controlling these traits neither is very close to it on chromosome 13.



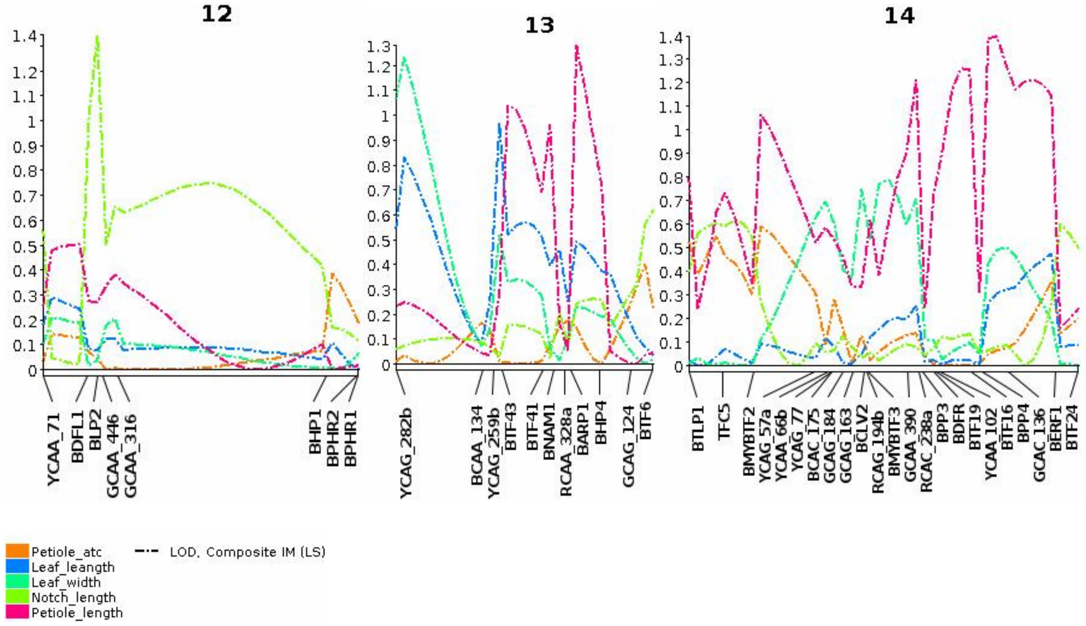
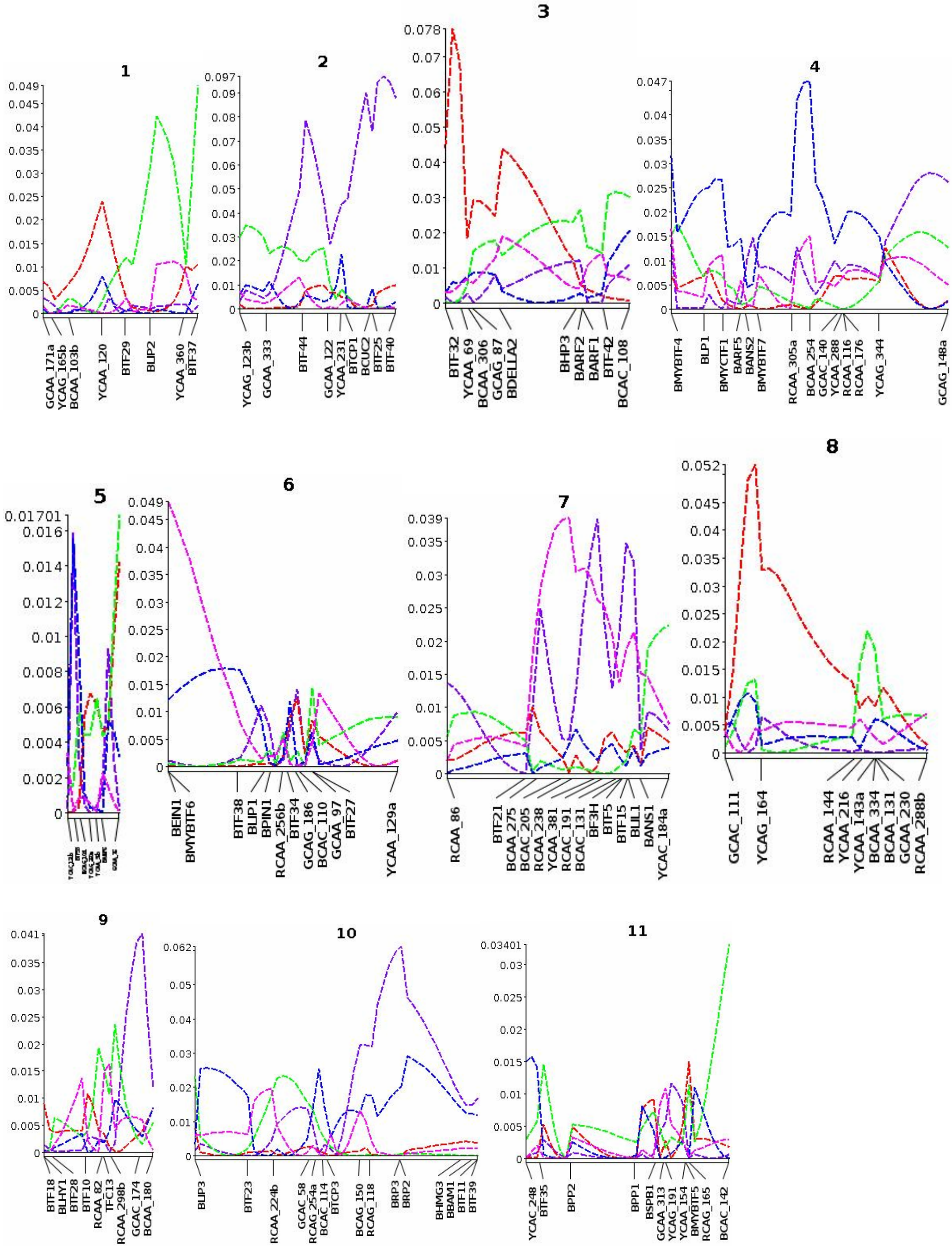

12
13
14
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
YCAA_71
BDFL1
BLP2
GCAA_446
GCAA_316
BHP1
BPHR2
BPHR1
YCAG_282b
BCAA_134
YCAG_259b
BTF43
BTF41
BNAM1
RCAA_328a
BARP1
BHP4
GCAG_124
BTF6
BTLP1
TFC5
BMYBTF2
YCAG_57a
YCAA_66b
YCAG_77
BCAC_175
GCAG_184
GCAG_163
BCLV2
RCAG_194b
BMYBTF3
GCAA_390
RCAC_238a
BPP3
BDFR
BTF19
YCAA_102
BTF16
BPP4
GCAC_136
BERF1
BTF24
Petiole_atc
Leaf_leangth
Leaf_width
Notch_length
Petiole_length
LOD, Composite IM (LS)

**Figure CXVIII CIM for BOB 360 for Elliptic Fourier PCAs.
The graph indicates that none of Elliptic Fourier PCAs is controlled by *BARP1* locus.**



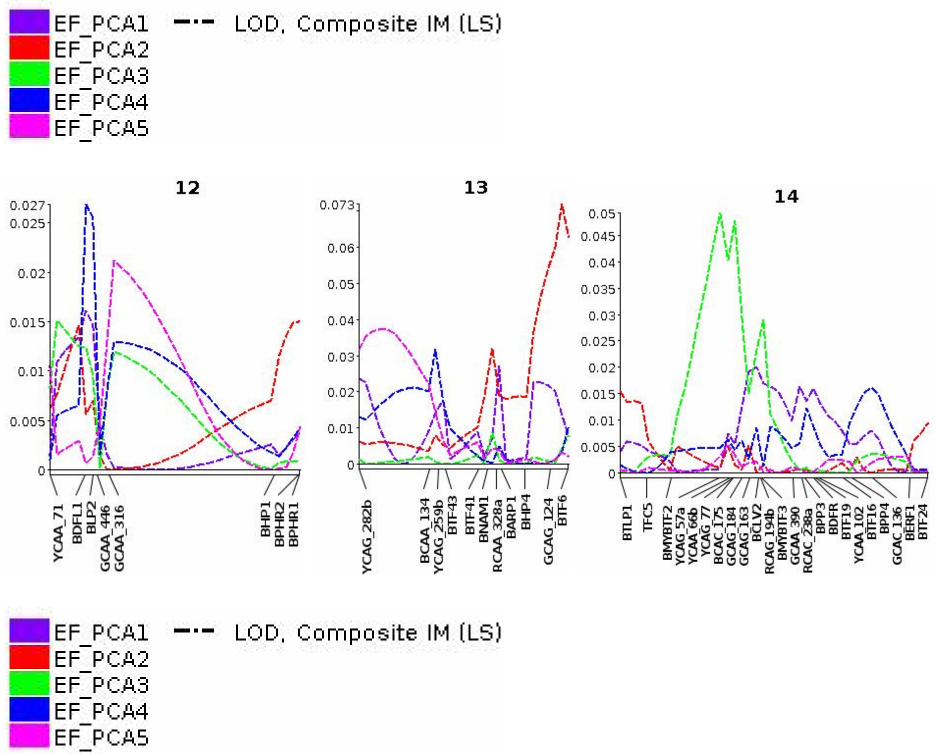
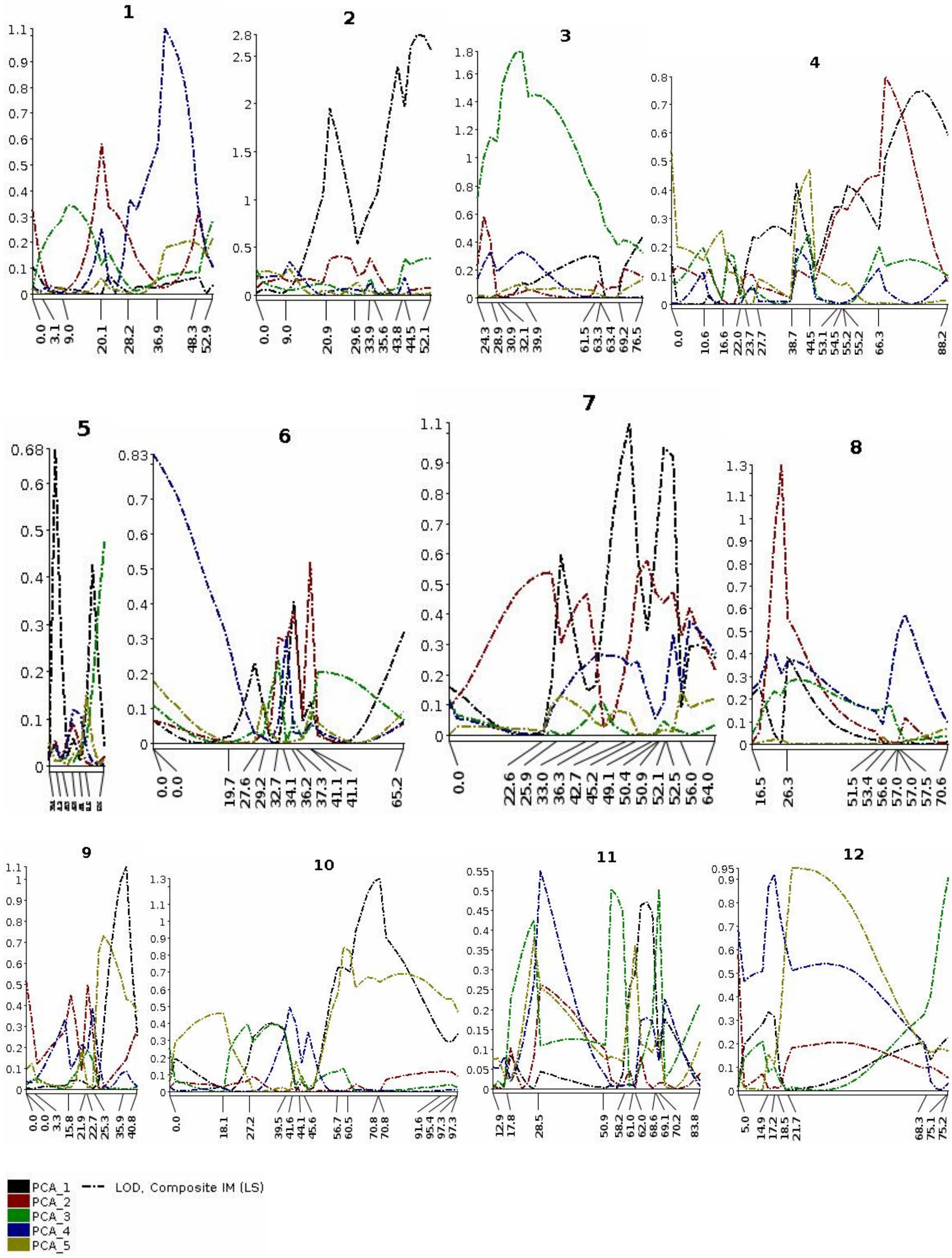
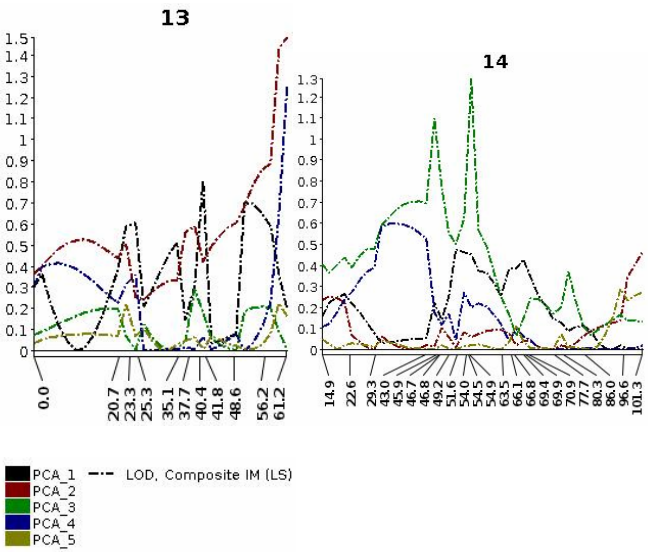

EF_PCA1
EF_PCA2
EF_PCA3
EF_PCA4
EF_PCA5
LOD, Composite IM (LS)
12
0.027
0.025
0.02
0.015
0.01
0.005
0
YCAA_71
BDFL1
BLP2
GCAA_446
GCAA_316
BHP1
BPHR2
BPHR1
13
0.073
0.06
0.05
0.04
0.03
0.02
0.01
0
YCAG_282b
BCAA_134
YCAG_259b
BTF43
BTF41
BNAM1
RCAA_328a
BARP1
BHP4
GCAG_124
BTF6
14
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
BTLP1
TFC5
BMYBTF2
YCAG_57a
YCAA_66b
YCAG_77
BCAC_175
GCAG_184
GCAG_163
BCLV2
RCAG_194b
BMYBTF3
GCAA_390
RCAC_238a
BPP3
BDFR
BTF19
YCAA_102
BTF16
BPP4
GCAC_136
BERF1
BTF24
EF_PCA1
EF_PCA2
EF_PCA3
EF_PCA4
EF_PCA5
LOD, Composite IM (LS)

Figure CXIX CIM for BOB 360 for PCAs.
The graph indicates that none of the shape PCAs are controlled by the *BARP1* locus.




13
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.0
20.7
23.3
25.3
35.1
37.7
40.4
41.8
48.6
56.2
61.2
14
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
14.9
22.6
29.3
43.0
45.9
46.7
46.8
49.2
51.6
54.0
54.5
54.9
63.5
66.1
66.8
69.4
69.9
70.9
77.7
80.3
86.0
96.6
101.3
PCA_1
PCA_2
PCA_3
PCA_4
PCA_5
LOD, Composite IM (LS)

Appendix C

List of Solution

C.1 Stock Solutions for DNA Southern blot analyses

1. Denaturation solution

NaOH	20g
NaCl	58.4
DdH2O to	1L

2. Neutralization solution

NaCl	175.4
Tris Base	60.6
ddH_2O	~0.9L
pH with HCl to	~7.4
ddH_2O to final volume	1L

3. 20X SSC

NaCl	175.25
Na-Citrate	88.2

dH_2O to just under	1L
pH to	~7.0
ddH_2O to final volume	1L

4. Formamide

Formamide stock solution is deionised by stirring with AG 501-X8(D) ion exchange resin (BIORAD) for 30 min and subsequently filtering/decanting. USE CAUTION when handling formamide. Absolutely wear gloves. Formamide is toxic. It is targeted to the reproductive organs.

5. 2M Phosphate Buffer (2M PB)

1)	Na_2HPO_4	28.3g
2)	ddH_2O	160ml
3)	Mix until dissolved(heat ok)	
4)	NaH_2PO_4	27.5g
5)	Mix until thoroughly dissolved	
6)	Bring vol to	200ml
7)	F, A/C	

6. 100X Denhard’s

1)	BSA	0.5g
2)	PVP	0.5g
3)	Ficoll	0.5g
4)	ddH_2O to	25.0ml
5)	Filter through cinder-glass Millipore	
6)	Store to -20˚C	

7. 10% SDS

1)	SDS	10g
2)	ddH_2O	100ml
3)	Sterile filter	

8. 50X Dextran sulphate

1)	Dextran sulphate	10g
2)	Sterile H_2O to	100ml
3)	Heat/Vortex to dissolve	
4)	Store -20˚C	

9. Fish (Herring Sperm DNA)

1) Dissolve 100mg of Herring sperm DNA in 10 ml ddH2O. A combination of vortexing, heat and time are required to get the DNA to completely dissolve.
2) Shear the DNA by vigorously passing it through a small gauge needle fitted onto the appropriate sized syringe. Do this 20 or more times.
3) Add an equal volume TE equilibrated phenol.
4) Vortex and spin.
5) Transfer the aqueous layer to a new tube.
6) Add an equal volume of TE equilibrated phenol- Sevag (50:50).
7) Vortex and spin.
8) Transfer the aqueous layer to a new tube.
9) Add an equal volume of sevag.
10) Vortex and spin
11) Transfer the aqueous layer to a new tube.
12) Add 1/10 volume 3M NaOAc (pH 5.2-6.0).
13) Add 2X volume EtOH.
14) Invert to mix well and spin for 1-2 min.
15) Decant off supernatant.
16) Wash pellet with 70% EtOH.
17) Spin, Wash and Dry the DNA pellet.
18) Resuspend DNA in T_{10} $E_{0.1}$ pH 7.5 to ~20mg/ml.
19) Boil 10 minutes to denature.
20) Place on ice; it is now ready for use. Store at -20˚C.

(Alternatively boil just before use.)

10. 5% PPi

1)	Sodium pyrophosphate	7.5g
2)	ddH2O to	150ml
3)	F, A/C	

11. Hybridization solutions

Stock Solutions	Hyb (~0.05 ml/cm^2)
Formamide	2.5ml (50%)
SSC(25X)	1.0ml(5X)
Na-pH 6.8(2M)	0.05ml(20mM)
Denhardt's (100X)	0.05ml(1X)
SDS (10%)	0.05ml(0.1%)
Dextran sulphate (50%)	1.0ml(10%)
ssDNA (10mg/ml)	0.05ml(100ug/ml)
dH_2O	

C.2 Stock Solutions for In situ hybridization

1. BM blocking solution (1X)

Reagent	Amount to add (for 200 mL)	Final concentration
Blocking buffer for in situ hybridization (BBISH) (10X)	20 mL	1X
Maleic acid buffer	180 mL	90%

Melt a frozen aliquot of 10X BBISH at room temperature and mix with maleic acid buffer by swirling. Store block solution at room temperature until needed. Prepare fresh before use.

2. Block solution 2

Reagent	Amount to add (for 400 mL)	Final concentration
Bovine serum albumin (BSA) (minimum 98%)	4 g	1%
5 M NaCl	12 mL	3%
Triton X-100	1.2 mL	0.3%
1 M Tris-Cl (pH 7.5)	40 mL	10%
H2O (RNase-free)	346 mL	86.5%

Dissolve BSA in H2O by stirring vigorously at room temperature. Add other reagents and mix by stirring at room temperature. Store block solution at room temperature until needed. Solution may be kept for 1 d at 4°C.

3. Buffer C

Reagent	Amount to add (for 200 mL)	Final concentration
5 M NaCl	4 mL	2%
1 M Tris (pH 9.7)	20 mL	10%
1 M MgCl2	10 mL	5%
H2O 166 mL		83%

Mix reagents by stirring at room temperature. Store solution at room temperature until needed. Prepare fresh before use.

4. Hybridization solution (12 slides)

Reagent	Quantity (for 12 slides)	Final concentration
In situ salt solution (10X)	240 µL	12.5%
Formamide (100%) (RNase-free)	960 µL	50%
Dextran sulfate (MW 500) (50% in H2O)	480 µL	25%
Denhardt's reagent (50X)	48 µL	2.5%
tRNA (Sigma) (10 mg/mL in H2O)	216 µ	11.25%

Prepare all reagent stock solutions in RNase-free H2O and store at –20°C. Incubate dextran sulfate solution at 85°C to reduce viscosity before pipetting. Mix all other reagents in a 2-mL RNase-free microcentrifuge tube and incubate at 85°C until needed; do not allow solution to cool down or it may solidify. Prepare fresh before use. The calculations provided are for a total of 12 slides; calculations can be easily adjusted for a different number of slides.

5. LB (Luria-Bertani) liquid medium

Reagent	Amount to add
$H2O$	950 mL
Tryptone	10 g
NaCl	10 g
Yeast extract	5 g

Combine the reagents and shake until the solutes have dissolved. Adjust the pH to 7.0 with 5 N NaOH (~0.2 mL). Adjust the final volume of the solution to 1 L with H2O. Sterilize by autoclaving for 20 min at 15 psi (1.05 kg/cm2) on liquid cycle.

6. NTE

Reagent	Amount to add (for 500 mL)	Final concentration
5 M NaCl (pH 8.0)	50 mL	10%
1 M Tris-Cl (pH 7.5)	5 mL	1%
0.5 M EDTA	1 mL	0.2%
RNAse-free H2O	444 mL	88.8%

All reagents should be RNase-free. Dissolve reagents at room temperature. Store solution at room temperature until needed. Prepare fresh before use.

7. Phosphate-buffered saline (PBS)

Reagent	Amount to add (for 1X solution)	Final concentration (1X)	Amount to add (for 10X stock)	Final concentration (10X)
NaCl	8 g	137 mM	80 g	1.37 M

KCl	0.2 g	2.7 mM	2 g	27 mM
Na2HPO4	1.44 g	10 mM	14.4 g	100 mM
KH2PO4	0.24 g	1.8 mM	2.4 g	18 mM

PBS can be made as a 1X solution or as a 10X stock. To prepare 1 L of either 1X or 10X PBS, dissolve the reagents listed above in 800 mL of H2O. Adjust the pH to 7.4 (or 7.2, if required) with HCl, and then add H2O to 1 L. Dispense the solution into aliquots and sterilize them by autoclaving for 20 min at 15 psi (1.05 kg/cm2) on liquid cycle or by filter sterilization. Store PBS at room temperature.

8. Proteinase K buffer (10X)

Reagent	Amount 25 mL to add (for	Final concentration
1 M Tris-Cl (pH 8.0)	25 mL	0.1 M
0.5 M EDTA (pH 8.0)	25 mL	0.05 M
H2O	200 mL	

All reagents should be RNase-free. Mix reagents and store buffer at room temperature.

9. TE buffer

Reagent	Quantity (for 100 mL)	Final concentration
EDTA (0.5 M, pH 8.0)	0.2 mL	1 mM
Tris-Cl (1 M, pH 8.0)	1 m	10 mM

H2O to 100 mL

10. Tissue fixing buffer

Reagent	Amount to add (for 40 mL)	Final concentration

Paraformaldehyde (16%) (Electron Microscopy Sciences)	10 mL	25%
DMSO	0.4 mL	1%
Phosphate-buffered saline (PBS)	4 mL	10%
H2O	25.6 mL	64%

Mix reagents in a chemical fume hood in a 50-mL conical tube. Keep on ice at all times. Prepare fresh before use.

Table of Contents

List of Figures

List of Tables