A Molecular Analysis of Flower Colour Development in an Ornamental Monocot (*Anthurium andraeanum*)

A thesis presented in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy
in
Plant Molecular Biology

at Massey University, Palmerston North,
New Zealand.

Vern Eddy Collette

2002

ABSTRACT

Colour in *Anthurium andraeanum* spathe and spadix was investigated at the molecular level. A cDNA library was constructed from poly $(A)^+$ RNA isolated from different stages of spathe tissue of the red-flowered anthurium cultivar, Altar. Full-length clones for the flavonoid biosynthetic genes, chalcone synthase, flavanone 3-hydroxylase, dihydroflavonol 4-reductase (*DFR*) and anthocyanidin synthase were isolated by heterologous screening. The expression pattern of these genes implicates *DFR* as a prime regulatory target in the spathe, having an independent regulatory mechanism to that of the other three genes. In the spadix, other regulatory targets are suggested. Additional analysis of *DFR* expression in the spathe revealed a diurnal rhythm to its transcript profile and a model of the possible functional significance of this is presented.

Molecular analysis of the genetic model for anthurium spathe colour was performed with three genotypically defined white lines recessive at the *O* and *M* loci, revealing a more complex genetic model than that originally proposed. The hypothesis that the *O* locus encodes a regulatory protein with specific targets is discussed along with various possible identities for *M*.

Several partial *Myb* cDNA clones were isolated, representing six distinct Myb groups in the anthurium spathe. A full-length cDNA clone for one *Myb* gene, *AaMyb1,* was obtained. *AaMyb1* encodes a R2R3 Myb protein. It had all the structural features in its DNA binding domain that are conserved in R2R3 Myb proteins as well as an acidic domain in the C-terminus that is a potential activation domain. In sequence comparisons with other Myb proteins, AaMYB1 had high similarity to anthocyanin related Mybs from *Zea mays* (maize). However, in transient assays, AaMYB1 was unable to restore wild type phenotype in an *Antirrhinum majus* line, mutated at the anthocyanin Myb locus *Rosea1*. The expression pattern of AaMYB1, in fact, suggests a role in regulating flavone production in the anthurium spathe.

Analyses were done to further investigate the regulation of the anthurium *DFR* promoter. Specific conserved *cis*-elements recognised by anthocyanin Myb regulators were found in the promoter fragment. However, transient expression assays showed that the anthurium *DFR* promoter was activated independently of ROSEA1. The possibility that *DFR* expression is controlled by several regulatory mechanisms, involving various signal transduction cascades, is discussed.

ACKNOWLEDGEMENTS

This acknowledgement is to all who in one way or another have assisted me in completing my study. Firstly, to the government of New Zealand, I am deeply grateful for offering me a Commonwealth scholarship to pursue my doctorate in this wonderfully scenic country. To Kiri Manuera, Margaret Gilbert and your respective teams, I appreciate all the kindness, assistance, interest and support shown to me. I am also thankful for the extension of my scholarship through Massey University, enabling me to produce an excellent thesis.

To my supervisors, Professor Paula Jameson, at Massey University, Dr. Kevin Davies and Dr. Kathy Schwinn of Crop & Food Research (C&FR), it was an honour to be guided by your intellect and proficiency. Many thanks for the mentoring and training provided which I am sure will contribute positively to my development as a scientist. I have always commented to my peers that I was assigned the best combination of supervisors a student could ask for. I hold each one of you in very high esteem and desire for you and your loved ones the very best for the future. My thanks also to Paula and family for allowing me to stay at their home when I first arrived in New Zealand.

To C&FR, I express my sincere thanks for funding my research and to all the staff for providing a most enjoyable work culture. To Dr. Chris Winefield, Steve Arathoon and Jan Mason, many thanks for the countless assistance in the lab and for the friendship and laughter that often characterised our interactions. To the members of the C&FR business house tennis team, thanks for wonderful times of recreation, hopefully you can secure first place in a subsequent tournament. MaryAnne, in the mean time, we would guard zealously our personal title of 'unbeaten mix doubles team over two seasons.'

Many thanks, to the staff at the Massey Plant Growth Unit for the excellent care of my plants, during the thesis. I understood the odds were against us in getting the plants established and thriving but clearly your expertise was more than sufficient.

I can write a thesis of equivalent length describing in detail the contributions made by my wife throughout the duration of this study. Words cannot begin to describe the great debt of love and gratitude I have for my wife and best friend, Lauren. She has joyfully given up the last four years of her life on my behalf. It is her love, encouragement, inspiration, intuition, enthusiasm and prayer that has been integral to my success. Her interest in my work has never diminished and her ability to grasp the concepts of molecular biology, though her background is business studies, is a testimony to the wonderful mind she possesses. I am grateful to her for the detailed corrections of my work. I dedicate this thesis to you my darling and pledge my support in your PhD endeavours.

To all the members of my family and friends both at home in Trinidad and Tobago and around the world. Thanks for the continued support and prayer. This success belongs to us all. Special commendation to my mother, Frances Collette ('the tallest lady'). We both know what this means to us. You are a pillar of strength, a symbol of honour and virtue. Thanks for being my mother, the best in the world.

My appreciation also goes to Dr. Umaharan of the University of the West Indies, St. Augustine, who nurtured my love for genetics and molecular biology. He continues to be an inspiration and a model for me. Also to Kairi Blooms, Trinidad, for their continued support of anthurium research. Such links between industry and university are critical to continued development. I am confident that the results of this work would contribute positively to your company.

Finally, I acknowledge my friend and my God, the ultimate designer of the wonderfully complex biological systems that we explore today through science, and Jesus Christ my Lord, for the multifaceted grace that was given to me at every stage of my thesis. My heart is filled with gratitude and praise to Him.

TABLE OF CONTENTS

ABBREVIATIONS

Chemicals

ANS	Anthocyanin synthase
BSA	bovine serum albumin
CH_3COOH	acetic acid
CHS	Chalcone synthase
CTAB	Cetyltrimethylammonium bromide
DFR	Dihydroflavonol 4-reductase
DMSO	dimethyl sulphoxide
DTT	dithiothreitol
EtBr	ethidium bromide
F3H	Flavanone 3-hydroxylase
IPTG	isopropyl-β-D-thiogalactoside
KAc	potassium acetate
L-Broth	Luria Broth
Liquid N_2	liquid nitrogen
2β ME	2β mercaptoethanol
$MnCL_2$	manganese chloride
PEG	polyethylene glycol
PMSF	phenyl methyl sulfonyl fluoride
PVP-40	polyvinylpyrrilidone
PVPP	polyvinylpolypyrrolidone
SSC	standard saline citrate
TBE	tris borate EDTA buffer
TE	tris EDTA buffer
TLC	thin layer chromatography
X-Gal	5'-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
X-Gluc	5'-bromo-4-chloro-3-indoyl-β-D- glucuronide

Terms/Techniques

CaMV	cauliflower mosaic virus
GUS	β-glucuronidase
GFP	green fluorescent protein
h	hour
min	minute
PLACE	plant cis-acting regulatory elements
rpm	revolutions per minute
s	second
TRANSFAC	transcription factor database
v/v	volume/volume
w/v	weight/volume

LIST OF FIGURES

On page following

LIST OF TABLES

Chapter 1

Introduction

1.1 COMMERCIAL ASPECTS OF ANTHURIUM PRODUCTION IN TRINIDAD AND TOBAGO

Anthurium andraeanum Linden ex André (anthurium) originated in Columbia and is a slow-growing perennial that requires shady, humid conditions such as found in tropical rainforests. It is a member of the family Araceae that also includes philodendron, monstera, and calla lily. It is characterised by its brilliantly coloured, heart shaped spathe (bract) and for this reason it is grown as an ornamental crop around the world. The traded value of anthurium is second only to that of spray tropical orchids among the tropical flowers, and the world import market size for anthurium is estimated to exceed US$20 million annually (World Market for Anthurium, Asia Regional Agribusiness Project (RAP) Market Information, 1997).

Today, Trinidad and Tobago is the largest foreign-based suppliers of anthurium to the United States of America, (USA). Currently, there are some ten hectares under intensive cultivation, producing about 2 million stems annually, generating US $750,000 as foreign capital and having estimated profits as high as 20-30% (Chris Avey, Kairi Blooms, Trinidad and Tobago, West Indies, personal communication). Clearly this industry has the potential to be a significant foreign exchange earner and a major employer within this local economy.

Prior to 1985, traditional cultivation of anthurium in Trinidad and Tobago involved using the natural canopy of cocoa to provide shade for the cut flowers. However, in the mid-eighties, the local floriculture industry began to change with the injection of large sums of capital. Modern anthurium shade houses were introduced enabling growers to produce a much higher quality of flower, with a longer post-harvest life than was previously possible. More importantly, newly imported varieties offered new and exciting colours and much higher production rates (Chris Avey, personal communication). These developments have resulted in a steady increase in acreage and export volume from 1985-1995.

Despite this tremendous growth potential, the industry is now facing a slump in production. From 1995 to the present, export levels have fallen by approximately 30% and are today

estimated to be less than half of production capacity (Chris Avey, personal communication). The low production levels are attributed to the unsuitability of the imported varieties to the harsh tropical conditions as well as their inferior bloom characteristics and their susceptibility to bacterial diseases such as anthurium bacterial wilt.

The local anthurium industry in Trinidad and Tobago faces yet another challenge. The international market demands different colours throughout the year. To effectively respond to this seasonal demand for anthurium blooms, growers must have a high degree of flexibility and the industry needs a wide range of cultivars. More importantly, the international market thrives on novelties. Novel colours fetch premium prices, often twice as high as conventional colours such as reds (Chris Avey, personal communication). The world's largest producer, the Netherlands, maintains their competitive advantage on the international market because they are the world's largest breeders of anthurium and, therefore, the source of most new colours and varieties.

The particular needs of the local industry in Trinidad and Tobago are to have a large range of colours, from plants that are highly productive, disease resistant, and moderately adaptable to the conditions prevalent in Trinidad and Tobago. This requires technology that can modify colour without bringing changes to those factors developed by breeders through a long and labourious breeding process. Flower colour modification by genetic engineering provides a unique way of modifying colour in an already adaptable background. Concurrently with this thesis, an efficient transformation procedure has been developed for anthurium (Chen and Kuehnle, 1996; Dr. Pathmanathan Umaharan, The University of the West Indies, St. Augustine, Trinidad and Tobago, personal communication). Therefore, the gene identification and transformation protocols, together, provide a solid scientific platform for developing an even wider range of colours for the anthurium growers than is presently obtainable, without the risk of affecting productivity.

This study is the first molecular investigation of flower colour in anthurium and, as such, hopes to lay the groundwork for future studies that have economic significance while making academic contributions to our understanding of flower colour development.

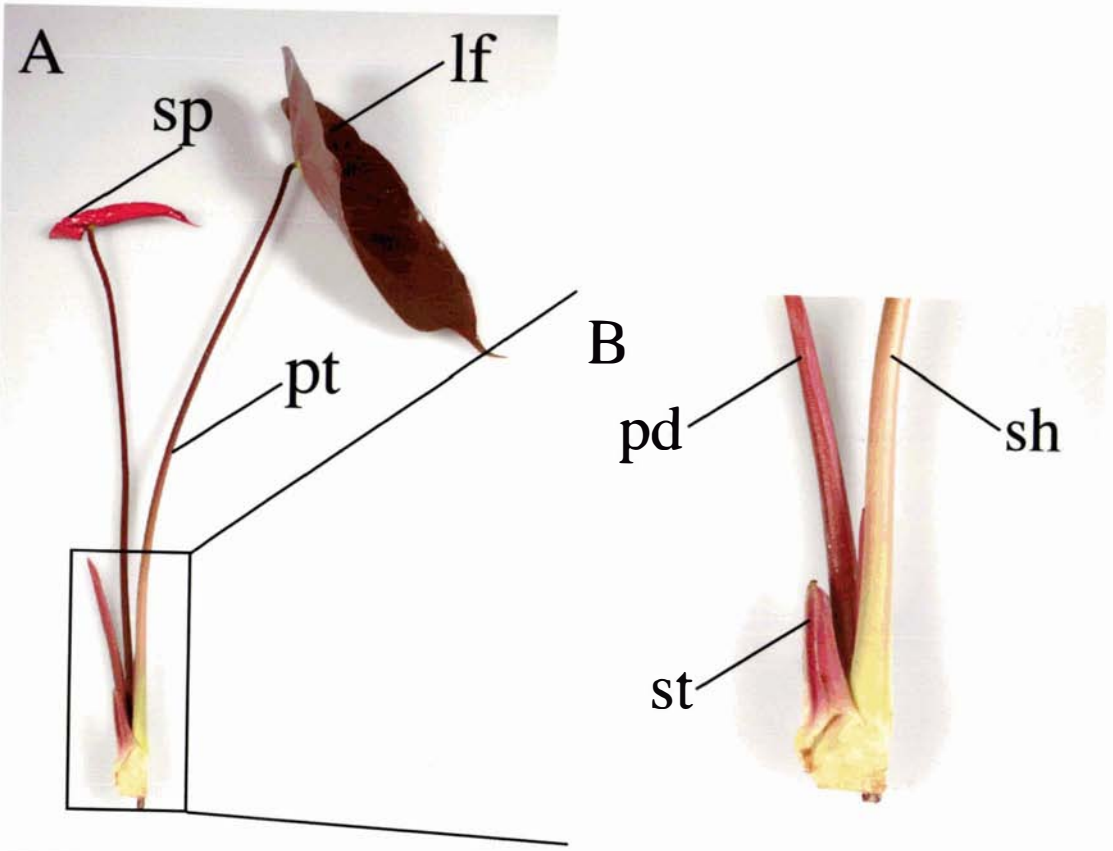
1.2 THE ANTHURIUM FLOWER

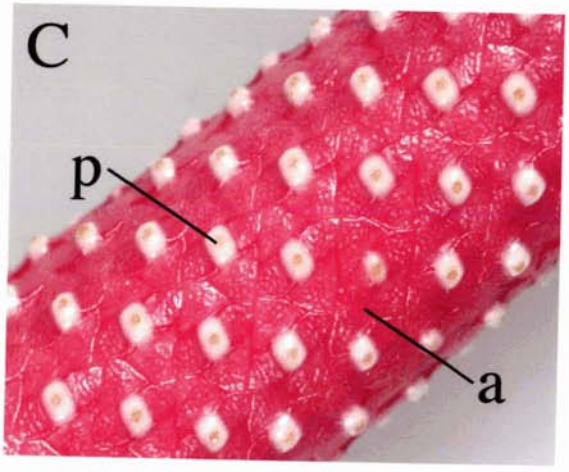
It is hoped that the results from this study will assist future efforts of genetic manipulation to modify flower colour in anthurium. However, the plant taxonomy and the morphological features of the flower along with the genetic and biochemical features of pigment production in flower tissue all together constitute an interesting plant system for basic molecular research of anthocyanin gene expression and regulation.

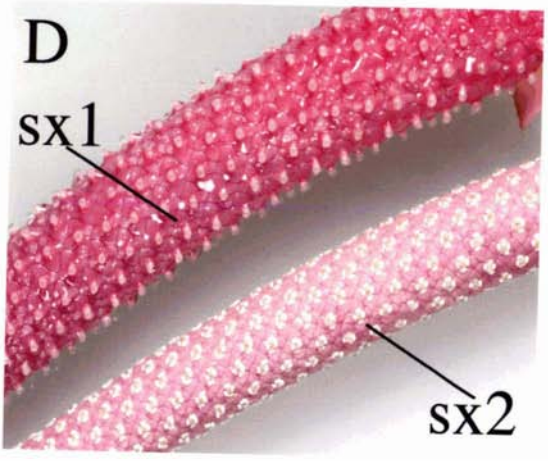
The commercial flower consists of a modified bract called the spathe that subtends a cylindrical protruding inflorescence rachis called the spadix (Higaki et al., 1984). Based on the state of the axillary bud, the growth of the plant has been divided into a juvenile phase that is followed by a generative phase. In the juvenile phase, a vegetative bud emerges in the axil of leaves, while in the generative phase, an inflorescence bud can be found in the leaf axils (Christensen, 1971). Once the juvenile phase has passed, flowers are produced in a cycle of leaf, flower, leaf, flower, throughout the year (Kamemoto and Nakasone, 1963). The term 'flower' is used throughout this thesis to refer to the commercial product, made up of the spadix, spathe and peduncle. The surface of the spadix is made up of fused segments, each being the point from which the true flower emerges (Figure 1.1). These are minute hermaphroditic flowers that are protogynous in development making anthurium highly outcrossing (Campbell, 1900; Kamemoto and Nakasone, 1963; Higaki et al., 1984). These true flowers are invisible to the naked eye and are insect pollinated. Once fertilised the ovary enlarges and forms a single berry with one or two seeds inside, in a process that completely disfigures the spadix (Figure 1.1). Once mature these detach from the spadix.

When the flower emerges in the leaf axil, the spathe is tightly curled around the immature spadix. It takes an average of six to nine weeks from the time of its emergence, for the peduncle to be fully extended and the spathe fully expanded, exposing a mature spadix (refer to Figure 2.1). The spathe is cordate and simple, although for some anthurium species it may be compound and also assume other shapes. Very prominent veins that originate at the spadix-spathe junction run along the surface of the spathe in an arcuate manner and converge at the apex of the spathe (examples of this can be seen in the flowers shown in

Figure 1.1. Various aspects of anthurium morphology and pigment distribution. Panel A and B shows the distribution of anthocyanin in various tissues apart from the spathe (sp). Pigment accumulates in the leaf (lf), petiole (pt), peduncle (pd), sheath (sh), and stipule (st). Panel C shows the segmented spadix and the location of pistil (p) and anther (a). In Panel D, the true flowers have reached reproductive maturity. Anthurium flowers are protogynous. Spadix 1 (sx1) shows the mature, sticky, stigmatic surface while the anthers, which mature much later than the stigma, are shown in spadix 2 (sx2). Panel E shows a developing ovule (ov), indicating successful pollination. As the ovule develops, it protrudes from the surface of the spadix.







E

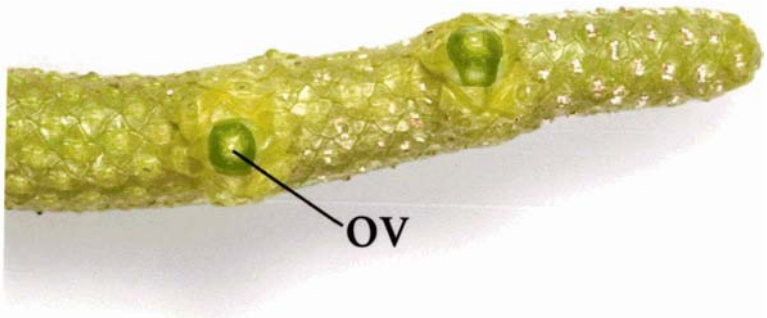


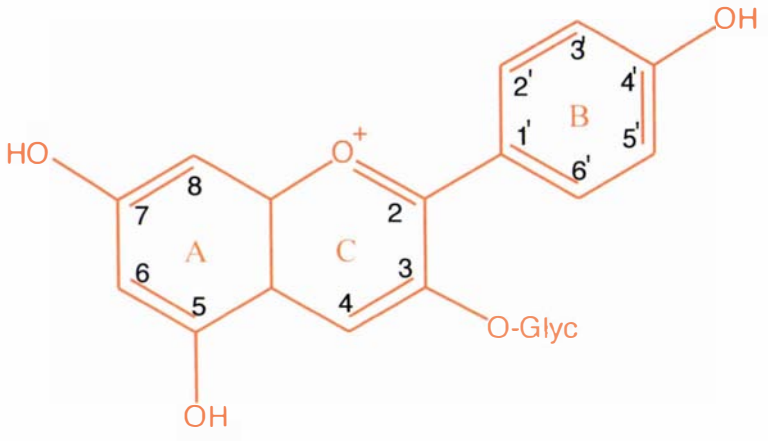
Figure 1.3) (Higaki et al., 1984). The spathe is also characterised by a thick cuticle on the upper and lower epidermis. The epidermis (both upper and lower) is one cell layer thick and beneath this is a single or double layer of isodiametric hypodermal cells. In anthurium, flower colour pigment accumulates only in the hypodermal cells of both the upper and lower epidermis.

1.3 FLOWER COLOUR PIGMENTS

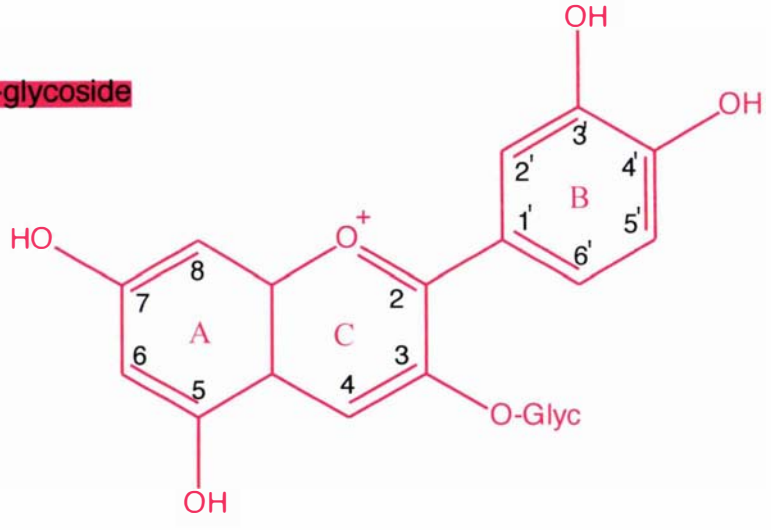
Flower colour was among the first plant traits chosen for inheritance studies, and early work on flower colour in *Antirrhinum majus* (antirrhinum) provided the first significant link between genetics and biochemical changes (Whedale, 1914; 1915). The main colour pigments have been well characterised and fall into four categories. Carotenoids, found in a large number of plant species, are hydrophobic, lipid soluble pigments derived from the isoprenoid pathway. They are located in plastids of the chromoplast or chloroplast type. Those in the chromoplast serve primarily as colour pigments and impart yellow, red, orange or purple colouration to tissues. Those in the chloroplast serve to protect the photosynthetic machinery from photo-oxidation. Chlorophylls, though serving a primary role in photosynthesis, can contribute to the colour of some flowers and fruit in addition to vegetative tissue. Betalains are the only colour pigments that are taxonomically restricted being found only in the order *Caryophyllales* (Clement et al., 1994). Betalains are synthesised from the amino acid tyrosine into the two subclasses, betacyanins (reds and purples) and betaxanthins (yellows and orange). The fourth class of pigment and the most widespread are the flavonoids. They comprise the largest and most diverse group of plant pigments. Flavonoids are hydrophilic with a basic structure consisting of two phenyl rings (referred to as the A- and B-rings) connected by a three-carbon unit that is closed to form the C-ring in all the flavonoids except the chalcones (refer to Figure 1.2). Flavonoids are derivatives of the phenylpropanoid pathway and contribute orange, coral, pink, red, purple, blue and blue-black colours to flowers and other types of plant tissue. Some flavonoids can contribute to yellow colours.

Figure 1.2. The three anthocyanin derivatives of the phenylpropanoid pathway. The three carbon rings (A, B and C) that comprise the basic ring structure of flavonoids is shown along with the numbering of some of the carbons. The molecules shown are representatives of the three classes of anthocyanin types. They differ only in the degree of hydroxylation for the B-ring.

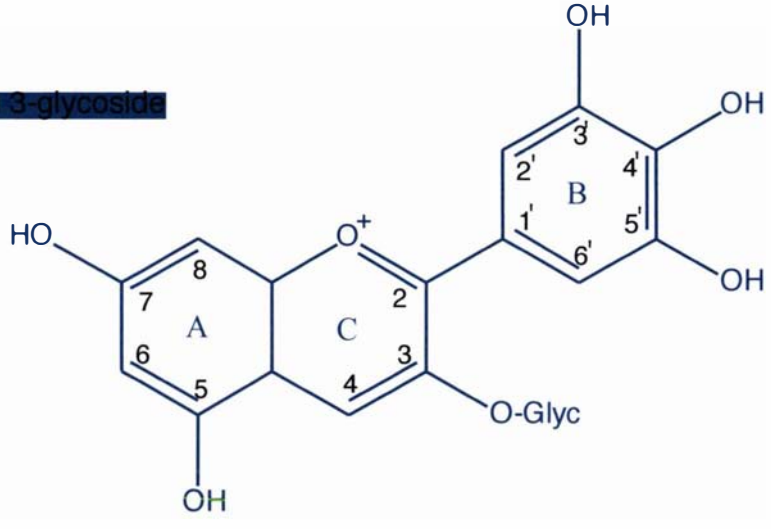
Pelargonidin 3-glycoside



Cyanidin 3-glycoside



Delphinidin 3-glycoside



1.4 THE FLAVONOIDS

The earliest recognition of existence of the compounds, which are known today as flavonoids, was in the seventeenth (17^{th}) century (Grew, 1682). Today, several book volumes and reviews are dedicated solely to flavonoid research reflecting the extensive body of knowledge and vast numbers of this type of plant pigment that have been described over time. Arguably the most important member of this group are the anthocyanins as they make a significant contribution to flower colour providing the basis for most of the orange, pink, red, magenta, purple and blue flower colours. Apart from the anthocyanins, aurones and chalcones are also coloured flavonoids giving rise to yellow and orange colour in some species (Kuhn et al., 1978; Bohm, 1993). Not all the flavonoids absorb light in the visible range. Some, such as flavanones, flavones, flavonols, dihydroflavonols and leucoanthocyanidins are colourless and, with the exception of flavones and flavonols, are usually transient. However, the flavanones, flavones and flavonols can contribute 'depth' to white and cream flowers and exert additional effects on final colour through copigmentation. While the colour function of anthocyanin is easily appreciated, flavonoids are known to have a multiplicity of additional functions in plant tissue. There is growing evidence for the role of flavonoids as UV-protectants. Flavonoid compounds tend to be localised in the epidermal layers of plants, a position that is logically consistent with them serving as shields to potentially harmful ultraviolet radiation (Cadwell et al., 1983). Flavonoids are also involved in pollen germination and pollen tube growth (Coe et al., 1981; Mo et al., 1992), and act as biochemical messengers facilitating plant microbe symbiosis (Peters et al., 1986; Phillips and Tsai, 1992; Pueppke, 1996), or as antimicrobial agents with either a phytoalexin or an allelopathic response (Mori et al., 1987; Dixon and Lamb, 1990). There is also growing evidence for their pharmaceutical benefits, as antioxidants, bioactives and anticarcinogens (Middleton and Kandawami, 1992; Middleton et al., 2000).

1.5 ANTHOCYANINS

Most anthocyanins are derived from cyanidin, pelargonidin or delphinidin, which differ only in the degree of hydroxylation of the B-ring (Figure 1.2) and generally produce red,

orange and blue colours respectively. To function properly as pigments anthocyanins must accumulate in the vacuole at sufficient concentration, and in a stable coloured state. Glycosylation and copigmentation are two important stabilising modifications that help to maintain the anthocyanin in the coloured state. Glycosylation of the unstable anthocyanidin commonly occurs first at the C-3 position to form the anthocyanin but can continue extensively at other positions on the anthocyanin and even colourless flavonoid molecules providing diversity in addition to stability.

Anthocyanins, like the other flavonoids display a high propensity to form molecular complexes by a process referred to as copigmentation. This is crucial to anthocyanin colour stabilisation and variation. Although anthocyanins are the principal flavonoid flower colour pigments, their chemistry in aqueous solution is such that they are only coloured at fairly acidic pH values. Between pH 3 to 7, the approximate pH range in the vacuoles of floral cell sap, the molecule becomes hydrated resulting in the formation of the colourless pseudobase (Harborne, 1988). Under normal conditions, therefore, the molecule must be protected from hydration to serve its role as a colour pigment. It is thought that copigmentation protects against hydration thereby stabilising the anthocyanin.

Various forms of copigmentation have been described. The anthocyanin can form a complex with itself, a form of copigmentation referred to as self-association. This has been observed with Di-glucosidic methylated anthocyanins, forming molecular aggregates that protect against hydration (Harborne, 1988; Goto and Kondo, 1991). Anthocyanins can also exhibit intermolecular copigmentation, forming complexes with colourless flavonols and flavones, especially the *C*-glycosides (Asen et al., 1970,1986; Brouillard and Dangles, 1993). Intramolecular copigmentation has been observed in aromatic acylated anthocyanins. Here, the acyl groups can sandwich the anthocyanin nucleus, conferring exceptional resistance to colour loss (Brouillard, 1981; Goto and Kondo, 1991; Bloor, 1996). Copigmentation involving complexes with metal ions can also contribute to pigment stability and final colour. The complexing of aluminium ions to red anthocyanins is key to the bluing of the red sepals of *Hydrangea macrophylla*

(Takeda et al., 1985). It is suggested that metal ion strengthens the interaction between the anthocyanin and the copigment (Brouillard and Dangles, 1993). It is this diversity of copigment complexes and glycosylation events that enables such an incredibly diverse range of flower colour to exist from only a limited number of anthocyanin molecules, at pH ranges where anthocyanins are virtually colourless in vitro.

1.6 ANTHURIUM FLORAL PIGMENTS

The main colours in spathes of anthurium are red/pink, orange/coral and white. A few cases of green, and even brown coloured spathes are also known. A selection of some of the colours is shown in Figure 1.3. Some white lines can develop a pink blush as the flower matures giving a patterned arrangement to the spathe. Blues and yellows are notably absent and while varied shades of red and orange are present, these are not always in a desirable genetic background.

Cultivars with a red spathe usually have a bright yellow spadix, while the colour of the spadix varies from orange to red in cultivars with a pink spathe. In the orange spathe group, spadices are either orange or yellow, whereas in cultivars with a white spathe the spadix can be attractively matched with pink to bright red spadices. Other colours such as purple, bronze and green are also found in the spadix (Kamemoto and Kuehnle, 1996).

In anthurium (and in the Araceae family) the major colour pigments in the spathe are anthocyanin derivatives. Carotenoid pigments are also present but only in spadix tissue and are responsible for the bright yellow spadix (Dr. David Lewis, Crop & Food Research, New Zealand, personal communication). All other spadix colours are flavonoid based. Chlorophyll also contributes to spathe colour, either alone, giving completely green spathe or in combination with other anthocyanin pigments to give a brown spathe (Kamemoto and Kuehnle, 1996).

Pelargonidin and cyanidin-derived anthocyanins are the main colour pigments in anthurium spathe. These are usually glycosylated with glucose and rhamnose forming pelargonidin and cyanidin 3-rutinoside (Iwata et al., 1979, 1985). The presence of one or both pigments,

Figure 1.3. The colour range of anthurium flowers. The first five flowers are lines used in this research. Note the different shades of pink spathe and particularly the varied spadix colours in the pink line. While the example for green spathe (meadow) is showing pigment, (note the pigmented veins) there are cultivars with spathe that are completely green. The presence of some anthocyanin pigment is evident in reddish brown spathe (butterfly).

Altar

Lido

Atlanta

Montana

Acropolis

Meadow

Butterfly

Panther

as well as their different concentrations, determines the spathe colour in anthurium (Iwata et al., 1985). Whilst only pelargonidin glycosides have been found in some cultivars with orange spathe, most orange cultivars have both pelargonidin and cyanidin derived anthocyanins but the ratio of pelargonidin to cyanidin is always less than one (Iwata et al., 1985). In contrast, the red cultivars, which have a higher concentration of cyanidin glycosides than pinks, also have pelargonidin glycosides, but in significantly less quantities to cyanidin glycosides (Iwata et al., 1985). In a quantitative analysis, Iwata et al. (1985) showed that colour expression in the orange group was as intense as in the red group. However, the concentration of the dominant pigment in each colour group was highly disproportionate. Cyanidin 3-rutinoside was found at nearly ten times the level of absorbance compared to pelargonidin 3-rutinoside, per gram fresh weight tissue. This suggests that the quantity of pelargonidin compared to cyanidin exerts a more pronounced effect on colour expression (Iwata et al., 1985).

Flavones are the other main flavonoid pigments in anthurium spathe. In a leaf survey of 142 species from 58 genera of the Araceae family, Williams et al. (1981) found that flavone *C*-glycosides comprised 82% of the major flavonoids. Flavone mono-*C* and di *C*-glycosides along with *O*-glycosyl derivatives were all represented. In a recent study, *C*-glycosides were identified as the dominant flavonoid in four anthurium lines (red, white, orange and purple). Analysis of the NMR spectra revealed the likely structure of 7,4'-dimethyl apigenin 7-*C*-glucoside 2" -*O*- rhamnoside previously reported as embinin (Dr. Stephen Bloor, Industrial Research Ltd, New Zealand, personal communication). These data are consistent with work done by Iwata et al. (1985). They investigated 50 white anthurium and they were all characterised by the absence of anthocyanins and a substantial accumulation of flavones.

In addition to the spathe and spadix, anthocyanin pigments have also been observed in the leaves, peduncle, stipules, roots and stems of anthurium (refer to Figure 1.1 to observe pigmentation in some of these structures). While in petals anthocyanins are nearly always found in epidermal cells, in vegetative tissue they tend to be located sub-epidermally (Kay et al., 1981; Hrazdina, 1982). However, Wannakrairoj and Kamemoto (1990a) found that in anthurium spathe, the location of anthocyanin differed from species to species and could be

located in the epidermis, hypodermis and/or mesophyll. In anthurium cultivars, anthocyanins were found exclusively in the hypodermal layers on both the abaxial and adaxial surfaces of the spathe (Higaki et al., 1984).

In addition to pelargonidin and cyanidin glycosides, peonidin 3-rutinoside was reported in *A. amnicola.* In combination with cyanidin 3-rutinoside, it accounts for the lavender spathe and purple spadix of this species (Marutani et al., 1987). The distribution of anthocyanins, however, in *A. amnicola* was restricted to the lower and upper epidermis.

1.7 THE GENETICS OF FLOWER COLOUR INHERITANCE IN ANTHURIUM

Plant breeding work has provided significant information concerning the inheritance of flower colour in anthurium (Kamemoto et al., 1988). Based on crosses using the five major spathe colour groups (red, orange, pink, coral and white) Kamemoto and Nakasone (1963), reported that white and orange lines bred true, red crossed with red produced red and orange in the ratio 3:1 and red crossed with orange produced red and orange in the ratio 1:1. Based on these ratios, they suggested a multiple allelic system in which red (R^r) was dominant to orange (R^o). This was to be the first hypothesis of flower colour inheritance in anthurium. In later studies, Sheffer and Kamemoto (1977) suggested, that among interspecific hybrids, complementary gene action with at least two different genes was operational. Continued investigation over the years with better control of crosses as well as larger populations has clearly demonstrated that two factors rather than a multiple allelic series are responsible for the five major colour groups. Correlating data from biochemical studies with the results from the breeding efforts, Kamemoto et al. (1988) concluded that two genes, *M* and *O,* were responsible for anthocyanin production in anthurium and that there is recessive epistasis of the *O* locus over the *M* locus. They proposed that gene *M* controls the production of cyanidin 3-rutinoside and that gene *O* controls the production of pelargonidin 3-rutinoside. Therefore, the possible white genotypes are: *MMoo, Mmoo* and *mmoo*. The orange genotypes are *mmOO* and *mmOo* and the genotypes for red are *MMOO, MmOO, MMOo* and *MmOo.*

Quantitatively, *M* and *O* genes in combination produce red to pink spathe colours, *O* produces orange to coral and *oo* produces white. However, the dosages of *M* and *O* affect the range of colours obtained. This is referred to as the dosage effect of *M* and *O* genes (Kamemoto et al., 1988). The dosage effect was investigated further using selected crosses of pink with coral lines, coral with coral as well as pink with white lines, combined with further biochemical work. From this Kamemoto et al. (1988) proposed that the incremental effect of *M* is greater than that of *O* and taking into account the recessive epistasis of the *O* locus over the *M* locus, they suggested the following: *mmoo* = *MMoo* = *Mmoo* < *mmOo* < *mmOO* < *MmOo* < *MmOO* < *MMOo* < *MMOO*. Therefore, in going from pink to red the genotype changes from the double heterozygous condition (*MmOo*) to the double homozygous dominant (*MMOO*).

Although *M* and *O* may be the two major genes affecting spathe colour in anthurium, Kamemoto et al. (1988) suggested several other genes operate to provide the diversity of flower colours. The intergrading forms within the red and orange colour groups suggest the presence of other genes modifying the pathway.

Evidence for the involvement of additional genes, was generated by Wannakrairoj and Kamemoto (1990b) in breeding studies aimed at elucidating the genetic system controlling purple spathe colour in anthurium. They proposed that a recessive allele, *p*, modifies the colour of anthocyanins controlled by the *M* and *O* loci. The purple spathe is produced when *P* is in the homozygous recessive state. This phenotype, however, is only expressed when the *O* locus has at least one dominant allele. Otherwise the spathe will be white. The proposed genotype for purple spathe is *M-O-pp*. If the p locus is dominant, *M-O-* is red, while *mmO-* is orange (Wannakrairoj and Kamemoto, 1990b).

1.8 THE BIOCHEMISTRY OF FLOWER COLOUR AS IT RELATES TO ANTHURIUM

The biochemistry and genetics of the phenylpropanoid pathway has been well documented for numerous species. The relevant data, as they apply to pigment production in the anthurium spathe are presented in the following sections. A diagram

showing portions of the phenylpropanoid pathway relevant to anthurium is represented in Figure 1.4.

1.8.1 Biosynthesis of substrates for the flavonoid pathway

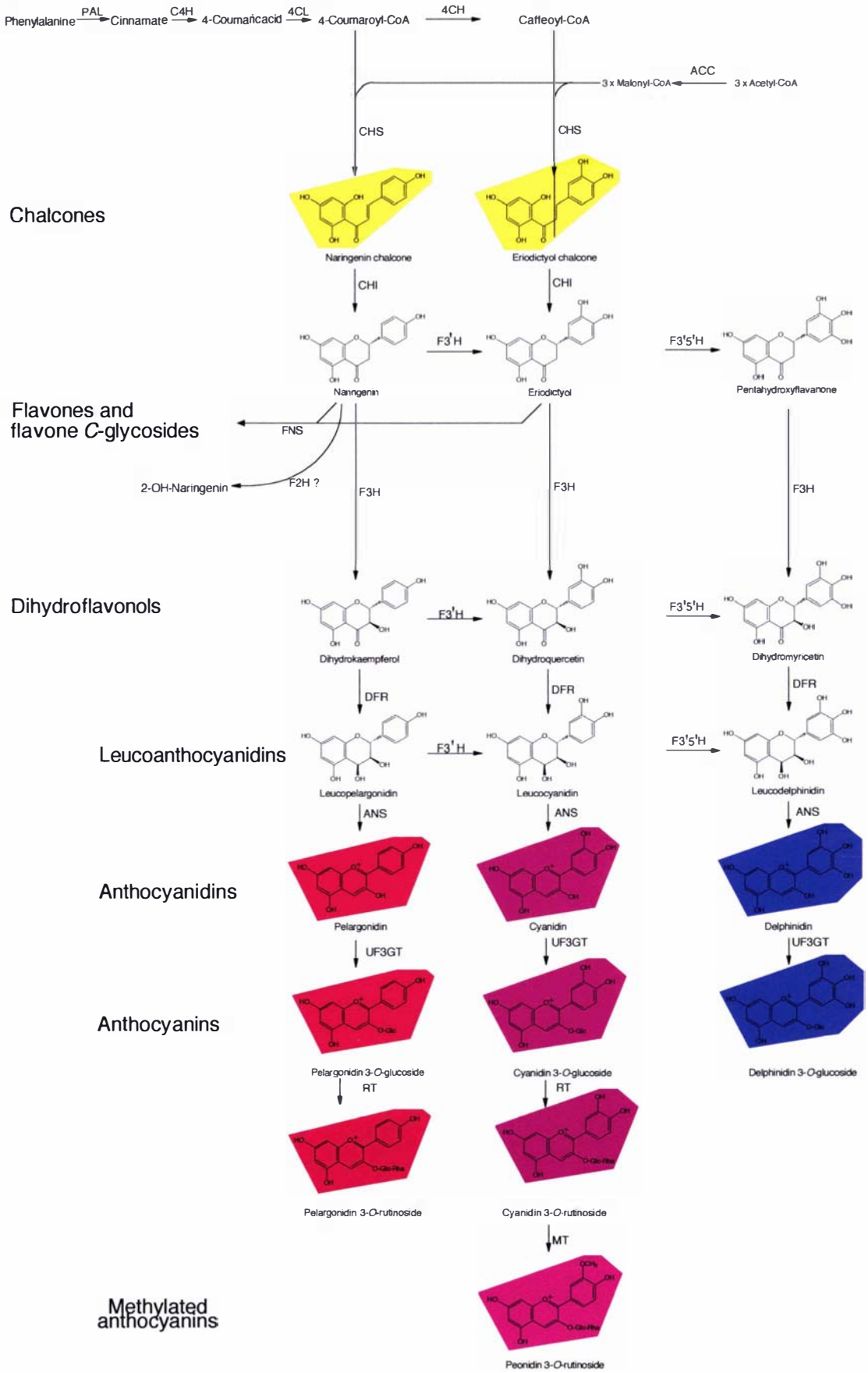
The phenylpropanoid pathway starts with the enzyme phenylalanine ammonia lyase (PAL), which catalyses the *trans*-elimination of ammonia from the amino acid phenylalanine to form *trans*-cinnamate (Hanson and Havir, 1972; Heller and Forkmann, 1993). The enzyme cinnamate 4-hydroxylase (C4H), a cytochrome P450 monooxygenase, transforms the *trans*-cinnamate to 4-coumaric acid (Nair and Vining, 1965; Heller and Forkmann, 1993). 4-coumaric acid is then ligated to CoA by the enzyme 4-coumarate ligase (4CL) to form a 4-coumarate-CoA ester (Hahlbrock and Griesbach, 1975; Heller and Forkmann, 1993). This is the substrate for a range of secondary metabolites including flavonoids.

1.8.2 Biosynthesis of chalcones

The primary flavonoid ring structure is a hybrid of two CoA ester derivative precursors. Malonyl-CoA, which supplies the A ring of the flavonoid, is formed from a reaction with acetyl-CoA and CO_2 catalysed by acetyl-CoA carboxylase (ACC). The B-ring and part of the heterocyclic 3-carbon bridge originates from the 4-coumaroyl-CoA, (Heller and Forkmann, 1993). The first committed step to flavonoid biosynthesis is mediated by the enzyme chalcone synthase (CHS). It catalyses the stepwise condensation of 4-coumaryl-CoA with three acetate units from malonyl-CoA to yield a yellow C_{15} intermediate, naringenin chalcone (Heller and Hahlbrock, 1980).

Although other substrates such as caffeoyl-CoA and feruloyl-CoA can be used by CHS as substrates, 4-coumaroyl-CoA is the most efficiently used and is, in a number of cases, the exclusive substrate. Interestingly, CHS is very closely related to the stilbene synthases (STSs) that makes the backbone of the stilbenes (such as resveratrol). In fact STSs always share the same substrate (4-coumaroyl-CoA). The striking similarities

Figure 1.4. Biochemical pathway for flavonoid production in anthurium. This figure is a modified version of one supplied by Dr. Kevin Davies, Crop & Food Research, New Zealand. It is not known in anthurium at what level (s) hydroxylation of the B-ring occurs to give the dihydroflavonols. In addition, while flavone *C*-glycosides have been reported in anthurium spathe, the exact identity of the flavone and the flavone *C*-glycoside is not known. Although methylated anthocyanins are included in the diagram, they have only been reported in *A. amnicola*. The portion of the pathway leading to delphinidin pigments have been included even though they are not found in anthurium spathe. They have been included given the interest in generating blue-coloured anthurium spathe. The full names of the enzymes are: PAL: phenylalanine ammonia lyase, C4H: cinnamate 4-hydroxylase, 4CL: 4-Coumarate: CoA ligase, 4CH: 4-Coumarate: CoA 3-hydroxylase, ACC: acetyl-CoA carboxylase, CHS: chalcone synthase, CHI: chalcone isomerase, FNS: flavone synthase, F2H: flavanone 2-hydroxylase, F3H: flavanone 3-hydroxylase, F3'H: flavonoid 3-hydroxylase, F3'5'H: flavonoid 3'5'-hydroxylase, DFR: dihydroflavonol 4-reductase, ANS: anthocyanidin synthase, UF3GT: UDP- glucose: flavonoid 3-*O*-glycosyltransferase, RT: rhamnosyl transferase and MT: methyltransferase.


Phenylalanine
PAL
Cinnamate
C4H
4-Coumaricacid
4CL
4-Coumaroyl-CoA
4CH
Caffeoyl-CoA
ACC
3 x Malonyl-CoA
3 x Acetyl-CoA
CHS
CHS
Chalcones
Naringenin chalcone
Eriodictyol chalcone
CHI
CHI
F3'H
F3'5'H
Naringenin
Eriodictyol
Pentahydroxyflavanone
Flavones and
flavone C-glycosides
FNS
2-OH-Naringenin
F2H ?
F3H
F3H
F3H
Dihydroflavonols
Dihydrokaempferol
Dihydroquercetin
Dihydromyricetin
F3'H
F3'5'H
DFR
DFR
DFR
Leucoanthocyanidins
Leucopelargonidin
Leucocyanidin
Leucodelphinidin
F3'H
F3'5'H
ANS
ANS
ANS
Anthocyanidins
Pelargonidin
Cyanidin
Delphinidin
UF3GT
UF3GT
UF3GT
Anthocyanins
Pelargonidin 3-O-glucoside
Cyanidin 3-O-glucoside
Delphinidin 3-O-glucoside
RT
RT
Pelargonidin 3-O-rutinoside
Cyanidin 3-O-rutinoside
MT
Methylated
anthocyanins
Peonidin 3-O-rutinoside

between both proteins is the subject of much research and discussion resulting in some previously classified *CHS* sequences being reclassified as stilbene synthases (discussed in Chapter 3).

1.8.3 Biosynthesis of flavones

The second step in flavonoid biosynthesis involves the isomerisation of the chalcone into a colourless flavanone, usually (2S)-flavanone naringenin through an intramolecular reaction in which the C-ring is closed (Figure 1.4) (Harborne, 1988). This reaction is catalysed by the enzyme chalcone-flavanone isomerase (CHI). However, this reaction can occur non-enzymically as evidenced by documented cases of spontaneous isomerisation occurring in vitro (Moustafa and Wong, 1967) and may also occur in vivo, as maize suspension cells, with no detectable CHI activity, can still synthesise anthocyanin (Grotewold et al., 1998). A similar situation has been reported in some Dianthus lines (Forkmann and Dangelmayr, 1980).

Flavanones represent a branch point in the pathway, as they are direct precursors for flavones and two flavonoid intermediates, namely flavan 4-ols and dihydroflavonols (Harborne, 1988; Heller and Forkmann, 1993). As mentioned, before, flavones (of the *C*-glycoside type) are the main non-anthocyanin flavonoids that accumulate in anthurium spathe tissue.

Flavones are synthesised by an oxidative reaction, which introduces a double bond between the C2 and C3 carbon (Figure 1.4) (Heller and Forkmann, 1993). Two very different classes of biosynthetic enzymes have been shown to catalyse this reaction. In *Petroselinum hortense* (parsley) a 2-oxoglutarate-dependent dehydrogenase (FNS I) accomplishes the desaturation reaction by abstracting hydrogen from the flavanone naringenin to form the flavone apigenin (Britsch, 1990). In contrast, it is a microsomal cytochrome P450 monooxygenase FNS II that catalyses the reaction (via abstraction of water) in a range of species including *Gerbera hybrida* (gerbera) (Martens and Forkmann, 1999), antirrhinum, *Torenia fournieri* (torenia) (Akashi et al., 1999) and *Glycine max* (soybean) (Kochs and

Griesbach, 1987). Indeed, the use of FNS II appears to be more common than the use of FNS I.

An interesting variation is seen in *Glycyrrhiza echinata* (licorice). According to Afchar et al. (1984), *C*-glucosyl flavones accumulate in the tissue of licorice, and Kerscher and Franz (1987) have shown that 2-hydroxylation of flavanones is an essential prerequisite for *C*-glycosylation. Subsequently, a P450 monooxygenase called (2S)-flavanone 2-hydroxylase (F2H) (if the substrate is naringenin) or licodione synthase (LS) (if the substrate is liquiritigenin) was characterised from elicitor-treated suspension-cultured cells of licorice (Akashi et al., 1998). Unlike the FNS II, the licorice P450 (F2H or LS) only catalysed the formation of the 2-hydroxy-flavanone intermediate (Figure 1.4) (Akashi et al., 1998; Martens and Forkmann, 1999). The flavone, apigenin, was only formed upon acid treatment of the respective enzyme assay (Akashi et al., 1998). Thus, FNS II produces flavones directly by a dehydrogenation reaction, while the closely related F2H proceeds via a 2-hydroxy- flavanone intermediate (once naringenin is the substrate). A similar enzymatic model is suggested for the production of isoflavones in legumes through the dehydration of 2-hydroxyisoflavone by the P450 isoflavone synthase (Kochs and Griesbach, 1986; Hashim et al., 1990).

Whether the production of *C*-glycosylated flavones in anthurium proceeds via a FNS I, FNS II or F2H enzyme has not been verified biochemically. However, given that production of flavones by P450 enzymes seems to be the more common route it would be no surprise if the same were to be found in anthurium spathe tissue.

1.8.4 Biosynthesis of anthocyanins B-ring hydroxylation

When flavanone substrates are hydroxylated at the C3 position, the colourless dihydroflavonols are formed (Figure 1.4). Flavanone 3-hydroxylase (F3H) catalyses the reaction and this cytoplasmic enzyme requires 2-oxoglutarate, Fe^{2+} and ascorbate as cofactors (Fritsch and Griesbach, 1975; Heller and Forkmann, 1993). Naringenin is converted to dihydrokaempferol (DHK) by this biochemical route.

Pelargonidin, cyanidin and delphinidin differ only in the degree of hydroxylation of the B-ring, being, 4', 3'4' and 3'4'5' hydroxylated respectively (Figure 1.2 and 1.4). The hydroxyl group in the 4' position of the B-ring in dihydrokaempferol is incorporated into the flavonoid structure from the CHS substrate 4-coumaroyl-CoA. The additional B-ring hydroxylation reactions at positions 3' and 5' are usually catalysed by two membrane bound P450 cytochrome monooxygenases acting at the C_{15} level (Heller and Forkmann, 1988). The two enzymes are flavonoid 3'-hydroxylase (F3'H) and flavonoid 3'5'-hydroxylase (F3'5'H) and they can hydroxylate naringenin and dihydrokaempferol in the respective 3' and 5' positions (Harrison and Stickland, 1978; Forkmann and Dangelmayr, 1980; Forkmann and Stotz, 1981; Heller and Forkmann, 1988). In addition, in vivo feeding experiments showed that leucoanthocyanidins can act as substrates for F3'H in *Petunia hybrida* (petunia) and F3'5'H in *Eustoma grandiflorum* (lisianthus) (Schwinn, 1994).

This broad range of substrate specificity of F3'H may represent the 'commitment' of the plant to produce cyanidin derivatives. However, the substrate that F3'H actually uses depends in part on the competition with other biosynthetic enzymes for the same substrate. As the hydroxylation of the B-ring determines which of the anthocyanidin types are formed, the availability of the genes encoding F3'H and F3'5'H may be of particular importance for the manipulation of flower colour.

When dihydrokaempferol (DHK) is used as the substrate for F3'H activity, dihydroquercetin (DHQ) is formed. Acting directly on DHK or on DHQ, F3'5'H forms dihydromyricetin (DHM). These dihydroflavonols determine the anthocyanin type that is formed (Figure 1.4). DHQ is the precursor for red-coloured cyanidin pigments while DHM is the precursor for the formation of blue/purple delphinidin pigments. In the absence of both hydroxylases, DHK accumulates which is a precursor of the orange-coloured pelargonidin pigments (van Tunen and Mol, 1991; Heller and Forkmann, 1993).

Only F3'H activity is suggested in anthurium, given the presence of cyanidin pigments, although it is not known at what stage in the pathway the hydroxylation of the B-ring occurs. Delphinidin-based anthocyanins do not accumulate in anthurium spathe tissue and may suggest, among other things, that F3'5'H activity is absent. Delphinidins have only been recorded in one genus of the Araceae family, *Schismatoglottis coccina* (Williams et al., 1981).

Several enzymatic steps are still needed to convert dihydroflavonols to anthocyanins, and the products must also be transported across the vacuolar membrane because dihydroflavonols are cystolic whereas anthocyanins are found exclusively in the vacuole (Kays et al., 1981). This stage of the pathway is very complex and requires the action of a number of different enzymes.

The first step towards anthocyanin synthesis from the dihydroflavonols is an NADPH-dependent reaction, involving a reduction at the 4-position of the C-ring by dihydroflavonol 4- reductase (DFR), which yields leucoanthocyanidin, a colourless and unstable intermediate (Figure 1.4) (Stafford and Lester, 1982; Heller and Forkmann, 1993).

An interesting feature of DFR is the substrate specificity it displays (Gerats et al., 1982; Heller et al., 1985; Forkmann and Ruhnau, 1987; Johnson et al., 1999). The amino acid motif responsible for this trait has been identified through mutagenic experiments with gerbera and a conserved asparagine within this motif is thought to be key in conferring substrate specificity to DFR, while a conserved glutamate if mutated to leucine, abolished enzyme activity (Beld et al., 1989; Johnson et al., 2001).

Feeding experiments performed on genetically defined acyanic mutants of *Mathiola incana* (mathiola), *Callistephus chinensis*, *Dianthus caryophyllus* (carnation), petunia and antirrhinum, have showed convincingly that leucoanthocyanidins are precursors to anthocyanin synthesis (Heller and Forkmann, 1988). Recently, There has been significant gain in our understanding of the biochemistry of this conversion. Anthocyanidin synthase (ANS) cloned from *Perilla frutescens* (perilla) was

functionally characterised as a 2-oxoglutarate-dependent oxygenase (Saito et al., 1999). Recombinant ANS and UDP-glucose:flavonoid 3-*O*-glucosyltransferase (UF3GT) protein expressed in *E. coli* under physiological conditions catalysed the formation of anthocyanidin 3-glucoside directly from the leucoanthocyanidin (Nakajima et. al., 2001). This model is further supported by the structural studies on ANS (Wilmouth et al., 2002) and excludes the need for an external dehydratase as was previously suggested (Heller and Forkmann, 1988, 1993; Saito et al., 1999).

Owing to their instability under normal physiological conditions anthocyanidins do not accumulate in plant cells. They are normally stabilised to form the corresponding anthocyanin by the addition of a sugar residue(s), usually at the 3' or 5' position, by various glycosyltransferases (GTs). The first reaction is the addition of a glucose moiety (UDP-glucose) catalysed by the GT enzyme UDP-glucose: flavonoid-3-glucosyltransferase (UF3GT) (Ferdoroff et al., 1984; Martin et al., 1991; Heller and Forkmann, 1993). Glucosylation is commonly followed by the addition of a rhamnose moiety to the glucose molecule catalysed by UDP-rhamnose: anthocyanidin-3-*O*-glucoside rhamnosyltransferase (3RT). Anthurium anthocyanidins are glycoconjugated at the 3' position with glucose and rhamnose resulting in the diglycosidic anthocyanin-rutinoside (Figure 1.4) (Iwata et al., 1979). Thus both UF3GT and 3RT enzyme activity are suggested in anthurium spathe tissue. Yet another class of GT must also exist in anthurium to catalyse *C*-glycosidation in the formation of the flavone *C*-glycosides. An enzyme for this step has not been reported in floral tissue.

The presence of peonidin glucosides, although only reported in one *Anthurium* species *(A. amnicola),* suggests the presence of *O*-methyltransferases in this species. These enzymes transfer a methyl group from S-adenosylmethionine to hydroxyl groups of flavonoids at the C-3' position. Peonidin is formed when cyanidin is methylated in the 3'-position (Figure 1.4) (Jonsson et al., 1982, 1984a,b). It is not known at what stage of anthocyanin biosynthesis in *A. amnicola* that methylation occurs. In petunia 3'-methylation occurs after the formation of the 3-*O*-rutinoside. The availability of *A. amnicola* cDNA clone for this enzyme may be useful in attempts to develop the purple hue in anthurium. Although in

other plant species subsequent ring modifications such as acylation, sulphonation and further glycosylation do occur, these are not known to occur in anthurium.

1.8.5 Transfer of anthocyanins to the vacuole

All the reactions described thus far occur within the cytosol, but anthocyanins accumulate within the vacuoles of cells where the environment is more stable and their full colour can be developed. Therefore, to complete the pathway, the anthocyanidin glycoside must be transported across the vacuolar membrane, from the cytosol to the vacuole. In maize, a glutathione-S-transferase (GST) attaches a glutathione tag to the anthocyanin molecule, allowing it to be transported into the vacuole possibly by a glutathione pump (Marrs et al., 1995). The *Bronze2 (Bz2)* locus in maize encodes the GST, and when mutated, produces bronze coloured kernels resulting from oxidation and condensation of the anthocyanins in the cytosol (Marrs et al., 1995). This gene is under similar regulatory control as the anthocyanin structural genes (Taylor and Briggs, 1990; Marrs et al., 1995). Similar genes have been cloned from petunia and *Arabidopsis thaliana* (arabidopsis). In petunia, the GST is encoded by *AN9*, which if mutated, results in acyanic flowers (Alfenito et al., 1998). In arabidopsis, the *TT12* was found to encode a GST necessary for flavonoid sequestration into the vacuoles of the seed coat endothelium (Debeaujon et al., 2001).

1.9 TEMPORAL EXPRESSION OF ANTHOCYANIN BIOSYNTHETIC GENES

Any initiative to manipulate flower colour must be based on a thorough understanding of the biochemical and molecular aspects of flower colour development in that species. Therefore, characterising the expression patterns of the flavonoid biosynthetic genes is central to a molecular investigation of flower colour development.

The control of anthocyanin structural gene expression is usually reflected in their temporal expression pattern. The temporal expression patterns of flavonoid biosynthetic genes have been documented for several plant species, such as antirrhinum (Martin et al., 1991; Jackson et al., 1992), petunia (Quattrocchio et al., 1993), maize (Dooner, 1983), lisianthus

(Davies et al., 1993; Oren-Shamir et al., 1999), arabidopsis (Pelletier and Shirley, 1996), *Vitis vinifera* (grapes) (Boss et al., 1996) and perilla (Gong et al., 1997).

For most dicots investigated, the cascade of genes involved in anthocyanin biosynthesis appear to be separated into two subsets that represent the division of the pathway into two independently regulated units. One group consists of the early biosynthetic genes (EBGs) and the other is comprised of the late biosynthetic genes (LBGs) (Beld et al., 1989; Jackson et al., 1992; Martin and Gerats, 1993; Quattrocchio et al., 1993; Mol et al., 1998). It is the LBG expression, which peaks later than the EBGs, that is correlated with pigment production. In antirrhinum, *CHS* and *CHI* constitute the EBGs and *F3H*, *DFR*, *ANS* the LBGs (Jackson et al., 1992). In petunia, a similar pattern exists except that *F3H* is coordinately regulated with *CHS* and *CHI* and the LBGs comprise the genes from *DFR* onward (van Tunen et al., 1988; Quattrocchio et al., 1993; Weiss et al., 1993; Quattrocchio et al., 1998). Similar patterns to these traditional dicot models are found in lisianthus and arabidopsis. In lisianthus, *CHS* and *CHI* form the EBGs, while *F3H* is coordinately regulated with the anthocyanin specific genes, and in arabidopsis, LBGs include *DFR* and *ANS*, while *CHS*, *CHI* and *F3H* constitute the EBGs.

There are dicot species, such as grape and perilla where this pattern does not hold true. In acyanic grape berries, all genes have reduced expression levels, whereas, prior to pigment accumulation in red berries, transcripts for all genes, from *CHS* to *ANS*, are strongly expressed. Anthocyanin production is concomitant with *UFGT* expression (Boss et al., 1996). While it is thought that the absence of *UFGT* expression is responsible for white grapes, it is not known if this is due to mutation(s) in the structural or regulatory gene(s) (Boss et al., 1996). In perilla, no distinct subdivision into LBGs and EBGs is observed as the transcript abundance of all structural genes examined was coordinately regulated (Gong et al., 1997).

The only monocot where similar research has been conducted is maize. In this case all the structural genes involved in anthocyanin biosynthesis were coordinately induced (Dooner, 1983; Paz-Ares et al., 1987; Ludwig et al., 1989). Therefore, the transcripts of all the

structural genes involved in anthocyanin production in the maize kernel have the same temporal distribution pattern (reviewed in Mol et al., 1998).

Comparing maize to the traditional dicot models, where modular control of the pathway is observed, inferences can be made as to the evolutionary significance of discrete regulatory units (Martin et al., 1991). In maize, the pigmentation of the kernel may primarily be associated with attracting animals for seed dispersal. Therefore, there may be no selective advantage in modular control of anthocyanin biosynthetic genes in this tissue. However, in flowers, flavones and flavonols absorb UV light and may also play a role as insect attractants. Therefore, modular control of anthocyanin production allows for regulated control of the biosynthesis of these pigments separate from that of anthocyanin production. On this basis, the regulatory control in maize may be very different from that in the leaf or flower of another monocot species and may be of considerable relevance to pigment production in anthurium.

Since flavonoid biosynthetic genes are controlled at the level of transcription, then the temporal distribution of the genes serves as a window to the pattern of regulation of anthocyanin biosynthesis in that species. Regulatory proteins involved in anthocyanin biosynthesis have been defined in only a few plant model systems and are discussed in the following sections.

1.10 GENE TRANSCRIPTION

Regulating the expression of genes is fundamental to an organism's ability to control its biological processes. Most regulation of expression occurs at the level of transcription and is mediated by transcription factors that either activate or repress transcription of the gene. Much of our knowledge in this area has come from studies of eukaryotic transcription activators and these studies have revealed a highly intricate system that is tightly regulated at many levels. Regulation of transcription is effected by sequence specific binding of the transcription factor to *cis*-acting elements in the promoter or enhancer regions of the target gene. This binding facilitates the assembly of the transcription factor complex, which includes several basic transcription factors and RNA polymerases, to initiate mRNA

synthesis. Consistent with their function, transcription factors have a DNA binding domain and a regulatory activation domain (or a repressor domain for factors that repress transcription). These domains are functionally separable and can be repositioned within the protein or exchanged between proteins without loss of function (Ptashne, 1988). Further to this, it is the activation domain that is thought to interact with the basic transcription machinery (Ptashne, 1988). Several families of transcription factors have been identified both in animals and plants, based on characteristic structural DNA-binding motifs (e.g. Myb, bZIP, bHLH, MADS-box) (Meshi and Iwabuchi, 1995). While the DNA binding domain tends to be specific and highly conserved for a certain class of transcription factor, the activating regions are less precisely defined, and exhibit extensive variation within a given class.

There is a growing body of knowledge on the precise steps involved in the transcription of a gene. The ability to perform in vitro transcription reactions (Buratowski et al., 1989) has shed light on the general order of assembly for the transcription factor complex. While several enzymes and a growing number of cofactors (Inostroza et al., 1992; Nakajima et al., 1997) are involved at various levels of gene transcription, for protein-encoding genes, the key enzyme involved in eukaryotic systems is RNA polymerase II (RNA Pol II). Though structurally complex (it is made up 12-sub units), it cannot recognise its target promoters directly. Added to which, it is required to modulate production of the RNA transcripts of individual genes in response to developmental and environmental signals (Nikolov and Burley, 1997). To accomplish these two functions it requires interaction with variety of proteins that comprise the basal transcription machinery which includes TATA-binding protein (TBP), general transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH), TBP-associate factors (TAFs), and Switch/Snif (SWN/SNF) proteins which are involved in chromatin remodeling events. The process of gene transcription can be summarised into three phases. The first is initiation, which involves binding to the TATA-box in the target DNA by TFIID and the assembly of the initiation complex. This is followed by the unwinding of the DNA through the helicase activity of TFIIH and subsequent synthesis of an RNA transcript by RNA Pol II, in the presence of nucleoside triphosphates. Transcription is terminated by the addition a poly $(A)^+$ signal located downstream of the

last exon. This signal is used to add a series of adenylate residues during RNA processing. Transcription often terminates at 0.5-2 kb downstream of the poly $(A)^+$ signal.

Two models have been advanced to explain the many protein-protein interactions associated with the transcription of a gene. The first is a step-wise assembly developed from in vitro transcription factor studies (Buratowski et al., 1989). This model assumes an ordered sequential assembly of the transcription pre-initiation complex. However, such a process appears to be inefficient given the extraordinarily large size of the complex. In addition, it is hard to envisage the assembly of such a complex within the time frame necessary at each promoter (Lemon and Tjian, 2000). The competing model is at the other end of the spectrum and is described as the pre-assembly model. Founded on the discovery that RNA Pol II existed as a holoenzyme, complexed with several members of the basal transcription machinery, such as SWI/SNF, the model assumes the recruitment of a completely assembled RNA Pol II pre-initiation complex (Koleske and Young, 1994; Parvin and Young, 1998). However, the validity of this model is also being challenged as new information becomes available (Zawel et al., 1995; Kimura et al., 1999). In addition to these two models, there is the possibility that genes may undergo an early chromatin arrangement involving interactions with transcription factors and chromatin-associated remodeling factors that render the gene competent for activation prior to binding of the initiation complex (Cosma et al., 1999; Krebs et al., 1999).

1.11 Myb TRANSCRIPTION FACTORS

Myb proteins play a key role in the regulation of structural gene expression for anthocyanin biosynthesis. They are a unique family of DNA binding proteins that bind DNA as monomers, unlike other sequence-specific DNA binding proteins (Howe et al., 1990), and induce a bend in the target sequence (Saikumar et al., 1994; Solano et al., 1995a,b). The cellular proto-oncogene c-Myb and its truncated version v-Myb, isolated from the avian myeloblastosis virus, were the first Myb proteins discovered and characterised as DNA binding proteins acting as transcription factors (Klempnauer et al., 1982; Biedenkapp et al., 1988).

1.11.1 Myb distribution and diversity of function

Myb proteins have been found in a wide range of eukaryotes such as humans (Nomura et al., 1988), chicken, mouse (Howe et al., 1990; Lipsick, 1996), *Xenopus* (Amaravadi and King, 1994), *Drosophila* (Peters et al., 1987), fungi (Tice-Baldwin et al., 1989; Ohi et al., 1994; Wieser and Adams, 1995), slime mould (Stober-Grasser et al., 1992; Guo et al., 1999), a wide selection of angiosperms including both dicots and monocots (Jackson et al., 1991; Avila et al., 1993; Baranowskij et al., 1994; Lin et al., 1996; Romero et al., 1998), gymnosperms such as black spruce (Charest et al., 1994) and mosses, for example *Physcomitrella patens* (Leech et al., 1993).

The first plant Myb gene to be identified was the maize anthocyanin regulator C1 (Paz-Ares et al., 1987). Since that time, the number of Myb proteins in plants has expanded enormously with arabidopsis alone assumed to have in excess of 100 Myb genes (Romero et al., 1998). A similarly large family is suggested for the maize genome (Rabinowicz et al., 1999), petunia (Avila et al., 1993) and tomatoes (Lin et al., 1996). In fact, for plants, Myb proteins constitute a large gene family of transcription factors that play a significant role in the transcription control of a diverse range of plant processes (Table 1.1).

This pattern seen in plants, is completely unlike that in vertebrates, where the Myb gene family is small and their function seems limited to controlling cellular proliferation and differentiation (Weston, 1998). The large size of the Myb family in plants, can in part, be explained by genetic redundancy and overlapping functions, such that structurally similar Mybs have similar functions, recognising the same target genes, and are involved in similar processes but are expressed in different tissues. AN2 is an R2R3 Myb related protein that acts as the main determinant of colour differences in the petunia corolla (Quattrocchio, 1994). AN2 regulates anthocyanin production primarily in the limb, although limited expression can be detected in the tube and pistil (de Vlaming et al., 1984, Quattrocchio et al., 1993). The paralogous AN4 appears to control anthocyanin production in a separate tissue, anthers, and it is the likely candidate to substitute for AN2 activity in tissues where

Table 1.1 The range of plant processes controlled by Myb proteins.

Plant process	Reference
Phenylpropanoid metabolism	Romero et al., 1998; Jin and Martin, 1999 and several references throughout the thesis
Plant defense	Yang and Klessig, 1996
Cellular differentiation	Oppenheimer et al., 1991; Wada et al., 1997
Cell shape	Noda et al., 1994; Mur, 1995.
Circadian clock	Wang et al., 1997; Wang and Tobin, 1998; Mizoguchi et al., 2002; Carre and Kim, 2002
Stress response	Urao et al., 1996; Itturriaga et al., 1996; Lu et al., 2002
Seed development and germination	Gubler et al., 1995; Suzuki et al., 1997
Leaf polarity	Waites et al., 1998; Galego and Almeida, 2002
Trichome development	Oppenheimer et al., 1991; Sawa, 2002
Lateral meristem initiation	Schmitz et al., 2002

AN2 is not found (Quattrocchio et al., 1993; Huits et al., 1994; Quattrocchio et al., 1998). Similarly in maize, two Myb proteins identified as C1 (Paz-Ares et al., 1986, 1987) and purple leaf (PL) (Cone et al., 1993) have been definitively characterised as transcriptional activators of the structural genes for anthocyanin biosynthesis (Sainz et al., 1997a). The *C1* and *Pl* gene products sharing 80% identity are functionally equivalent regulating the same genes but in different tissues (Cone et al., 1993). C1 is expressed in the aleurone of the kernel and the embryo and controls pigment production in these tissues (Chen and Coe, 1977; Coe, 1985) while PL is required for pigmentation of most of the plant body (Cone et al., 1993). Further evidence for their functional similarity was provided by the *PL-Bh* allele, that can substitute for C1 activity in the kernel (Cocciolone and Cone 1993).

However, it is possible for plant Mybs with very similar DNA binding domains to be involved in very different physiological processes. For example, GLABRA1 (GL1) and ROSEA1 are classed in the same Myb subgroup based on the similarity of their Myb domain (Jin and Martin, 1999) but GL1 is involved in trichome specification and probably

mediates its activity by modifying cell size, while ROSEA1 regulates anthocyanin biosynthesis (Oppenheimer et al., 1991; Schwinn, 1999). Therefore, while some redundancy may exist, diversity and flexibility are accomplished in having a large gene family. The overlapping function and apparent redundancy of Myb proteins in plants may represent a unique feature by plants to selectively use Myb transcription factors to regulate their specific physiological processes (Martin and Paz-Ares, 1997).

1.11.2 Defining features of Myb proteins: The Myb domain

Myb proteins have a modular structure consisting of an N-terminal DNA binding domain, and a C-terminal transactivation domain (only for transcriptional activators). Animal Mybs may also have a repressor domain in the C-terminus. The distinguishing feature of all Myb proteins is the DNA binding domain that specifically binds to sequences of DNA in the promoter of target genes (Frampton et al., 1989; Sakura et al., 1989; Howe et al., 1990). The transcriptional activation domain, in the C-terminal end, is believed to interact with the target DNA in such a way as to promote the initiation of transcription (Ptashne, 1988).

1.11.2.1 Three repeat Mybs

The DNA binding domain of the prototypic vertebrate cellular proto-oncogene c-Myb comprises three imperfect repeats (R1, R2 and R3) and regularly spaced tryptophan residues, which provide a hydrophobic core (Anton and Frampton, 1988; Ogata et al., 1992, 1994). Each repeat adopts a helix-helix-turn-helix configuration for interaction with the major groove of DNA (Ogata et al., 1994). The first repeat is not essential for DNA binding (Biedenkapp et al., 1988) and cooperative action between R2 and R3 is the suggested mechanism for binding, as neither of the two repeats can bind DNA independently (Tanikawa et al., 1993). There is evidence for direct interaction at specific contact points to form the cooperative unit that recognises the target sequence in the major grooves of the DNA (Ogata et al., 1994). It is the R2 and R3 domains that are required for sequence specific binding, with the third helix in each repeat serving as the recognition helix (Ogata et al., 1994). Plant proteins with three Myb motifs have recently been discovered in arabidopsis, where they control the cell cycle (Braun and Grotewold, 1999; Kranz et al.,

2000; Ito et al., 2001; Lu et al., 2002). In fact, they have been observed in all major evolutionary lineages in plants from bryophytes to monocots (Kranz et al., 2000).

The three-repeat plant proteins were shown to be more similar to vertebrate c-Myb than to other plant Mybs (Braun and Grotewold, 1999; Kranz et al., 2000). In addition, the putative amino acid motif in the R2 repeat that determines recognition specificity and the target genes of three repeat Mybs is also conserved in the plant counterparts (Kranz et al., 2000). This suggests R1R2R3 Mybs may have a conserved function in eukaryotes, that is cell cycle control and differentiation (Kranz et al., 2000), whereas R2R3 Mybs regulate plant specific processes that evolved during plant speciation (Kranz et al., 2000).

1.11.2.2 Two-repeat Mybs

Unlike the three-repeat configuration in animal Mybs, most plant Mybs have two repeats (R2 and R3) which are most similar to the second and third repeats of the prototypic c-Myb (Williams and Grotewold, 1997). In fact, to date, plants appear to be the only source of R2R3 Myb genes (Riechmann et al., 2000). The R2R3 Myb group is the largest Myb group in plants and given the extensive range of processes that they regulate, it had been suggested, that plant R2R3 Mybs were the equivalent of the R1R2R3 Mybs and that the three Myb-repeat DNA binding domain did not exist before the divergence of animal and plant lineages (Rosinski and Atchley, 1998). However, the discovery of the three-repeat Mybs throughout the plant kingdom, coupled with their strikingly similar gene intron/exon structure to those of animal three-repeat Mybs, suggest that the three-repeat Myb domain formed prior to the divergence of plant and animals.

The discovery of three-repeat Myb proteins across all plant lineages does support the hypothesis that the two-repeat Mybs may have evolved from loss of the first repeat (R1) from an ancient R1R2R3 ancestor (Lipsick, 1996). R1 has been shown not to be essential for Myb protein function (Biedenkapp et al., 1988). Therefore, the vast number of R2R3 Mybs may have resulted from the duplication of entire genes within plant systems (Lipsick, 1996).

1.11.2.3 Single repeat Mybs

A growing number of single-repeat Mybs (R1 or R2) have been identified in plants (Baranowskij et al.,1994; Kirik and Baumlein, 1996; Feldbrugge et al., 1997; Wada et al., 1997; Wang et al., 1997; Schaffer et al., 1998; Lu et al., 2002; Sawa 2002). It has been suggested that one-repeat Mybs may have a different mechanism for DNA binding compared to that of two or three-repeat Mybs (Jin and Martin, 1999), but this is not supported by current research. NMR analysis of the DNA binding domain of the single-repeat Mybs, TFR1 and TFR2, bound to telomeric DNA has revealed three helices with architecture similar to the three repeats of c-Myb (Nishikawa et al., 2001). Sequence specific binding is possible with only the single repeat and again, the third helix of TFR1 is the recognition helix sitting in the major grove of the target DNA, a configuration that parallels that seen in three and two-repeat Mybs (Bianchi et al., 1997; Nishikawa et al., 2001).

1.11.2.4 Myb domain features for plant and animal Mybs

Certain key differences within the DNA binding domain of plant R2R3 Myb distinguish them from their animal counterparts. One such difference is that in most plant Mybs, there is a conservative substitution of the tryptophan residue in the third repeat (Martin and Paz-Ares, 1997). Additionally, there is an extra leucine residue between the second and third helices of the R2 repeat in plants, coupled with a glutamine to serine substitution in the R2 DNA recognition helix (Williams and Grotewold, 1997). Such changes may alter the recognition specificities of the recognition helix as well as the co-operative activity between the R2 and R3 repeats that underlines the DNA binding activity of these transcription factors. In fact a chimaeric protein of v-Myb and maize P gene exhibited a novel DNA binding specificity (Williams and Grotewold, 1997). Apart from these differences between plant and animal Mybs, there is greater similarity between the same repeat in different proteins than between the repeats in the same protein, suggesting that each repeat has a unique function (Martin and Paz-Ares, 1997).

1.11.3 Defining features of Mybs: The C-terminus

The marked level of conservation observed in the DNA binding domain does not apply for the C-terminal end of Myb proteins. In fact proteins that share very high levels of conservation between DNA binding domains can exhibit significant divergence at the C-terminal end (Rosinski and Atchley, 1998). However, certain motifs either outside the Myb domain or within the C-terminal end have been identified as being conserved within different Myb subgroups (Marocco et al., 1989; Jackson et al., 1991; Avila et al., 1993; Kranz et al., 1998; Nesi et al., 2001).

While the DNA binding properties are a function of the N-terminal Myb domain, transactivation domains are usually positioned in the C-terminus. Several studies have shown that the activation domains function by directly contacting the general transcriptional factors (Stringer et al., 1990; Lin and Green, 1991; Xiao et al., 1994). These activation domains can be rich in acidic amino acids, proline or glutamine (Ptashne, 1988). For Mybs associated with anthocyanin biosynthesis, an acidic activation domain is most common. Acidic activation domains usually bear a significant net negative charge (Gill and Ptashne, 1987; Ma and Ptashne, 1987; Gill et al., 1990). The ability of such domains to form an amphipathic α-helix was consequently suggested as being central to their ability to serve as activation domains (Ptashne, 1988). However, only in some cases has it been confirmed that these domains are essential for transactivation (Goff et al., 1992; Baranowskij et al., 1994; Urao et al., 1996). In addition, questions of the importance of an amphipathic α-helix to the function of Myb transcription factors have been raised as recent experiments suggest that specific amino acids within the C-terminus are of equal or more importance than the formation of an α-helix (Sainz et al., 1997b). Also transcription activation was maintained in Myb proteins for which the α-helix had been disrupted (Sainz et al., 1997b).

It is useful to bear in mind that all Myb-related proteins need not be transcriptional activators. Some may in fact serve as repressors as was demonstrated for arabidopsis AtMYB4 (Jin et al., 2000) and *(Fragaria x ananassa)* strawberry FaMYB1 (Aharoni et al.,

2001). AtMYB4 was suggested to act by direct repression and by competition with the activators for the binding motifs in the promoters of target genes (Jin et al., 2000). The peptide sequence NLELRISLPDDV was shown to be a prerequisite for the repression activity (Jin et al., 2000). A similar sequence was also found in the Myb repressor FaMYB1 from strawberry (Aharoni et al., 2001). The balance between Myb activators and repressors may provide extra flexibility in transcriptional control as well as provide a mechanism for regulating transcription factor activity.

1.11.4 Myb recognition specificity

Although Myb proteins share a highly conserved DNA binding domain, they show considerable variation in the specific *cis*-elements they recognise. Examining DNA binding specificities by electrophoretic mobility shift assays Solano et al. (1995a) and Romero et al. (1998) showed that Myb proteins from mammals, invertebrates and slime moulds bind to the site T/C AAC G/T G A/C/T A/C/T (MBSI), while the motifs recognised by plant Mybs include, MBSI, TAACTAACT (MBSII), G G/TTA/T GGT A/G (MBSIIG) and its related C/TACCA/TAACC. The majority of plant R2R3 recognise MBSIIG (Solano et al., 1995a; Romero et al., 1998; Jin and Martin, 1999).

As previously stated, Myb proteins with similar DNA binding domains may yet regulate different target genes (Jin and Martin, 1999). Other factors, in addition to the differences in the DNA binding domain, may contribute to the flexibility of target motifs recognised by R2R3 Mybs. For example, if transcription factors are assembled as a complex with specific proteins within the cell prior to binding their target DNA, then the variability in the matrix of proteins present within that cell would help determine the nature of the complex formed. Therefore, two Myb proteins with very similar DNA binding domains can, by virtue of the proteins available within the cell to form the protein complex, recognise different target sequences and so be involved in distinct physiological functions (Jin and Martin, 1999).

1.11.5 Myb interactions with other proteins

Gene activation in eukaryotes is combinatorial, depending on cooperation between different transcription factors. In animal systems, Myb related proteins interact with several different classes of transcription factors to activate their target genes (Lipsick, 1996; Davies et al., 1999; Pinson et al., 2000). In plants two types of transcription factors, basic helix-loop-helix (bHLH) and basic region/leucine zipper (bZIP) proteins have been identified as transcriptional coactivators for Myb, to date. The regulation of anthocyanin biosynthesis in both monocot and dicot species is mediated through interaction between Myb and bHLH proteins (discussed in detail in the following section). Myb/bHLH factors are also involved in regulating trichome development (Oppenheimer et al., 1991; Walker et al., 1999; Payne et al., 2000), and have also been implicated in the activation of drought and other stress responsive genes (Abe et al., 1997; Smolen et al., 2002).

The bHLH transcription factors consist of two domains characteristic of MYC transcription factors, the N-terminal basic domain and the C-terminal HLH motif that mediates protein interaction through homo and heterodimerisation (Ludwig et al., 1989; Murre et al., 1989a,b; Weintraub et al., 1991). The basic domain is responsible for DNA recognition and binding through the formation of an α-helical structure that interacts with the major groove of the DNA (Ferre-D'Amare et al., 1993). Yeast two hybrid experiments with maize anthocyanin Myb regulators have shown that the N-terminal region of the Myb domain interacts with the N-terminal region of its bHLH partner and that the HLH motif is not required for this interaction (Goff et al., 1992). In fact, no transactivation domain was found in the maize bHLH protein regulating anthocyanin biosynthesis, suggesting that it mediates activation through protein-protein interaction (Goff et al., 1992). Further studies with maize suggest that the Myb/bHLH complex may bypass inhibitory effects of the Myb domain on the Myb activation domain (Goff et al., 1991; Grotewold et al., 1994; Sainz et al., 1997a)

The second class of proteins, bZIP, are characterised by a 25 amino acid stretch of basic residues that specifies and contacts the ACGT core sequence in the target DNA, and a leucine rich zipper domain responsible for protein dimerisation (Ellenberger, 1994). This

motif has been found associated with genes encoding flavonoid biosynthetic enzymes (Schulze-Lefert et al., 1989). The binding site preferences of plant bZIP proteins depends on sequences flanking the ACGT core and two of the main types are the G and C box elements (Schindler et al., 1992a,b). Myb/bZIP interaction has been implicated in the regulation of early flavonoid genes in antirrhinum (Sablowski et al., 1994)

1.11.6 Regulation of anthocyanin biosynthesis in model species

The combinatorial action of Myb and bHLH proteins has been shown to be an evolutionarily conserved mechanism in plant species for the activation of structural genes involved in anthocyanin biosynthesis. The model was first demonstrated in maize (Goff et al., 1992; Cone et al., 1993) and since then has been observed in several plant model systems such as petunia (Quattrocchio et al., 1993, 1998; Spelt et al., 2000), antirrhinum (Martin et al., 1991; Schwinn, 1999), arabidopsis (Nesi et al., 2001) and perilla (Gong et al., 1999a,b). The defined proteins are listed in Table 1.2.

Table 1.2 Myb and bHLH proteins involved in regulating anthocyanin biosynthesis in five plant species.

Plant species	Myb	bHLH
petunia	AN2 (Quattrocchio, 1994) AN4 (Quattrocchio et al., 1993; Huits et al., 1994)	AN1 (Quattrocchio et al., 1998) JAF13 (Spelt et al., 2000)
antirrhinum	ROSEA1, ROSEA2 and VENOSA (Schwinn, 1999)	DELILA (Goodrich et al., 1992) MUTABILIS (Schwinn, 1999).
arabidopsis	*PAP1, *PAP2 (Borevitz et al., 2000)	Unknown
maize	C1 (Paz-Ares et al., 1986) PL (Cone et al., 1993)	R/B family (Styles et al., 1973; Ludwig and Wessler, 1990)
perilla	MYBP1 (Gong et al., 1999a)	MYC-RP, MYC-GP (Gong et al., 1999b)

*PAP1 and PAP2 were cloned by activation tagging. When over expressed, they resulted in widespread accumulation of purple anthocyanidin pigments throughout the entire plant tissue (Borevitz et al., 2000). There has been no report on knockout mutants for PAP1 and PAP2 to observe their effect on transcript levels for flavonoid biosynthetic genes in arabidopsis. Therefore, their precise function(s) under physiological conditions is/are not known.

The cytosolic WD repeat proteins is yet another class of proteins that have been shown to affect anthocyanin biosynthesis in petunia and arabidopsis (de Vetten et al., 1997; Walker et al., 1999). While their mechanism of action is not yet clear, they are thought to act upstream of anthocyanin Mybs (de Vetten et al., 1997) and affect anthocyanin production possibly through post-translational modification (of the Myb, bHLH or Myb/bHLH complex) or nuclear transport of anthocyanin Mybs (de Vetten et al., 1997; Walker et al., 1999). Recently, Sompornpailin et al. (2002) provided evidence for the physical interaction of MYB-RC and a perilla WD repeat protein. This protein-protein interaction resulted in the transport of the WD protein into the nucleus (Sompornpailin et al., 2002).

While the mechanism of Myb/bHLH interaction is conserved, different sets of structural genes are targeted, depending on the species as was suggested by the difference in temporal expression for genes in the pathway. In petunia, the target genes are from DFR onward, (Quattrocchio et al., 1993; Spelt et al., 2000), while in antirrhinum, a Myb/bHLH pair regulates genes from F3H to ANS (Jackson et al., 1992; Schwinn, 1999). However, in maize and perilla, all the anthocyanin biosynthetic genes appear to be targeted and regulated coordinately (Dooner, 1983; Ludwig et al., 1989; Gong et al., 1997).

In species where the anthocyanin production is controlled through the regulation of LBGs, the EBGs still need to be active for pigment to be produced. Therefore, for petunia and antirrhinum, the regulation of early genes in the pathway appears to be independent and may involve separate regulatory elements. In this regard, another class of transcription factors (bZIP proteins) has been implicated as an interaction partner for Mybs. Cooperative action between Myb and bZIP in plant systems have been demonstrated for the light induced transcription activation of *CHS* in parsley as a prerequisite for flavonoid accumulation in leaf epidermal cells (Weisshar et al., 1991; Feldbrügge et al., 1997). They

are also suggested as partners for AmMYB305 and AmMYB340 in regulating flavonol biosynthesis because bZIP recognition motifs have been found in the PAL promoter sequence (Sablowski et al., 1994).

1.11.7 In vivo transcription factor studies for anthocyanin Mybs

Using primary structural features to ascribe function to Myb transcription factors could be very misleading, making functional characterisation essential. Several techniques for functional assays have been developed and are reviewed in Schwechheimer et al. (1998). One technique that provides rapid, first hand information is an in vivo transient expression assay using particle bombardment, protoplast transformation or electroporation. Using one or more of these methods, functional data on a wide range of transcription factors has been obtained (Schwechheimer et al., 1998 and references therein). For conclusive results, assays should be performed with tissue that lacks endogenous activity of the transcription factor being investigated. In addition, cofactors known to be required for activator function should be supplied either exogenously or endogenously from the tissue. For example, transient expression assays with maize C1 or petunia AN2 were only effective when the corresponding bHLH partners or functional equivalents were supplied (Sainz et al., 1997a; Quattrocchio et al., 1998).

Assays to determine if the putative regulator induces anthocyanin biosynthesis is a common test for functionality. However, caution must be applied in interpreting negative results from these assays, as several unknown endogenous factors or interaction dynamics may interfere with the assay. Experiments could also be set up with various promoter- reporter constructs to assay for the ability of the transcription factors to recognise and activate specific target promoters (Sablowski et al., 1994; Quattrocchio et al., 1998; Schwinn, 1999). It is important to bear in mind that the ability to recognise and bind to a given promoter does not always mean this is the native target of the Myb or that it regulates the process for which the structural gene is involved (Sablowski et al., 1994).

1.12 INTRODUCING NOVEL COLOURS

Many ornamental species have a limited colour range or lack specific colours such as blue or yellow. Introducing novel colours is a key objective of molecular biology research into flower colour development. The extensive interspecific crosses that have occurred in anthurium over hundreds of years have produced numerous shades of red- and orange-coloured cultivars. Conspicuously absent from the colour hue in anthurium are blues, yellows and patterned spathe types (although some instances of the latter have been recorded (Kamemoto and Kuehnle, 1996). An ever-increasing array of strategies for colour modification has been developed (reviewed in Davies and Schwinn, 1997). While a detailed overview of these strategies is outside the scope of this thesis, the following sections are a selection of some key approaches.

1.12.1 Targeting specific genes

Each of the structural anthocyanin genes is a potential target for genetic manipulation. They could be introduced to overcome mutational blocks in the pathway or substrate specificities, or they can be silenced using a range of sense and antisense technology (van der Krol et al., 1988, 1990; Elomaa et al., 1993; Courtney-Gutterson et al., 1994; Deroles et al., 1998). F3'5'H has been a major target of molecular flower breeding efforts in the hope of introducing delphinidins, the dihydroflavonol precursor for the highly desirable blue phenotype. However, the presence of delphinidin-derived anthocyanins is not always concomitant with blue coloured flowers (Forkmann, 1991). *Tulipia sp, Impatiens sp, Cyclamen sp, Eustoma sp* and *Pelargonium sp* all produce delphinidin derivatives, yet lack true blue colours. While there is a high correlation between blue flower colour and the presence of delphinidin-derived anthocyanins (Heller and Forkmann, 1993), successful generation of novel blue colours depend on other factors such as self association, intra and intermolecular copigmentation, association with metal ions and high vacuolar pH (Goto and Kondo, 1991; Brouillard and Dangles, 1993; Holton and Tanaka 1994).

According to Davies and Schwinn (1997) several challenges are encountered in trying to develop blue colours in ornamentals. Firstly, several traits have to manipulated

simultaneously which can be complicated. F3'5'H has to be added such that sufficient quantities of dihydroquercetin (DHM) are produced which may require F3'H activity to be silenced. Then, a suitable DFR that can catalyse DHM has to be present and the final anthocyanin molecule requires a specific vacuolar pH and copigmentation patterns for true blues to develop. Secondly, the unavailability of cDNA clones for *C*-glycosylation enzymes limits the use that could be made of them to provide copigments. Thirdly, vacuolar pH is closely linked to the stable expression of blue colours (Hellar and Forkmann, 1993). This is highlighted in pelargonium, which has delphinidin-derived derivatives but low vacuolar pH and therefore lacks true blues (Davies and Schwinn, 1997).

1.12.2 Modifying genes involved in copigment synthesis

The cloning of genes involved in the synthesis of non-anthocyanin flavonoids provides another means by which colour modification can be achieved other than targeting genes involved in the primary reaction of the flavonoid pathway. This approach takes advantage n of the copigmentation role of flavonols. An antisense strategy to eliminate flavonol synthase (FLS) activity in petunia and *Nicotiana tabacum* (tobacco) led to a decrease in flower colour and an increase in anthocyanin resulting in altered flower colour possibly by changing anthocyanin copigment complexes (Holton et al., 1993).

In addition, the flavones, especially the *C*-glycosides, seem particularly effective as copigments for generating blue colours. In vitro experiments showed that the existing colour range of lisianthus could be extended from shades of pink and purple to blues by substituting the usual flavonol copigments with flavone *C*-glycosides, swertisin and isoorietin, in a solution of lisianthus delphinidin (Asen et al., 1986). In fact, anthocyanins in blue flowers are always found in association with flavones (Harborne, 1976). Therefore, copigmentation with flavones appears key to the blue pigment.

1.12.3 Manipulating the pathway by modifying regulatory genes

Since it is the expression pattern of specific transcription factors that control the activity of flavonoid biosynthetic genes, then modifying regulatory genes may introduce spatial and temporal novelty into anthocyanin production. When the maize regulatory gene *(R)* was expressed in arabidopsis under the control of the constitutive 35S promoter, transformants were obtained that had increased anthocyanin levels (Lloyd et al., 1992). No phenotypic changes occurred when the maize C1 was similarly expressed. In tobacco however, constitutive expression of both C1 and Lc in combination, produced marked increases in anthocyanin levels and novel distribution patterns (Lloyd et al., 1992).

Similarly, petunia plants transformed with *Lc* cDNA driven by the CaMV promoter showed an up-regulation in anthocyanin biosynthesis (Bradley et al., 1998, 1999). Expression of antirrhinum *Roseal* in two heterologous hosts (petunia and lisianthus) resulted in altered pigmentation patterns, providing the first evidence that such changes could be induced solely by a Myb protein (Schwinn, 1999). The use of regulatory proteins to introduce new colour patterns and novel colours is developing as a viable molecular approach for flower colour manipulation.

In summary, while the anthocyanins may be the principal colour pigments, the final colour developed is determined by a complex of factors ranging from pH to specific complexes with metal ions. All these factors come into play in various combinations when attempting to manipulate flower colour by genetic engineering. Therefore, such an effort cannot be divorced from a thorough understanding of the molecular and biochemical aspects of pigment production in the targeted tissue.

1.13 THESIS AIMS

The initiative for molecular biology research in anthurium was born out of industry demands in the West Indies to expand the range of flower colours available to the local growers, thereby increasing their competitiveness in the international market. Using a

molecular approach, two possible routes to introduce novel phenotypes are to target structural gene or regulatory gene activity. Both alternatives are most effective with genes cloned from the plant of interest. Therefore, the first aim of this study was to construct and screen a cDNA library for structural and regulatory clones involved in anthocyanin biosynthesis in anthurium.

Anthurium is an interesting model for anthocyanin gene regulation studies. In contrast to the traditional ornamental models studied, anthurium is a monocot, with a modified leaf as the flower and an extended flower development period of eight-nine weeks. In this regard, the second aim of this work was to determine if these developmental features were associated with an expression pattern of the anthocyanin biosynthetic genes that might lead to the formation of a unique regulatory model for anthurium spathe tissue. Also, given the very different biochemical and developmental profiles between the spadix and the spathe, a comparison of the expression and regulatory patterns of the biosynthetic genes in these tissues would determine if the regulatory model developed for the spathe tissue was applicable to the spadix.

Although preliminary genetic data suggest that two genes are involved in controlling anthocyanin biosynthesis in the anthurium spathe (Kamemoto et al., 1988), there have been no documented attempts to identify and characterise the two genes at the molecular level. An aim, therefore, was to perform a molecular analysis with genotypically defined, white anthurium lines to test the genetic model. Such information would be a significant step towards understanding flower colour inheritance in anthurium.

A significant objective of the study was to gain insight into the regulation of anthocyanin pigment production in anthurium by isolating cDNA clone(s) for regulatory genes of the Myb-protein type and performing specific experiments to determine functionality. Furthermore, it would be useful information to determine whether antirrhinum anthocyanin Myb regulatory proteins could activate anthurium anthocyanin structural gene promoters.

Chapter 2

General material and methods

2.1 PLANT MATERIAL

The anthurium cultivars used in this study were obtained from Rainbow Park Nurseries in Auckland and the University of Hawaii anthurium collection, and their general characteristics which are of significance for this research are represented in Table 2.1 Mature flowering cultivars of red, white, orange and pink anthurium were established at the Plant Growth Unit, Massey University. They were grown in a long-term medium of fines 50%, fibre 30% and pumice 20%. Also included in the mixture were dolomite, lime and gypsum, each at 2 kg m^{-3} and slow release fertiliser, Osmocote (Scotts Company, USA) at 200 g m^{-3}. Plants were given a weekly feed of Peters Professional water-soluble fertiliser (Scotts-Sierra Horticultural Products, USA), having an NPK ratio of 20/8.7/16.6 and containing trace elements. Light measurements averaged 18000-20000 $\mu E\ m^{-2}\ s^{-1}$ between April and October each year while a 70% shade cloth was employed during the summer months. The genotypically defined Hawaii lines, used in experiments discussed in Chapter 5, were not established as plants at the Massey greenhouse. They were only used for RNA extraction and this was performed in Hawaii.

Table 2.1 Anthurium cultivars and genetic lines used in this research

Cultivar/line	Source	Spathe Colour	Spadix Colour	Genetics
Altar	R.P.N	Red	Yellow	Unknown
Atlanta	R.P.N	Orange	Yellow	Unknown
Lido	R.P.N	Pink	Orange	Unknown
Acropolis	R.P.N	White	Pink	Unknown
Montana	R.P.N	Light pink	Pink/purple	Unknown
1244	U.of H	White	Pink	*oom-*
1349	U.of H	White	Pink	*oom-*
1250	U.of H	White	Pink	*oom-*

R.P.N., Rainbow Park Nurseries; U.of H., University of Hawaii

Flower development was divided into six stages based on developmental characteristics (Figure 2.1). The stages chosen had to be consistent across the cultivars being used in the study. Since acyanic cultivars were included in the work, stages defined by criteria such as

Figure 2.1. The six developmental stages of the anthurium flower. The development of the flower was divided into six stages that were consistent across the cultivars used in the research. The cultivar Altar is shown. The stages are defined as follows:

Stage 1: Flower not yet emerged

Stage 2: Flower is first visible

Stage 3: Flower separates protrudes from the sheath separating from the base of the leaf.

Stage 4: Flower peduncle elongates

Stage 5: The spathe is half unfolded

Stage 6: The spathe is fully expanded (newly open)


1
2
3
4
5
6

the first appearance of colour could not be used. The flower of all cultivars had a consistent growth pattern, from the time of emergence to full expansion of the flower. Therefore, the growth pattern of the flower formed the basis upon which the six stages were chosen.

2.2 RNA EXTRACTION PROCEDURES

High quality intact RNA was a prerequisite to successfully constructing a cDNA library for anthurium. RNA was also needed for northern analysis and for making first stand cDNA in PCR experiments. Different RNA extraction protocols were used to isolate RNA from spathe and spadix tissue.

2.2.1 RNA extraction from spathe tissue

Spathe tissue from each of the six developmental stages of the flower (Figure 2.1) was harvested from the plant, frozen in liquid N_2 and then used immediately or stored at -80 °C until required.

Total RNA was isolated by a modification of the method described by Prescott and Martin, (1987). The RNA extraction buffer consisting of 5% (w/v) SDS, 0.5 mM EDTA, 150 mM LiCl and 5 mM Tris HCl (pH 9) was supplemented with 0.45 g PVPP and 450 µL of 2β ME to prevent oxidation of the plant phenolics. A mortar and pestle, pre-cooled using liquid N_2, was used to grind frozen tissue (0.5 g per sample) to a fine powder. The ground tissue was added to 15 mL of the RNA extraction buffer in a Nalgene 50 mL tube. This was followed by vigorous shaking and centrifugation at 10,000 rpm for 15 min in a Sorvall SS-34 rotor.

The supernatant was removed to a new tube and was spun again to remove any traces of plant tissue/debris. The supernatant was again transferred to a new tube and extraction buffer added to bring the volume back to 15 mL. Three chloroform (without phenol or isoamyl alcohol) washes were then performed, each involving the addition of an equal volume of chloroform, thorough mixing and centrifugation as before. This step may have helped remove cuticular waxes, of which anthurium spathe has a lot.

The supernatant was then transferred to a new Nalgene tube and washed with an equal volume of phenol:chloroform:isoamyl alcohol (24:24:1 v/v), thoroughly mixed and centrifuged. The upper phase was carefully separated and placed in a sterile 50 mL Corex tube. Two rounds of overnight nucleic acid precipitation were performed. The first was at -20 °C using an equal volume of isopropanol and 1/10 volume of 3 M NaAc. This precipitated both RNA and DNA from the aqueous solution. The nucleic acids were pelleted by a 30 min centrifugation at 4 °C and 10,000 rpm in an SS-34 rotor. The supernatant was discarded and the pellet allowed to air dry for 8-10 min before resuspending in water. The second precipitation was a selective RNA precipitation by the addition of 8 M LiCl to the solution, to a final concentration of 2 M in a sterile microcentrifuge tube. The RNA was pelleted by centrifugation at 4 °C at maximum speed of 14,000 rpm in a microcentrifuge. The RNA pellet was washed twice with 70% (v/v) ethanol and then dissolved in 40 μL of sterile water.

When the protocol by Prescott and Martin (1987) was first tried without any modification, the RNA yields were low and the quality inferior, with an $A_{260/280}$ absorbance ratio less than 1.8. In addition, large quantities of genomic DNA co-precipitated with the RNA. However, a number of improvements were made to the protocol for use with anthurium flowers.

Anthurium flowers are very waxy and appear to have high polyphenolic levels as evidenced by the quick browning of the tissues when cut. Tissue polyphenolics are known to undergo rapid oxidation, forming complexes with proteins, polysaccharides, alkaloids and anthocyanins that interfere with RNA extraction (Loomis, 1974; Salzman et al., 1999). Assuming this to be the case with anthurium tissues, the addition of 2β ME, a sulfhydryl reducing agent, may have inhibited the activity of oxidase enzymes in the tissue and prevented the oxidation of plant phenolics and may have also served to inactivate ribonucleases by breaking disulphide linkages (Dawson et al., 1986; Venkatachalam et al., 1999). The extra chloroform washes coupled with increased extraction buffer may have diluted out the waxes, while the PVPP and the two clearance centrifugation steps included in the protocol helped to remove the tissue debris.

These changes to the protocol led to intact RNA with distinct 23 S and 16 S RNA bands. In addition, total RNA yields increased to 200-350 μg g^{-1} of tissue with $A_{260/280}$ ratios between 1.8 and 2.0 being consistently obtained from tissues at all stages. Also negligible amounts of genomic DNA coprecipitated.

2.2.2 RNA extraction from spadix tissue

Initial attempts to isolate RNA from spadix tissue using the RNA extraction protocol developed for the spathe proved unsuccessful. The true flowers are located on the spadix and as the stigma matures large quantities of a sticky substance (possibly polysaccharide) are produced to trap pollen grains. This material may have interfered with the RNA extraction as only low yields of poor quality RNA were obtained from young spadix tissue (Stages 1 to 3 refer to Figure 2.1). Furthermore, both yield and quality became worse when older tissue was used (Stages 4 to 6). Evidently this protocol could not efficiently separate RNA from tissues with such high levels of polysaccharides. The hot borate method of RNA isolation developed by Wilkins and Smart (1996) was then tried. This method is good for isolating high quality RNA from recalcitrant plant species containing copious levels of phenolics, polysaccharides or other secondary metabolites. The protocol uses proteinase K which together with the borate extraction buffer, enhances dissolution of the cell wall, prohibits the oxidation of endogenous phenolics and their binding to RNA, inhibits ribonuclease activity and denatures protein. RNA is then separated from all other cellular constituents and purified by selective precipitation in the presence of LiCl, followed by a precipitation step in KAc for the removal of positively charged polysaccharides.

Solutions

Extraction buffer

0.2 M Na borate decahydrate

30 mM EGTA

2% (w/v) SDS

These were dissolved in sterile water, the pH adjusted to 9.0 with 5 M NaOH and then autoclaved for 40 min at 121°C at 216 Kpa

Other reagents used during the protocol:

DTT to 10 mM (Stored at -20 °C)
Nonidet P-40 to 1.0% (v/v)
PVP to 2% (w/v)
Proteinase K (20 mg mL^{-1} stored at -20 °C)
TE Buffer [10 mM Tris, (pH 8.0); 1 mM EDTA, pH 8.0]
10 mM Tris, pH 7.5 with conc. HCl
2 M KCl
2 M KAc, (pH 5.5 with conc. HCl)
2 M and 8 M LiCl, ice-cold
3 M NaAc pH 6.0 (with glacial CH_3COOH)
100% (v/v) and 70% (v/v) EtOH ice-cold

At the start of the procedure, a water bath or rotary oven was set at 42 °C. A stock solution consisting of 4 mL of extraction buffer (per 0.5 g of plant tissue) plus 2% (w/v) PVP, 1% (v/v) Nonidet P-40 and 10 mM DTT was prepared, thoroughly mixed and left at 85 °C until required. In the interim an aliquot of 100 µL proteinase K was added to each Nalgene screw capped sample tube and left on ice. The frozen tissue (0.5 g per sample) was then ground to fine powder in a mortar with a pestle that was pre-cooled with liquid N_2. The next steps were crucial and had to be conducted quickly to avoid RNA degradation.

The ground tissue was transferred to another mortar and pestle that was at room temperature. Immediately, 2.5 mL of the extraction buffer (at 85 °C) was added and the powder was further pulverised for 10-15 s and then transferred to Nalgene tubes containing the proteinase K, and mixed well. The mortar and pestle were rinsed with a further 1 mL of extraction buffer and this was added to the homogenate in the Nalgene tubes. Tubes were then placed at 42 °C to incubate for 1.5 h with constant mixing. Subsequently, 280 µL of 2 M KCl was added to each tube, the contents mixed gently and the tubes placed on ice for 1 h.

Following this, homogenates were centrifuged at 14,000 rpm (using a Sorval SS-34 rotor) at 4 °C for 20 min to remove plant debris and precipitated material. The supernatant was removed and centrifuged again at the same speed and temperature for 10 min to remove any remaining precipitate. The first selective precipitation of RNA was then performed. The supernatant was first transferred to a 15 mL RNase free glass Corex tube and 1/3 volume of 8 M LiCl was added to give a final concentration 2 M. The tubes were sealed and inverted to mix the contents then left overnight at 4 °C.

After precipitation, the RNA was pelleted by centrifugation, again at 14,000 rpm (using a Sorval SS-34 rotor) for 20 min at 4 °C. The supernatant was discarded carefully without disturbing the pellet and the pellet taken through a series of washes each using 3 mL of ice cold 2 M LiCl. For each wash, once the LiCl was added, the pellet was detached from the wall of the tube and swirled in the solution then centrifuged as before for 10 min at 4 °C. The supernatant from each wash was discarded prior to adding a fresh aliquot of 2 M LiCl.

After the last cycle of washing and centrifugation, 2 mL of 10 mM Tris (pH 7.5) was added to dissolve the pellet. As the pellets never completely dissolved, an aliquot of 200 µl 2 M KAc (pH 5.5) was added to remove contaminating carbohydrates. After a gentle mixing and incubation on ice for 15 min, the tubes were again centrifuged at 14,000 rpm for 10 min at 4 °C to pellet the precipitated carbohydrates. The supernatant was then transferred to a sterile 15 mL Corex tube and 2.5 volumes of 100% (v/v) ethanol were added. The tubes were sealed with parafilm so they could be vigorously mixed then placed at -20 °C overnight. On occasion, tubes were placed at -80 °C for 2 h. This gave slightly lower yields than the precipitation at -20 °C but did not compromise the quality of the RNA.

To pellet the RNA, tubes were centrifuged again at 14,000 rpm using a Sorval SS-34 rotor at 4 °C for 30 min, after which the ethanol was removed leaving behind a relatively clear pellet on the sides of the tubes. An aliquot of 3-5 mL of 70% (v/v) ethanol was used to gently wash the RNA. Samples were then centrifuged as described above but for 5 min.

The ethanol was discarded, the tubes inverted and the RNA pellets were air dried for a maximum of 10 min. Pellets were dissolved in 200 μL of sterile water and then transferred to sterile microcentrifuge tubes.

2.3 ANALYSING THE ISOLATED RNA

The purity and quantity of RNA was determined by spectrophotometry while gel electrophoresis was used to determine whether the isolated RNA was degraded or intact.

2.3.1 Quantifying RNA by spectrophotometry

An aliquot of the RNA preparation was diluted in 750 μL of sterile water. Using clean quartz cuvettes and sterile water as the standard, the base line for the spectrophotometer was determined. The absorbance of the diluted preparations, were then determined from 200 nm to 320 nm. Two absorbance ratios were then used to evaluate the purity of nucleic acids. The ratio of absorbance at 260 nm and 280 nm was used to measure the degree of contamination of the samples with proteins and phenol. The 200 nm to 320 nm ratio was used to indicate contamination by carbohydrate. The concentration of RNA in the preparations was calculated on the basis that an A_{260} of 1.0 is equivalent to 40 μg mL^{-1} of single stranded RNA.

2.3.2 Checking for RNA degradation by agarose gel electrophoresis

The gel apparatus was soaked in 2% (v/v) 'Absolve' solution for least 2 h to inactivate any contaminating RNase. 1 μL of RNA was mixed with 2 μL of 10 x loading dye and separated on a 1% (w/v) agarose gel containing 10 μg mL^{-1} EtBr.

2.4 PURIFYING MESSENGER RNA

Total RNA (1 mg) was taken through two rounds of purification using Amersham Pharmacia Biotech mRNA Purification Kit (Amersham Pharmacia Biotech, England) to obtain poly $(A)^+$ RNA used in making the cDNA library described in subsequent sections.

In addition, poly $(A)^+$ RNA (5 μg) was required when measuring transcript levels for Myb genes as these tend to be expressed at very low levels within the cell. For this, only one round of purification was performed.

The key feature of this procedure was the use of spin columns for affinity purification of poly $(A)^+$ RNA on oligo(dT)-cellulose. Total RNA was first denatured, cooled, adjusted to a salt concentration of 0.5 M NaCl using the sample buffer provided, and then added to one of the pre-packed columns. Poly $(A)^+$ RNA was selectively bound to the oligo(dT)-cellulose and unbound ribosomal RNA was removed by two washes. The first wash was with a high salt buffer and the second with a low salt buffer. The poly $(A)^+$ RNA was then recovered by elution with warm elution buffer and subjected to another round of purification. The purification procedure effectively removed the ribosomal RNAs along with the small amounts of genomic DNA that were present. The yield of poly $(A)^+$ RNA was measured with a spectrophotometer as described in Section 2.3.1 and an aliquot tested on a 1% (w/v) agarose gel as described in Section 2.3.2. The poly $(A)^+$ RNA was visible as a smear on the agarose gel (data not shown).

2.5 ISOLATION OF cDNA CLONES FOR ANTHURIUM FLAVONOID BIOSYNTHETIC GENES

Isolating cDNAs or genomic clones for anthocyanin biosynthesis in anthurium flowers is critical to understanding and manipulating flower colour using molecular tools. A cDNA library contains cDNAs that represent a subset of expressed genes. The library was constructed from poly $(A)^+$ RNA isolated from the plant tissue where the genes of interest will most likely be expressed, in this case the spathe.

With this objective in mind, an anthurium cultivar that expressed the full complement of the known flavonoid biosynthetic genes present in *A. andraeanum* was chosen as the primary source of tissue for the library. Biochemical studies by Iwata et al. (1979, 1985) showed that red anthurium has both pelargonidin and cyanidin-derived anthocyanins. Therefore, an anthurium cultivar producing cyanidin-derived anthocyanins would have all the genes from *CHS* through to *GST* including *F3'H*. On the contrary, some orange

coloured anthurium cultivars have only pelargonidin derived anthocyanins and this may be due to the absence of a functional *F3'H* activity. *A. amnicola* is one species that produces peonidin derived anthocyanins (Marutani et al., 1987). Therefore, in addition to *F3'H*, the presence of peonidin suggests the activity of a methyltransferase. However, because this cultivar was not readily available, spathe tissue from a red coloured anthurium line, Altar, was used as the source of poly $(A)^+$ RNA for the construction of the cDNA library.

The cDNA library was constructed from poly $(A)^+$ RNA isolated from total RNA from each of the six stages of the spathe tissue of Altar by the method described in Section 2.2.1. The library was constructed using a Stratagene UNI-ZAP cDNA synthesis kit (Stratagene Cloning Systems, USA). The procedure can be summarised into three main stages:

a) the conversion of poly $(A)^+$ RNA into double stranded cDNA with *Xho* I and *Eco* RI restriction sites at the 3' and 5' ends, respectively.
b) size fractionation
c) ligation of double stranded cDNA into the vector followed by packaging into λphage and titering of the library

2.5.1 Preparation of double stranded cDNA

Double stranded cDNA was made in a two-step reaction comprising of first strand synthesis and second strand synthesis. Maloney murine leukemia virus reverse transcriptase (MMLV-RT) was used to catalyse the synthesis of first strand cDNA from poly $(A)^+$ RNA in a final reaction volume of 50 µL. The reagents used included 5 µL of 10 x first strand buffer, 3 µL methyl-nucleotide mixture (10 mM dATP, dGTP, dTTP plus 5 mM 5-methyl dCTP), 2 µL of oligo(dT_{18}) linker-primer, I µL of RNase block ribonuclease inhibitor (40 U μL^{-1}), 5 µL anthurium poly $(A)^+$ RNA and 1.5 µL MMLV-RT (50 U μL^{-1}). The linker-primer attached to the oligo(dT_{18}), was designed to engineer a *Xho* I restriction site to the 3' end of the cDNA fragments. The analogue, 5-methyl dCTP replaced the normal dCTP in the reverse transcriptase reaction, to hemi-methylate the cDNA thereby protecting it from restriction enzymes (*Xho* I) used in subsequent cloning steps.

The procedure was carried out simultaneously with 'test' poly $(A)^+$ RNA provided with the kit as a control. Once the MMLV-RT was added to the reaction with anthurium poly $(A)^+$ RNA, a 0.5 μL aliquot was then removed and radiolabeled with [α-^{32}P]dCTP (0.37 MBq μL^{-1}) as the first strand synthesis control reaction.

All three reactions were incubated at 37 °C for 1 h. Subsequently, the radioactive first strand reaction was stored at -20 °C while the other two reactions were taken through to second strand synthesis.

Two enzymes are key for synthesis of the second strand. The first, RNase H, was used to nick the RNA strand bound to the first strand cDNA. The fragments created served as primers for the second enzyme DNA polymerase I, which synthesised the second strand by nick translation. The reaction was performed with 20 μL of 10 x second-strand buffer, 6 μL of second strand dNTP mixture (10 mM dATP, dGTP, dTTP plus 26 mM dCTP), 114 μL sterile water, 2 μL [α-^{32}P]dCTP (0.37MBq μL^{-1}), 2 μL of RNase H (1.5 U μL^{-1}) and 11 μL DNA polymerase I (9.0 μL^{-1}).

The second-strand nucleotide mixture has been supplemented with dCTP, to reduce the probability of 5-methyl dCTP becoming incorporated in the second strand. The reactions were incubated at 16 °C for 2.5 h then immediately placed on ice.

The enzyme, *Pfu* DNA polymerase (2 μL, 2.5 U μL^{-1}), in the presence of blunting dNTP mix (23 μL), was added to the entire second strand reaction to remove overhangs formed at the ends by DNA polymerase I, creating blunt ended double stranded cDNA. The reaction was incubated at 72 °C for 30 min.

To remove all proteins from the reaction, an equal volume (200 μL) of phenol:chloroform:isoamyl alcohol (24:24:1 v/v), was added to the reaction. The mixture was vortexed and centrifuged at 14,000 rpm in a microcentrifuge to separate out the aqueous phase from the proteinaceous phase. The former was transferred to a fresh microcentrifuge tube to which an equal volume of chloroform was added to remove any

residual phenol. This mixture was vortexed then centrifuged at 14,000 rpm in a microcentrifuge and again the upper aqueous phase removed and transferred to sterile microcentrifuge tube. The double stranded, blunt ended cDNA was then precipitated overnight at -20 °C with 1/10 volume 3 M NaAc and 2.5 volumes of 100% (v/v) ethanol.

The cDNA was pelleted by centrifugation at 14,000 rpm for 60 min at 4 °C in a microcentrifuge. The supernatant was removed and the pellet washed with 70% (v/v) ethanol. As the pellet was very small and barely visible with the naked eye, a hand-held Geiger counter was used to monitor its location during the 70% (v/v) alcohol washes. After allowing the pellet to air dry for 15 min it was resuspended in 9 μL of *Eco* RI adapters solution and incubated at 4 °C for 60 min, to ensure the entire pellet went into solution. An aliquot of 1 μL of this solution was taken out and, along with the first strand control, was resolved on an alkaline agarose gel. The gel was subsequently capillary-blotted to a nylon membrane then used for autoradiography.

Resolving an aliquot of the first and second strand reactions with an alkaline agarose gel was necessary to check that adequate cDNA synthesis had occurred and that cDNA of an appropriate size was formed. Figure 2.2 shows the autoradiograph of cDNA from the first and second strand synthesis reactions, exposed for different times under different conditions.

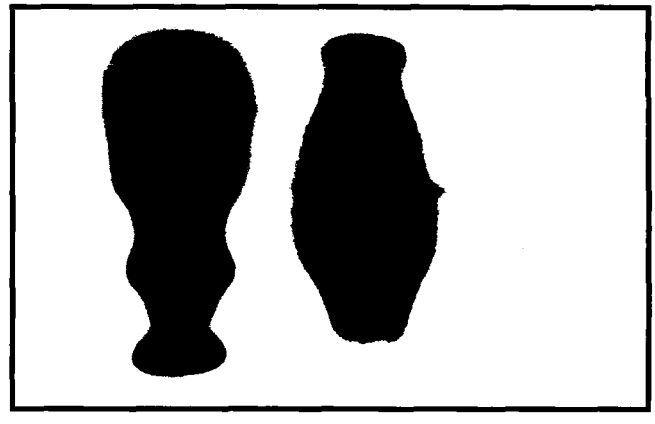
Once synthesis of both strands was confirmed, T4 DNA ligase (1 μL, 4 U μL^{-1}) was used to perform blunt end ligation of the *Eco* R1 adapters to both blunt ends of the double stranded cDNA in the presence of 10 x ligase buffer (1 μL) and 10 mM rATP (2 μL). The reaction was incubated for two days at 4 °C. The *Eco* RI adapters (supplied with the Stratagene kit) consisted of a 9- and 13-mer that are complementary to each other with an *Eco* RI cohesive end. The 9-mer oligonucleotide was phosphorylated to allow ligation with other blunt end termini such as the cDNA and other adapters. The 13-mer oligonucleotide was kept de-phosphorylated to prevent it from ligating to other cohesive termini.

At the end of the two-day period, the T4 DNA ligase was heat inactivated by incubating the

Figure 2.2. Autoradiograph of double stranded cDNA. Film was exposed directly to gel and then developed overnight at -80 °C. Size of the λ (Lambda) *Hin*d III DNA markers is shown on the left in bp. An exposure time of 20 min at room temperature was sufficient for detecting first strand synthesis. However, the second strand cDNA was only visible after an overnight exposure at -80°C. This is because the first strand reaction had 10 fold more radioactivity than the second strand reaction. Therefore, the film required a longer exposure time for the double stranded cDNA to be visible.

Lambda *Hin*dIII markers First strand cDNA Second strand cDNA

bp



4361

2322

2027

564

reaction at 70 °C for 30 min. The *Eco* RI ends of the 13-mer oligonucleotide were then phosphorylated using T4 polynucleotide kinase (1 μL, 10 U μL^{-1}). The reaction was supplemented with 10 x ligase buffer (1 μL) and 10 mM rATP (2 μL) and sterile water (6 μL) and incubated at 37 °C for 30 min. This was necessary to facilitate cloning into the vector arms.

The kinase was then heat killed in the same way as the T4 DNA ligase. The reaction was then subjected to *Xho* I digestion to release the *Eco* RI adapters and residual linker primers from the 3' end of the cDNA. The reaction was set up with 28 μL of *Xho* I buffer supplement and 3 μL *Xho* I at 40 U μL^{-1}.

2.5.2 Size fractionation

This was a crucial part of the library construction procedure, as the cDNA could have been lost if air bubbles were present in the column. Unattached adapters and small cDNA fragments were separated from the cDNA of the desired size range by passage through a drip column containing the gel filtration medium Sepharose CL-2B. The column was set up and packed with resin as instructed in the Stratagene protocol. The cDNA, containing a marker dye, was gently added to the surface of the resin bed. The leading edge of the dye, along with a hand held Geiger counter, was used to trace the movement of the cDNA through the column.

A pooled sample was collected from the time the leading edge of the dye reached the '-0.4' mL mark on the Stratagene resin column, and subsequently, individual samples were taken (three drops per sample) until the trailing edge of the dye reached the 0.3 mL mark on the column. These fractions were expected to contain the cDNA clones greater in size than 400 bp but not the small cDNA clones, adaptors and unincorporated nucleotides.

An aliquot of 8 μl of each collected fraction (the pooled and the individual samples) was analysed by electrophoresis on a 5% (w/v) agarose gel. When viewed, distinct smears, corresponding to different sizes of cDNA fragments were visible suggesting the

fractionation procedure was effective (data not shown). One sample (Sample 1) had cDNA fragments ranging in size up from 0.65 kb with cDNA visible at 4 kb. This fraction was chosen for cDNA library construction, as taking fractions with cDNA less than 0.65 kb risked including small cDNA fragments or adapters.

The selected fraction was washed with an equal volume of phenol:chloroform:isoamyl alcohol (24:24:1 v/v), centrifuged at 14,000 rpm for 15 min in a microcentrifuge and the upper aqueous phase transferred to a fresh microcentrifuge and washed again with an equal volume of chloroform. The cDNA was then precipitated at -20 °C overnight with 2.5 volumes of 100% (v/v) ethanol and 1 μL of glycogen (20 mg mL^{-1}).

After overnight precipitation, samples were centrifuged at 14,000 rpm for 60 min at 4 °C to pellet the cDNA. The supernatant was removed and the pellet carefully washed with 80% (v/v) ethanol. Again, a hand held Geiger counter was used to keep track of the pellet. Having purified and precipitated the cDNA fragments in Sample 1, the cDNA pellet was air dried for 15 min and resuspended in 10 μL of sterile water. An aliquot of 1 μL of Sample 1 cDNA was resolved on a 1% (w/v) agarose gel. The smear of cDNA was of the appropriate size (data not shown) and ready for ligation into the cloning vector.

2.5.3 Ligation, packaging and titering

Ligation.

Only Sample 1 was chosen for ligation into the dephosphorylated arms of the Uni-ZAP XR cloning vector. The first reactions were set up with 0.6 μL and 1.4 μL of cDNA solution in a total reaction volume of 5 μL. A control reaction using the 'test' insert provided with the Stratagene Kit was simultaneously set up. T4 DNA ligase (0.5 μL, 4 U μL^{-1}) catalysed both reactions in the presence of 10 x ligase buffer (0.5 μL), 10 mM rATP (pH 7.5, 0.5 μL) and 1 μg of the Uni-ZAP XR vector (1 μg μL^{-1}). The reactions were incubated at 4 °C for 48 h to optimise the ligation.

Packaging.

Ligated cDNA was packaged into lambda (λ) phage using the Gigapack III Gold packaging extract supplied with the Stratagene kit. The packaging extract was removed from -80 °C storage and 20 μL, from the ligation, was added to the packaging extract just when the extract started to thaw. After a gentle mix and a brief spin, the tubes were left at room temperature for 2 h. Subsequently, 500 μL of SM buffer along with 20 μL of chloroform were added to each tube. Each packaging extract represented part of the primary library.

Titering

To titer the library, 1/100 and 1/1000 dilutions of each packaged ligation was made in SM buffer using a 10 μL aliquot of the primary library. The bacterial host used for plating out the library was XL1-Blue MRF', and stocks of this were prepared as directed in the Stratagene protocol and stored in 10 mM $MgSO_4$ for 1-2 week periods.

To allow the phage to attach to the bacteria, 10 μL of each dilution was incubated at 37 °C with 200 μL of XL1-Blue MRF' (A_{600} of 0.5). Subsequently, this was mixed with LTop agarose containing 250 μL X-Gal (50 mg mL^{-1}) and 10 μL IPTG (200 mg mL^{-1}) for blue/white selection of the recombinants. After a quick inversion, the tube contents were poured onto an agar plate, allowed to set, then inverted and incubated overnight at 37 °C.

The titer of this first packaging gave a total of 500 plaque-forming units (pfu) μL^{-1} giving a total of 250,000 pfu. Subsequently, several ligation and packaging reactions were performed to give a primary library that was in excess of 1million pfu.

2.5.4 Library plaque lifts

Aliquots from the primary library, consisting of approximately 20,000 plaques, were plated out in LTop agarose on 20 agar plates (1,000 plaques/plate) and incubated overnight at 37 °C. The following day, plates were removed and cooled at 4 °C for 1-2 h to prevent the layer of LTop agarose from peeling off during the plaque lifts.

The lifts were done in a laminar flow hood using Hybond-N^+ positively charged nylon membranes (Amersham Pharmacia Biotech, England). Membranes and plates were marked using a ballpoint pen to secure their orientation. Two lifts were done on each plate, with the first lift being left for 2 min and the second for 4 min. Once taken off the plates the membranes were placed on filter paper soaked with denaturing solution for 2 min. Subsequently, they were removed and placed for 5 min on another filter paper soaked with neutralisation solution. This was followed with two 2 x SSC washes, each for 5 min. The DNA was then fixed to the membranes using Hoefer UV-C 500 UV cross linker for 1 min at 70,000 μJ cm^{-2} UV.

Denaturing solution

1.5 M NaCl

0.5 M NaOH

Neutralising solution

1.5 M NaCl

0.5 M Tris-HCL pH 7.2

0.001 M EDTA

2.5.5 Radioactive probe preparation

DNA to be labeled was cut out of plasmid DNA using the appropriate restriction enzymes, electrophoretically separated and purified from 1% (w/v) agarose gels according to the procedures described in Section 2.7 and 2.10. Once the concentration of the DNA was determined by gel electrophoresis (Section 2.8.3), 60-100 ng of DNA was used to make [α-^{32}P]dCTP radiolabeled probes. Two protocols were used to incorporate radiolabeled dCTP into probes.

For the first protocol, probes were prepared using [α-^{32}P]dCTP and the Quickprime T7 kit (Amersham Pharmacia Biotech, England). A volume of insert solution (1-3 μL) corresponding to 100 ng of DNA insert was made up to 36 μL with sterile water in an microcentrifuge tube and denatured at 95 °C for 4 min. The tube was immediately placed on ice. Reagent buffer (10 μL), [α-^{32}P]dCTP (3 μL) and T7 DNA polymerase (1 μL) were

then added. The contents of the tube were then carefully mixed, pulsed in a microcentrifuge and then incubated in a water bath at 37 °C for 15 min.
The second protocol was with *Redi*prime II random prime labeling (Amersham Pharmacia Biotech, England). The system provides individually dispensed reaction mixes that are dried in the presence of a stabiliser and a dye. The reaction mixes have been formulated using an improved exonuclease-free Klenow to give probes with high specific activity after shorter incubation times. DNA was diluted to a concentration of 2.5-25 ng in 45 µL of TE buffer (10 mM Tris HCl, pH 8.0, 1 mM EDTA) then denatured by heating for 5 min at 95 °C, followed by incubation on ice for 5 min. The denatured DNA was added to the reaction tube along with 5 µL of [α-^{32}P]dCTP (0.37 MBq µL^{-1}). The content of the tubes were mixed by pipetting and incubated at 37 °C for 10 min.

Subsequent steps were common to both kits. To separate the unincorporated nucleotides from labeled insert, the contents of the microcentrifuge tube were transferred to a ProbeQuant G-50 Micro Column (Amersham Pharmacia Biotech, England) (after the buffer in the micro-column was removed by centrifugation for 3 min). The column with the radiolabeled probe solution was placed in an empty sterile microcentrifuge tube and spun for 3 min at 3,000 rpm. At the end of the spin, the column was discarded and the solution containing the radiolabeled DNA was used for probing the membranes.

2.5.6 Screening the library

The library was screened using heterologous full-length cDNA clones for anthocyanin biosynthetic genes as probes. This was the preferred approach for securing anthurium cDNA clones as a number of cDNAs for flavonoid biosynthetic genes were already cloned from various plant species and available in our laboratory.

Once plaque lifts were completed and the cDNA crossed linked to the membranes, they were then prehybridised in Church and Gilbert solution (Church and Gilbert, 1984) for 2 h at 55 °C. At the end of the prehybridisation period, the solution was discarded and fresh Church and Gilbert solution added. The probes were then denatured by incubating at 95 °C

for 4 min and then added to the solution with the membranes. Hybridisation was performed at 55 °C overnight.

Solutions

Stocks of 20 x SSC and 10% (w/v) SDS were used to make lower dilutions used in washing of membranes.

20 x SSC Stock:

NaCl	3.0 M
Tri sodium citrate	0.3 M

Church and Gilbert solution (Church and Gilbert, 1984)

Solution 1: 0.5 M $NaH_2P0_4.2H_2O$

Solution 2 : 0.5 M Na_2HPO_4

The pH of Solution 2 was adjusted to 7.4 with Solution 1. Then for every 100 mL final volume 200 µl EDTA (pH 8 with conc. HCl) and 7 g solid SDS were added. The solution was left to mix with heating to help dissolve the SDS. Once the solution was completely cleared it was sterilised using a 0.45 µm filter disk.

Lower stringency conditions were used for washing as this was a screen using heterologous probes. Membranes were given two 15 min washes at 55 °C, first with 3 x SSC (w/v)/ 0.1% (w/v) SDS solution then in 2 x SSC (w/v)/ 0.1% (w/v) SDS. Membranes were then subjected to autoradiography at -80 °C using X-ray film.

Large cores of the plaques showing the strongest hybridisation signal (positives) from this primary screen were taken from the plates and placed in a solution of 500 µL SM buffer and 20 µL chloroform. This was vortexed to release the phage into solution and left at room temperature for 2 h then stored at 4 °C.

SM Buffer (per liter):

0.1 M NaCl

0.01 M $MgSO_4 . 7H_2O$

0.05 M Tris-HCl (pH 7.5)

0.01 % gelatin from a 2% (w/v) gelatin stock
Added sterile water to a final concentration of 1 L

To perform secondary screens, 10^{-3} dilutions were made of each positive plaque picked from the primary screen and 10-50 μL was plated out in XLI-Blue MRF' as previously described. Plaques were plated at a low density allowing single picks to be made. Plaque-lifts, prehybridisation and probing were performed as described previously. Some picks were taken through a tertiary screen, following the same protocol for the secondary screen.

2.5.7 In vivo excision of pBluescript phagemid from the Uni-Zap XR Vector

Once positives were obtained from the secondary screen and single plaque picks were made, an in vivo excision was performed as outlined in the Stratagene protocol. In vivo excision enabled the cloned insert to be recovered as a recircularised phagemid. The process is effected in the presence of filamentous (F1) bacteriophage-derived proteins. In the Uni-Zap XR vector, the sequence usually functioning as the origin of replication for the F1 phage has been divided into two parts: 1) the site of initiation and 2) the site for the termination of DNA synthesis. These two regions are genetically engineered to flank the entire pBluescript plasmid sequence and the cloned cDNA insert (when one is present). When the Uni-Zap XR vector and the F1 phage are both introduced into *E. coli* cells, the F1 phage proteins recognise and bind to the initiator region of its origin of replication, nick the DNA, thereby initiating DNA synthesis which proceeds from this point through to the termination signal at the 3' end, copying all the intervening sequences. The single stranded DNA molecule is circularised by the gene II product from the F1 phage, forming a circular DNA molecule containing the DNA between the initiator and terminator. In the case of the Uni Zap XR vector, this includes all the sequences of the pBluescript SK (-) phagemid and the cDNA insert, assuming that one is present.

The phage from a single plaque was released into a solution of 500 μL SM buffer and 20 μL chloroform. An aliquot of 250 μL of phage supernatant was incubated with 200 μL XL1-Blue MRF' cells (A_{600} of 1.0) and 1μL of EX-Assist helper phage at 37 °C for 15 min.

Subsequently 5 mL of L-Broth were added and the mixture subjected to vigorous shaking at 250 rpm at 37 °C for 3 h, or in some cases overnight. After incubation the tubes were centrifuged and the debris pelleted. A 1 mL aliquot of the supernatant, containing the excised phagemid, was stored as the in vivo stock in a sterile 1.5 mL microcentrifuge tube at 4 °C. The excised phagemid was replicated in SOLR cells (a non-suppressing *E. coli* strain which therefore prevents replication of the helper phage) to obtain single bacterial colonies. Aliquots of 200 µL SOLR cells (A_{600} of 1.0) were inoculated with of 1-10 µL of the in vivo stock and incubated for 15 min at 37 °C. An aliquot of this was then plated out on L-ampicillin plates (100 µg mL^{-1}) and incubated overnight at 37 °C to allow bacterial growth

For plasmid DNA isolation, a single colony was picked from the in vivo plates and cultured in L-Broth and ampicillin (100 µg mL^{-1}) overnight with shaking at 37 °C. Plasmid DNA was prepared by one of the techniques described in Section 2.6.1 to 2.6.4, and digested with one of the following pairs of restrictions enzymes to check for inserts: *Eco* RI / *Xho* I, *Xba* I / *Xho* I or *Apa* I / *Pst* I as described in Section 2.7.

2.5.8 Sequencing

Cloned DNAs were sequenced as a commercial service from the University of Auckland or the University of Waikato using the ABI automated DNA sequencer. The method employs the chain termination DNA sequencing principle with each of the four di-deoxynucleotides labeled with a unique fluorescent dye. Dye terminator reactions were then performed with either T3/T7 or M13 Reverse/M13 Forward primers of the pBluescript vector or with sequence specific primers. The ABI PRISM® 377 DNA Sequencer automatically analyses the DNA molecules labeled with the multiple fluorescent dyes. After samples are loaded onto the system's vertical gel, they undergo electrophoresis, laser detection, and computer analysis. Electrophoresis separation can be viewed on screen in real time and final data presented as a multicoloured trace with each nucleotide represented by a specific colour. The sequence data received were analysed using the DNASTAR (DNASTAR Inc. USA) suite of programmes.

2.6 PLASMID DNA PREPARATION

Plasmid DNA isolation was commonplace throughout the project to secure DNA for sequencing, making inserts for probing, preparing DNA for transient expression assays and making a variety of gene constructs. The method of plasmid DNA extraction varied according to the purpose for which the DNA was required and the quantity of DNA needed.

2.6.1. Plasmid miniprep by alkaline lysis

This method was followed without modification from that specified in Sambrook et al. (1989) and was often use to allow a quick check for the presence of inserts.

Reagents

Solution I:	50 mM glucose
	25 mm Tris (pH 8.0 with conc. HCl)
	10 mM EDTA (pH 8.0 with glacial CH_3COOH)
	Autoclaved at 121 °C for 20 min and stored at 4 °C.
Solution II:	0.2 M NaOH
	1 % (w/v) SDS
	Freshly prepared
Solution III	3 M KAc
	2 M acetic acid
	Autoclaved at 121 °C for 20 min and stored at 4 °C
LB Medium	10 g tryptone
	5 g yeast extract
	10 g NaCl
	Adjust pH to 7.5 then autoclave

To test for the presence of an insert, an individual *E. coli* colony was inoculated into 3 mL of L-Broth medium containing 100 μg mL^{-1} ampicillin and incubated overnight at 37 °C in orbital incubator set at 250 rpm. The culture was centrifuged at 4,000 g in a Jouan bench top centrifuge for 15 min to pellet the cells. The supernatant was removed and the tube inverted on a paper towels to allow the medium to drain completely from the pellet. The pellet was then resuspended in 200 μL of ice cold Solution I (see above) and then transferred to a 1.5 mL microcentrifuge tube and vortexed. To lyse the bacterial cells, 200 μL of freshly prepared Solution II was added, and then the tube was capped and gently inverted 4 to 6 times. Following this, 200 μL of ice cold Solution III was added to precipitate contaminating cellular components. The bacterial chromosomal DNA gets bound up with the precipitate and is removed along with it. The tube was gently inverted and then centrifuged for 10 min at 14,000 rpm in a microcentrifuge. Vortexing was avoided to prevent shearing of the chromosomal DNA. The supernatant was removed and added to an microcentrifuge tube containing 600 μL of phenol:chloroform:isoamyl alcohol (24:24:1 v/v). This mixture was vortexed, centrifuged and 500 μL of the top aqueous layer transferred to a fresh tube and mixed with an equal volume of isopropanol by vortexing. In most cases, samples were centrifuged at 4 °C 14,000 rpm for 20 min, immediately after the addition of the alcohol. On occasion, samples were allowed to sit on ice for 15 min prior to centrifugation. Once the plasmid DNA was pelleted, the supernatant was removed by aspiration and the pellet washed twice with 75% (v/v) ethanol and either air dried or dried under vacuum. The plasmid DNA was then resuspended in 40 μL of sterile water containing DNase-free RNase (20 μg mL^{-1} final concentration).

In some cases the mini-prep protocol was modified to accommodate up to 10 mL of bacterial culture. The bacterial cells from the 10 mL culture were resuspended in 200 μL of Solution I and subsequent steps used 400 μL and 300 μL of Solutions II and III, respectively. After centrifugation, 700 μL of the supernatant was removed and transferred to a fresh tube containing an equal volume of phenol:chloroform:isoamyl alcohol (24:24:1 v/v). The mixture was vortexed and centrifuged as described previously. 600 μL of the top layer was removed and added to an equal volume of isopropanol in a clean microcentrifuge

tube in order to precipitate plasmid DNA.

2.6.2 QIAprep spin miniprep

Another miniprep procedure used frequently to isolate high quality plasmid DNA for sequencing was the QIAprep Spin Miniprep Kit Protocol (QIAGEN, Australia). The procedure is based on the modified alkaline lysis method of Birnboim and Doly (1979). Bacteria are lysed under alkaline conditions, and the lysate is subsequently neutralised and adjusted to high salt binding conditions, ready for purification on the QiAprep silica-gel membrane. The latter selectively binds DNA while RNA, cellular proteins, and cellular metabolites are not retained on the membrane but are washed through. Using 3 mL overnight cultures of *E. coli,* the protocol consistently gave total yields ranging from 5-15 μg of high quality plasmid DNA.

2.6.3 Alkaline lysis/PEG treatment DNA preparation

This is another approach that was also used to secure good yields of high quality plasmid DNA for sequencing.

Reagents

GTE buffer

50 mM glucose,

25 mM Tris-HCl (pH8.0 with conc. HCl)

10 mM EDTA (pH 8.0 with glacial CH_3COOH)

0.2 M NaOH/ 1% (w/v) SDS (freshly made)

3.0 M KAc (pH4.8)

RNase A (DNase-free) (10 mg mL^{-1})

Chloroform

Isopropanol 100% (v/v)

EtOH 70% (v/v)

4.0 M NaCl

13% (w/v) PEG8000 (sterilised by autoclaving)

Pelleted bacterial culture (1.5-4.5 mL) was resuspended in 200 μL of GTE buffer by pipetting up and down. Cells were lysed with freshly prepared 0.2 M NaOH/1% (w/v) SDS and neutralised with the addition of 300 μL of 3.0 M potassium acetate, pH 4.8. Cellular debris was removed by centrifugation at maximum speed for 10 min at room temperature and the supernatant transferred to a clean tube. RNase A (DNase-free) was added to a final concentration of 20 μg mL^{-1} and incubated at 37 °C for 20 min. The supernatant was then extracted twice, with 400 μL chloroform and the DNA precipitated with an equal volume of isopropanol. After pelleting the DNA by centrifugation at maximum speed at room temperature, the isopropanol was removed and the pellet washed with 70% (v/v) ethanol and then vacuum dried. Subsequently, the pellet was resuspended in 32 μL of deionised water to which 8 μL of 4 M NaCl and 40 μL of autoclaved 13% (w/v) PEG 8000 were added. The sample was then thoroughly mixed and left on ice for 20 min. Plasmid DNA was then pelleted by centrifugation at 14,000 rpm for 20 min at 4 °C. The supernatant was removed, the pellet washed with 500 μL of 70% (w/v) ethanol and resuspended in 20 μL sterile water.

2.6.4 QIAGEN Plasmid Midi and Maxi protocols

Once the desired cDNA clone was obtained, large quantities of plasmid DNA were isolated as a stock for future work, using the QIAprep Plasmid Midi or Maxi DNA isolation kit. The principle is similar to that of the QIAprep Mini prep procedure outlined in Section 2.6.2. The modified alkaline lysis procedure is followed with the binding of plasmid DNA to QIAGEN anion-exchange resin under appropriate low salt and pH conditions. RNA, proteins, dyes and low molecular weight impurities are removed by a medium salt wash. Plasmid DNA is eluted in a high salt buffer and then concentrated and desalted by isopropanol precipitation. This protocol is designed for preparation of up to 100 μg of high or low copy plasmid.

Three buffers (P1, P2 and P3) were used in the first stages of this method. An RNase A solution, which was provided with the QIAprep kit, was added to Buffer P1, to give a final RNase concentration of 100 μg mL^{-1}. Buffer P2 contained SDS, which may have

precipitated due to low temperature storage, and so was always warmed to 37 °C before use.

For midi or maxi prep procedures, a starter culture was set up. This was done by inoculating 2 to 5 mL of LB medium (supplemented with 100 μg mL^{-1} ampicillin) with a single fresh colony containing the clone of interest from a freshly streaked selective plate (L-agar and 100 μg mL^{-1} ampicillin). The starter culture was incubated at 37 °C for 8 h at 300 rpm.

Subsequently, the starter culture was diluted 1/500 or 1/1000 into 100 mL (for midi) 500 mL (for maxi) of LB medium with 100 μg mL^{-1} ampicillin and incubated at 37 °C for 12 to 16 h at 300 rpm. For both cultures a flask volume at least four times that of the culture, was used.

2.7 RESTRICTION ENZYME DIGEST OF PLASMID DNA

Restriction enzyme digests of plasmid DNA were generally performed to check orientation of constructs, to isolate fragments for radioactive labeling, differentiate clones by restriction patterns, check for the presence of absence of inserts from cDNA library screens or generate complementary sites on fragments and vectors prior to ligation and cloning. Reactions were generally performed with 1-2 μg of DNA (or 10-12 for making inserts), 5 U of restriction enzyme, 1 x appropriate buffer in 10-20 μL total reaction volume. Digestion was usually done at 37 °C for 1-2 h unless otherwise specified.

Subsequently, the digestion reaction was mixed with 2-3 μl of 10 x loading buffer and fractionated on 1% (w/v) TBE agarose gel containing EtBr (10 μg mL^{-1}), to confirm the presence and size of inserts. A 1 kb or a 100 bp DNA ladder (Invitrogen, New Zealand) was used based on the expected sizes of the cloned products. DNA products that were smaller than 300 bp were resolved on 2% (w/v) agarose gels. The band patterns were revealed under UV light and documented using a gel imaging system.

2.8 METHODS FOR DETERMINING DNA CONCENTRATION

Two methods, based on the well-known absorbance properties of DNA, were used to quantify yields of plasmid DNA from the various preparation methods. However, DNA was also quantified on occasion by gel electrophoresis. This method was employed to determine the concentration of insert fragments for ligations or making probes.

2.8.1 Quantifying plasmid DNA by spectrophotometry

The method described for quantifying RNA in Section 2.3.1 applies here as well. The concentration of DNA in the preparations was calculated on the basis that an A_{260} of 1.0 with a path length of 1 cm is equivalent to 50 μg mL^{-1} of double stranded DNA.

2.8.2 Fluorometry

Fluorometric measurements were performed using the Hoescht's fluorometer (Hoefer Scientific Instruments, USA). The mini fluorometer uses a special dye (bis-benzimidizole commonly known as the Hoechst 33258) that binds specifically to DNA and preferentially to A-T rich regions. RNA does not compete with DNA for binding with the dye. All measurements were done in 3 mL of the dye in a quartz fluorometry cuvette supplied with the fluorometer. The reference standard was set using 10 μL of a monomeric plasmid DNA (10 μg mL^{-1}) in 3 mL of the dye stock solution. Against this baseline an aliquot of 1 μL of prepared DNA sample was diluted in 3 mL of the Hoescht dye preparation and measured to give the concentration of DNA in ng μL^{-1}.

2.8.3 Quantifying plasmid DNA by agarose gel electrophoresis

To determine the concentration of plasmid DNA by gel electrophoresis, 4 μL of DNA preparation was separated simultaneously on a 1% (w/v) agarose gel (containing 10 μg mL^{-1} EtBr) alongside 4 μL of a low mass DNA ladder. Both samples were mixed with 2 μL of 10 x loading dye. Intensity of EtBr staining (in UV light) was compared between sample DNA and mass ladder fragments.

2.9 POLYMERASE CHAIN REACTIONS

PCR reactions were carried out with one of two polymerases. The first and most commonly used was *Taq* DNA polymerase (Invitrogen, New Zealand) That has the property of adding an adenine onto the end of PCR products and this can be used to facilitate easy and effective ligation into prepared pGEM-T Easy vector (Promega, USA). The second enzyme used was Platinum *Pfx* DNA polymerase (Invitrogen, New Zealand), which is a high fidelity, high processivity enzyme due to its proofreading 3-5 exonuclease activity (Cline et al., 1996). For this reason it was used when adding specific restriction sites to existing cDNA clones from the cDNA library but also in the attempts to isolate an anthurium UFGT. Unlike *Taq* polymerase Platinum *Pfx* generates blunt ended fragments therefore these PCR products were normally cloned into pBluescript which had been digested with *Sma* I.

2.9.1 *Taq* polymerase PCR

The template for the PCR reactions was either plasmid DNA or first strand cDNA. Reactions were set up in a total volume of 50 µL in thin walled sterile PCR tubes and consisted of 5 µL of 10 x PCR buffer with Mg^{2+}, 1 µL of 10 mM dNTP mix, 1 µL of 50 pmol μL^{-1} downstream primer, 1 µL of 50 pmol μL^{-1} upstream primer, 1-5 µL of template DNA (0.5-1 µg per reaction) and 0.5-1 µL of *Taq* polymerase (2.5 U μL^{-1}). Sterile water was added to bring the volume of the reaction to 50 µL. Once components were added, samples were mixed, briefly centrifuged and immediately placed in a thermocycler. Generally, a three step PCR was carried with the following cycling parameters; 95 °C/30 s (to denature DNA and primers), 55-60 °C/1 min to allow annealing and 72 °C/30 s or 1 min (to allow elongation depending on the size of the PCR fragment) and one final elongation at 72 °C/2-5 min. The PCR was performed with 25-45 cycles depending on the abundance of the target. An aliquot of 5-8 µL of PCR product was analysed on a 1% (w/v) agarose gel as described in Section 2.8.3 except that a 1 kb DNA ladder was used instead of the mass ladder.

2.9.2 Platinum *Pfx* polymerase PCR

The enzyme is sold in an inactive form specifically bound to an antibody. This provides an automatic hot start for the PCR therefore increasing the specificity, sensitivity and yield. PCR reactions were set up in a 50 µL reaction volume with 5 µL 10 x *Pfx* amplification buffer, 1.5 µL of 10 mM dNTP mixture, 1 µL 50 mM $MgSO_4$, 1 µL of each primer at 50 pmol μL^{-1}, 0.5-1 µg DNA template and 0.5 µL of the *Pfx* polymerase (2.5 U μL^{-1}). The contents were mixed and briefly centrifuged. The PCR cycle was set up to first run a single cycle at 95 °C/2.5 min s to release the enzyme from the antibody. This was followed by 35-40 cycles of 95 °C/30 s, 50 °C/30 sec and 68 °C/30 sec. The reaction was completed with a single cycle of four min at 68 °C. PCR products were then analysed by gel electrophoresis.

In both PCR approaches, bands of interest were purified from the agarose gel (Section 2.10) and ligated into the required vector (Section 2.12). For sequencing, PCR inserts cloned into vectors, were transformed into XL1-Blue MRF' competent cells by the Heat shock method (Section 2.13) and subsequently cultured for plasmid DNA isolation. Plasmid DNA was subsequently isolated by one of the procedures described previously in Sections 2.6.1-2.6.4 and sent for sequencing using the ABI automated DNA sequencer described in Section 2.5.8.

2.9.3 3' RACE PCR

3'- RACE PCR was carried out according to the protocol described by Frohman et al. (1988). The procedure was employed in attempts to isolate anthurium; *CHI, FNS, F3'H, UF3GT*, and Myb cDNAs as discussed in Chapters 3, 6 and 7.

Several first strand cDNA stocks were made as template for the PCR reactions using RNA isolated from various combinations of the six stages of spathe development. First strand cDNA stocks were made for each individual flower stage, Stages 1 and 2 together, Stages 3 to 6 together and from all six stages pooled together.

First strand cDNA synthesis was primed using a dT_{17} adaptor sequence with *Sal* I, *Xho* I and *Cla* I sites attached to the 5' end of the primer as described by Frohman et al. (1988). The RNA directed DNA polymerase, Expand Reverse Transcriptase (Roche Molecular Diagnostics, USA), was used to catalyse the synthesis of the first strand cDNA. This enzyme has been genetically engineered with a point mutation in the RNase H sequence so as to reduce the RNase activity to undetectable levels. This, according to the manufacturers, ensures higher levels of full-length cDNA transcripts compared to the native Moloney Murine Leucaemia virus reverse transcriptase from which it was derived.

To set up the synthesis reaction, reagents were added into an RNase free 0.2 mL thin walled PCR tube. Firstly, 1 μg of total RNA was added to the tube followed by 1 μL of a 50 pmol μL^{-1} dT_{17} stock. Total volume was made up to 10 μL with RNase free water and then denatured using a thermocycler at 65 °C for 10 mins. The denatured RNA and primer were then immediately stored on ice. To this 10 μL reaction was added 4 μL of 5x Expand reverse transcriptase buffer, 2 μL of 100 mM DTT, 2 μL of 10 mM dNTP stock, 1 μL of RNase inhibitor at concentration of 40 units μL^{-1} and finally 1 μL of the Expand RT enzyme (50 U μL^{-1}). The reaction was made up to a total of 20 μL with sterile water. Ingredients were mixed by vortexing and centrifuged briefly, and then incubated in the thermocycler for 45-60 mins at 42 °C. Once completed the enzyme was heat inactivated by heating at 95 °C for 2 min.

The cDNA was purified using High Pure PCR Kit (Roche Molecular Diagnostics, USA). The cDNA synthesis reaction was made up to 100 μL with sterile water and then 500 μL of binding buffer was added. The tube was vortexed and taken through the purification procedure as described in Section 2.10. The cDNA was eluted in 100 μL of water and stored at -20 °C.

2.10 PURIFICATION OF DNA FROM AGAROSE GELS

The High Pure PCR Purification kit was used to isolate DNA from agarose gels for making probes or for subsequent cloning and further manipulation. The band of interest was excised from the gel with a sterile razor, taking the smallest possible gel fraction. An

amount of 300 μL of binding buffer was added per 100 mg of gel slice. The sample was vortexed, to resuspend the gel slice and incubated for 10 min (or until the agarose dissolved completely) at 56 °C. Once completely dissolved, isopropanol was added to the mixture at 150 μL per 100 mg of gel slice, and thoroughly vortexed. The High Pure filter tube was then inserted into the collection tube and the sample was transferred into the upper reservoir of the filter tube. This was centrifuged for 30 s at maximum speed in a standard microcentrifuge. The eluate was discarded and 500 μL of washing buffer was added to the upper reservoir of the filter tube and centrifuged as above. Again the eluate was discarded and 200 μL of wash buffer added to the upper reservoir of the filter tube and centrifuged again. The filter tube was then placed in an microcentrifuge tube and 100 μL of sterile water added to elute the DNA. This was left for 1-2 min and then centrifuged as before. The eluted, purified DNA was then used for further work.

2.11 PREPARATION OF COMPETENT CELLS

This method is based on the simple and efficient method (SEM) for preparing competent *E. coli* cells devised by Inoue et al. (1990).

Reagents for SEM Buffer

10 mM Pipes

55 mM $MnCl_2$

15 mM $CaCl_2$

250 mM KCl

All the components of the buffer except $MnCl_2$ were mixed and the pH adjusted to 6.7 with KOH. The $MnCl_2$ was dissolved in the mixture, then the solution was sterilised by filtration through a 0.45 μm filter.

A 200 mL L-Broth culture with tetracycline 12.5 μg uL^{-1} was inoculated with 400 μl of an overnight XL1-Blue MRF' culture and left to grow for 24 h (until an A_{600} of 0.6 was reached) at 22 °C in a shaking incubator (200-250 rpm). The flask was then removed and placed on ice for 20 min. The culture was then transferred to a 500 mL Sorval centrifuge

bottle and spun at 2500 g in a Sorval GSA rotor for 10 min at 4 °C. The pellet was resuspended in 20 mL ice cold SEM buffer, incubated on ice for 10 min and then centrifuged as above. The pelleted cells were again gently resuspended with 6 mL ice cold SEM, and DMSO was added with gentle swirling to a final concentration of 7% (v/v). Cells were then quickly aliquoted out into microcentrifuge tubes and immediately frozen in liquid N_2 and stored at -80 °C.

2.12 DNA LIGATION REACTIONS

Ligation reactions were performed when cloning PCR products for sequencing or when making constructs for transient gene expression studies. In the former situation, the pGEM-T Easy vector was used. The vector had been cut with *Eco* RV and 3' terminal thymidine added to both ends. These single 3' thymidine overhangs prevent recircularisation of the vector while providing compatible ends for PCR products generated by *Taq* polymerase (which adds a terminal adenine to the 3' strand end of the PCR products). These two factors greatly increased the ligation efficiency. Insertional inactivation of the α-peptide coding region of β-galactosidase enzyme allowed selection of recombinants by blue/white screening on LB plates containing IPTG/X-Gal in addition to the appropriate antibiotic.

Irrespective of the vector being used, ligation reactions were set up to optimize the ligation efficiency by using an insert:vector ratio of 3:1 based on the concentration of the DNA fragments. PCR products were analysed on 1% (w/v) agarose gels and products of correct size cut out and purified as described in Section 2.10. The concentration of the purified PCR product was estimated by comparison to the appropriate DNA mass ladder on a 1% (w/v) agarose gel as described in Section 2.8.3

When using the Promega pGEM-T Easy vector the general ligation reaction (in 10 μL) was as follows:

2X Rapid ligation buffer T4 DNA ligase	5 μL
pGEM-T Easy vector	1 μL

PCR product	X μL
T4 DNA ligase 3U uL^{-1}	1 μL
Water to bring final volume to 10 μL	Y μL

Reactions were mixed by gently tapping on the sides of the tube and incubated either at room temperature for 1 h at 4 °C or overnight for large fragments (>1 kb) to increase the number of colonies after transformation. When ligating into other vectors, such as pRT99, a 10 x ligation buffer was used to a final concentration of 1 x.

2.13 HEAT SHOCK TRANSFORMATION

Competent cells were gradually thawed on ice and an aliquot of plasmid DNA or ligation reaction (1-10 μL) was added to 100 μL of competent stock cells and gently mixed. The tubes were left on ice for 30 min, then heat shocked by incubating at 42 °C for 60 s, followed by an incubation on ice for 2 min. Approximately 300 μL of L-Broth was added and the cells incubated for 1 h in a orbital incubator at 37 °C and 250 rpm. The cells were then plated directly onto LB plates containing the appropriate antibiotic and incubated overnight at 37 °C. For blue/white selection 40 μL of a 50 mg mL^{-1} X-Gal stock and 20 μL of a 50 mg mL^{-1} IPTG stock was spread over the surface of the plates with a sterile glass spreader and allowed to air dry prior to plating transformed cells.

2.14 NORTHERN ANALYSIS

Northern hybridisation uses DNA:RNA hybridisation to determine the steady state levels and size of RNA gene transcripts. This procedure is based on the use of formaldehyde/formamide as denaturants. These also ensure efficient transfer of RNA to the membrane. When measuring transcript levels of biosynthetic genes, 20 μg of total cellular RNA was used. However, for regulatory genes, which are expressed at very low levels within the cell, poly $(A)^+$ RNA was purified and used to determine transcript abundance and expression patterns. All equipment used in these procedures was RNase free and all solutions were made with RNase free water. Essentially the procedure could be divided into four stages as described in the following sections.

2.14.1 Gel electrophoresis of RNA

Denaturing 1% (w/v) agarose gels were used to resolve RNA molecules ranging in size from 0.5 kb to 10 kb. These gels contained 0.66 M formaldehyde and were made with 1 x MOPS. Two sizes of gels were prepared based on the number of RNA samples that were being analysed. For each 1.95 g agarose, 15 mL 10 x MOPS, 130.5 mL water and 7.65 mL formaldehyde 37% (v/v) was used. 10 x MOPS (pH 7.0) contains 0.2 M MOPS, 50 mM KAc and 10 mM EDTA.Na_2

All components used were mixed in a sterile Schott bottle, swirled around gently to rehydrate the agarose and then microwaved until the agarose was completely melted. The liquified agarose was then cooled to 50-55 °C prior to adding formaldehyde. Immediately thereafter the gel was poured into RNase free gel tanks and allowed to set.

All RNA samples were kept on ice during preparation. 20 µg total RNA was added to sterile microcentrifuge tubes and denaturing solution was added to 3 x final volume of RNA solution in each tube. A stock of denaturing solution was made up as follows:

Reagents	Per sample	For 20 samples
10 x MOPS	2.0 µL	40.0 µL
Formamide (deionised)	10.0 µL	200 µL
Formaldehyde (37% (v/v) stock)	3.5 µL	70 µL
Ethidium bromide (1 µg mL^{-1})	1.0 µL	Nil
Ethidium bromide (10 µg/mL)	Nil	2 µL

Once the denaturing solution was added the samples were incubated at 65 °C for 20 min to remove any secondary structures from the RNA. Following this, samples were placed on ice and 5-10 µL of RNA loading buffer were added to each tube. An aliquot of 5 µL of RNA ladder (0.75 kb-9.74 kb) (Invitrogen, New Zealand) was used in all experiments and prepared for electrophoresis as the sample RNA. All samples were then immediately loaded

into the denaturing gel using filter sterilised RNA loading buffer and electrophoresed in 1 x MOPS first at 60 V for 1 h and then overnight at 25 V.

RNA Loading buffer
30% (w/v) Ficoll
1 mM EDTA
0.25% (w/v) Bromophenol blue
0.25% (w/v) Xylene cyanol FF

The following day the gel was removed, viewed, marked (to define orientation) and documented on a U.V transilluminator using a gel imaging system.

2.14.2 Blotting of RNA

The blotting apparatus was set up according to Sambrook et al. (1989) using 10 x SSC. The gel was carefully removed from the gel tray, inverted and placed on top a 3MM Whatmann paper wick that was placed on a sponge saturated with 10 x SSC. All air bubbles that were trapped between the gel and the paper wick were removed by gently rolling a sterile glass tube over the surface of the gel. Strips of parafilm were cut and placed adjacent to the sides of the gel to prevent the buffer from short circuiting around the gel rather than passing through it. On top of the gel was placed an equally sized piece of positively charged Hybond N^+ membrane soaked in 10 x SSC. Once again all air bubbles were removed, as this would hinder the transfer of RNA to the membrane. The membrane was then marked in the same places as the gel and three pieces of 3MM Whatmann paper (of equal size with the membrane) soaked in 10 x SSC were placed on the membrane. Stacks of paper towels, all of equal size to the gel and membrane, were placed on the Whatmann paper at a height slightly above the top of the container ensuring a gentle pressure was exerted on the gel when the lid was attached. This was essential to set up the capillary action needed for efficient transfer of RNA to the membrane.

Gels were blotted overnight, or longer if very thick gels were used. Following this the paper towels and filter paper were discarded and the membrane removed and washed gently in

2 x SSC. RNA was fixed to the membrane using a Hoefer UV-C, 500 UV cross-linker for 1 min at 70,000 μJ cm^{-2}. The membrane, protected by a plastic covering, was viewed on a UV transilluminator to confirm transfer of RNA, to mark bands of the RNA ladder and to photograph. After this the membranes were probed immediately or stored in sealed plastic and kept at 4 °C.

2.14.3 Hybridisation

cDNA clones obtained from the library screen were used as DNA probes for northern analysis. Full-length inserts were cut out of plasmid DNA using the appropriate restriction enzymes and purified, and the insert concentration determined as described in Sections 2.7, 2.10 and Section 2.8.3, respectively. Aliquots of 60-100 ng were used to make [α-^{32}P]dCTP radiolabeled probes.

The making of radiolabeled probes, followed by prehybridisation and hybridisation of membranes was as described in Section 2.5.5. However, prehybridisation and hybridisation were performed with high stringency at 65 °C, as opposed to 55 °C for library screening.

Once membranes were prehybridised, the solution was discarded and fresh Church and Gilbert solution (10-20 mL) was added. The probes were then denatured by incubating at 95 °C for 4 min and then added to the solution with the membranes. Hybridisation was performed at 65 °C overnight in a rotary oven.

Hybridisation was followed by three 15 min high stringency washes of the membranes at 65 °C. The first wash used 2 x SSC buffer, the second 1 x SSC and the final wash 0.1 x SSC. All wash solutions also contained 0.1% (v/v) SDS. The membranes were then exposed to photographic film for autoradiography at -80 °C overnight (longer exposures were used in some instances) and developed the following day according to standard procedures.

Chapter 3

Characterisation of cDNA clones for anthurium flavonoid biosynthetic genes

3.1 INTRODUCTION

Isolation and characterisation of genes involved in flavonoid biosynthesis in the anthurium spathe is necessary in order to develop a full understanding of the flower colour development in this species. Once a cDNA clone is isolated, information such as copy number in the genome, temporal and spatial distribution of gene transcripts and phylogenetic patterns can be obtained. In addition, it is common to perform biochemical assays with the expressed protein to confirm functionality. Several more experiments can be performed depending on the objectives that may shed light on the endogenous role of the gene. Further to this, the availability of the cDNA clones is a vital resource in subsequent transformation experiments to manipulate flower colour.

For some cDNA clones it is possible to ascribe an identity based on analysis, and comparison of their deduced amino acid sequences with those proteins with defined functions. Many of the structural genes involved in anthocyanin biosynthesis in other species have been identified in this way, aided by the extensive sequence databases available. However, characterisation by genetic studies or expression of the protein in a heterologous host is often the only definitive approach to accurately ascribing identity and function.

3.2 ISOLATION OF AN ANTHURIUM *CHS* cDNA CLONE

Heterologous screening of the anthurium library with an apple full-length CHS cDNA insert gave nine independent positive signals in the primary screen, six of which followed through to secondary and tertiary screens. Following in vivo excision and plasmid DNA preparation, two of these gave fragments of the right size to be full-length clones (1.5-1.6 kb). These had the same restriction pattern suggesting they were derived from the same mRNA transcript (data not shown). These were sequenced with T3 and T7 primers generating sequence data for each clone from the 5' and 3' ends respectively. Upon analysis of the nucleotide and deduced amino acid sequences, the two clones were found to encode identical putative *CHS* full-length clones. Sequence data were not obtained for the shorter

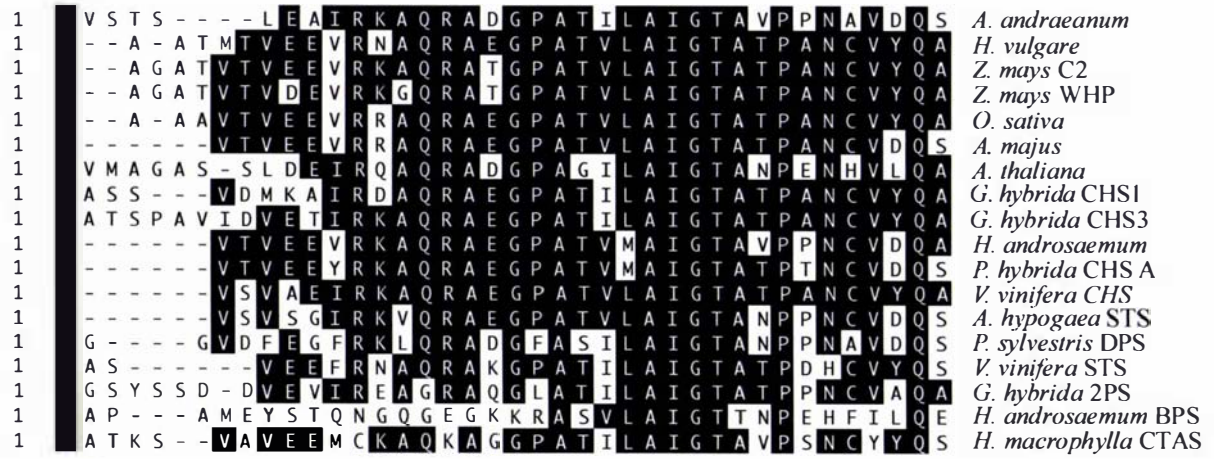
clones. One putative full-length cDNA was selected and named *AaCHS1*. The nucleotide sequence for the open reading frame is presented in Appendix II.

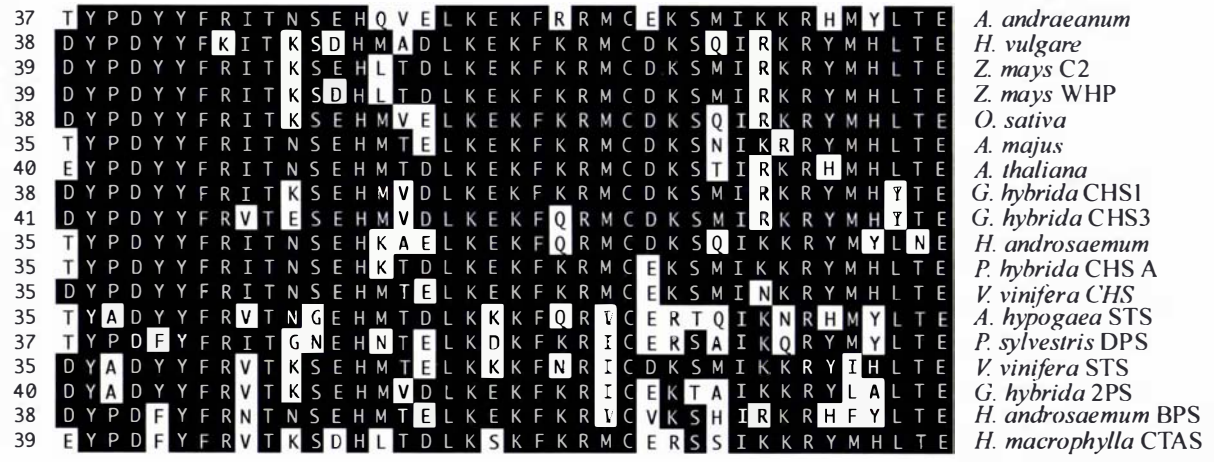
The *AaCHS1* cDNA clone obtained from the library was 1.5 kb in length with an estimated G+C content of 66.7%. The cDNA clone contained 139 bp of 5' untranslated sequence and the 3' untranslated sequence was 245 bp in length with a poly $(A)^+$ extension of 19 bp. A conserved polyadenylation signal (AATAAT) was located 148 bp downstream of the stop codon (Joshi, 1987). The longest open reading frame was 1176 bp and encoded a 391 amino acid deduced polypeptide with a predicted mass of 43 kD and an estimated isoelectric point of 6.5.

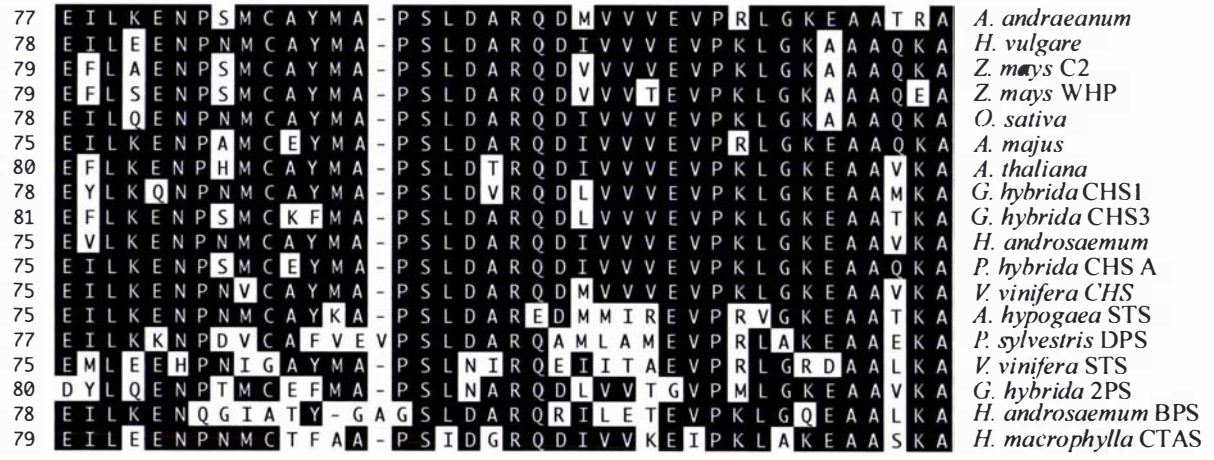
Chalcone synthase is a polyketide synthase, and AaCHS1 has all the structural features in its primary protein structure that define this diverse family of enzymes. First and most significant is the conserved cysteine residue (marked with an arrow in Figure 3.1) that defines the active site of polyketide synthases along with the three absolutely conserved amino acids (phenylalanine, histidine and asparagine) that are involved in the decarboxylation of malonyl-CoA (Ferrer et al., 1999). The CHS/STS signature motif (Bairoch, 1991; Fliegmann et al., 1992) was also conserved in AaCHS1 (Figure 3.1). There are some consistent differences between CHSs and STSs in residues 253-259 and 262-268, marked with broken lines in Figure 3.1. These are solvent exposed residues and provide the biochemical basis for the two related enzymes (CHSs and STSs) to utilise the same substrate but have alternate cyclisation pathways that gives rise to different end products (Ferrer et al., 1999). The deduced polypeptide encoded by the *AaCHS1* clone had the conserved residues present in all functionally characterised CHSs within this region.

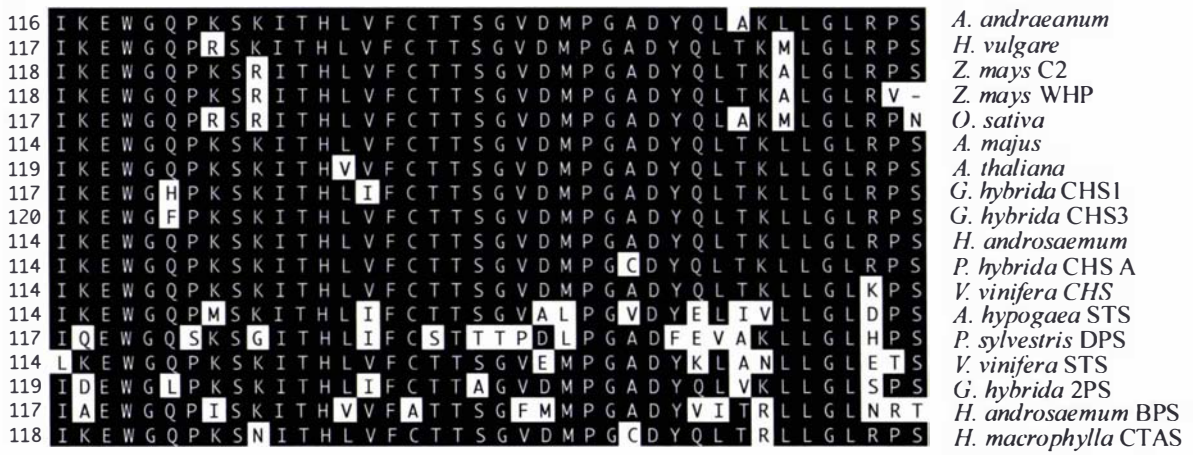
The *AaCHS1* cDNA clone hybridised strongly to a heterologous *CHS* probe and when BLAST searches were performed with either the cDNA sequence or the deduced polypeptide sequence, the best hit was always for CHS proteins. Southern analysis was performed but no conclusive statement could have been made on the number of *CHS* copies in the anthurium genome (data not shown). Given the high levels of sequence conservation reported for *CHS*, the alignment as shown in Figure 3.1 was designed to compare the

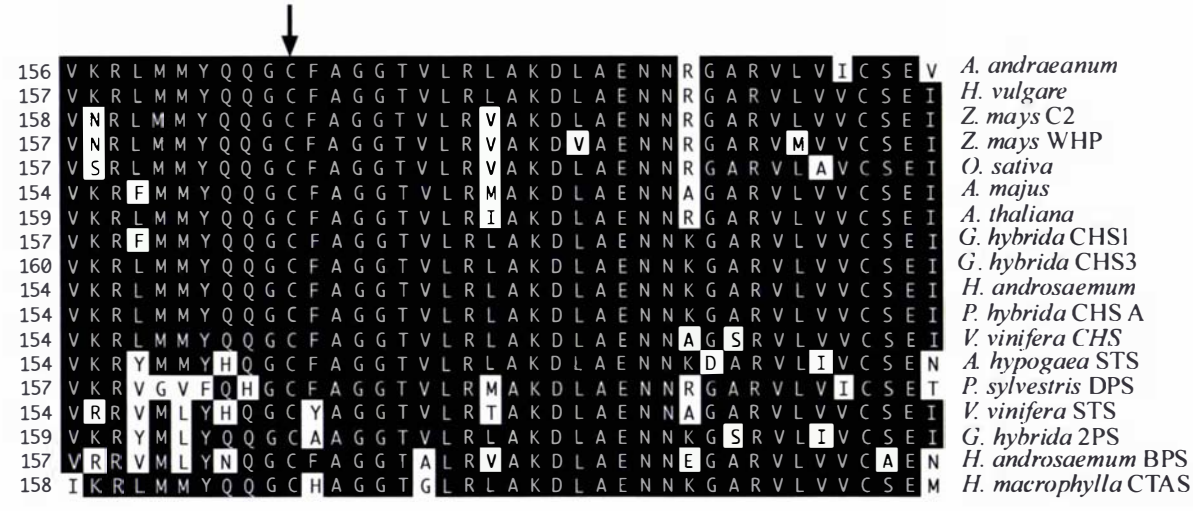
Figure 3.1. Multiple alignment of the deduced amino acid sequence of AaCHS1 with confirmed CHS and CHS-related proteins. The alignment was performed using the Clustal Method (Higgins and Sharp, 1989). Black shading indicates identical amino acids. Dashes indicate gaps in the sequence for alignments. The cysteine residue essential for catalytic activity is marked with an arrow and the three residues involved in decarboxylation of malonyl-CoA are marked with asterisks. The CHS/STS signature motif (Bairoch, 1991; Fliegmann et al., 1992) is underlined with a solid line. The broken line, identify the residues lining the cyclisation pocket that vary between CHS and STS proteins. The CHS sequences and their accession numbers are: *Anthurium andraeanum* (this work), *Hordeum vulgare* (P26018), *Oryza sativa* (X89859), *Zea mays* C2 and WHP (P24825, X60204 respectively), *Antirrhinum majus* (P06515), *Arabidopsis thaliana* (P13114), *Gerbera hybrida* CHS1 and CHS3 (Z38096 and Z38098 respectively), *Hypericum androsaemum* (AF315345) *Petunia hybrida* CHS-A (AF233638), *Vitis vinifera* (AB015872). The STSs: *Arachis hypogaea* STS (A00769), *Pinus sylvestris* dihydropinosylvin synthase (X60753), *Vitis vinifera* STS (P28343). Other CHS-related proteins: *Gerbera hybrida* 2PS-2-pyrone synthase (Z38097) and *Hypericum androsaemum* BPS- benzophenone synthase (AF352395).

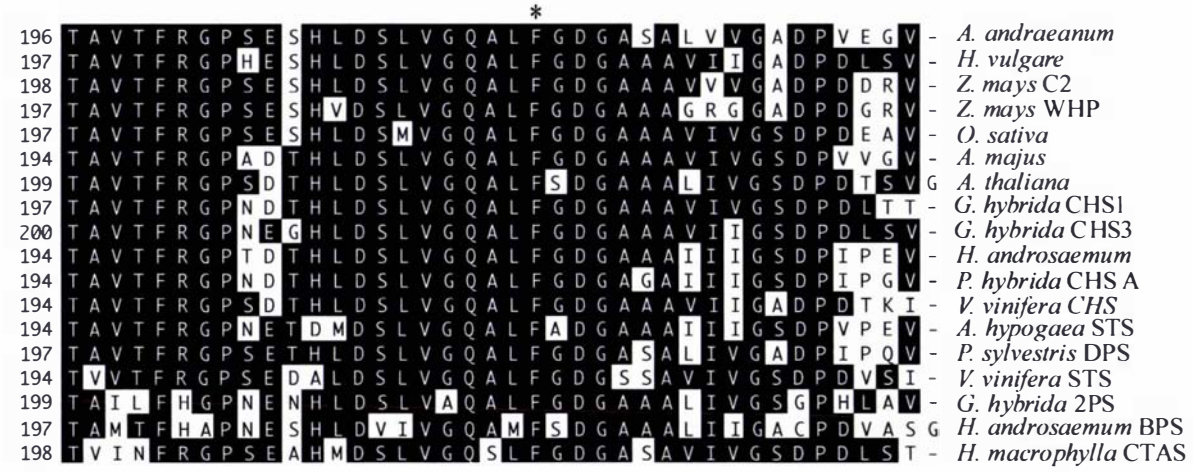

1 VSTS----LEAIRKAQRADGPATILAIGTAVPPNAVDQS A. andraeanum
1 --A-ATMTVEEVRNAQRAEGPATVLAIGTATPANCVYQA H. vulgare
1 --AGATVTVEEVRKAQRATGPATVLAIGTATPANCVYQA Z. mays C2
1 --AGATVTVDEVRKGQRATGPATVLAIGTATPANCVYQA Z. mays WHP
1 --A-AAVTVEEVRRAQRAEGPATVLAIGTATPANCVYQA O. sativa
1 ------VTVEEVRRAQRAEGPATVLAIGTATPANCVDQS A. majus
1 VMAGAS-SLDEIRQAQRADGPAGILAIGTANPENHVLQA A. thaliana
1 ASS---VDMKAIRDAQRAEGPATILAIGTATPANCVYQA G. hybrida CHS1
1 ATSPAVIDVETIRKAQRAEGPATILAIGTATPANCVYQA G. hybrida CHS3
1 ------VTVEEVRKAQRAEGPATVMAIGTAVPPNCVDQA H. androsaemum
1 ------VTVEEYRKAQRAEGPATVMAIGTATPTNCVDQS P. hybrida CHS A
1 ------VSVAEIRKAQRAEGPATVLAIGTATPANCVYQA V. vinifera CHS
1 ------VSVSGIRKVQRAEGPATVLAIGTANPPNCVDQS A. hypogaea STS
1 G----GVDFEGFRKLQRADGFASILAIGTANPPNAVDQS P. sylvestris DPS
1 AS-------VEEFRNAQRAKGPATILAIGTATPDHCVYQS V. vinifera STS
1 GSYSSD-DVEVIREAGRAQGLATILAIGTATPPNCVAQA G. hybrida 2PS
1 AP---AMEYSTQNGQGEGKKRASVLAIGTTNPEHFILQE H. androsaemum BPS
1 ATKS--VAVEEMCKAQKAGGPATILAIGTAVPSNCYYQS H. macrophylla CTAS

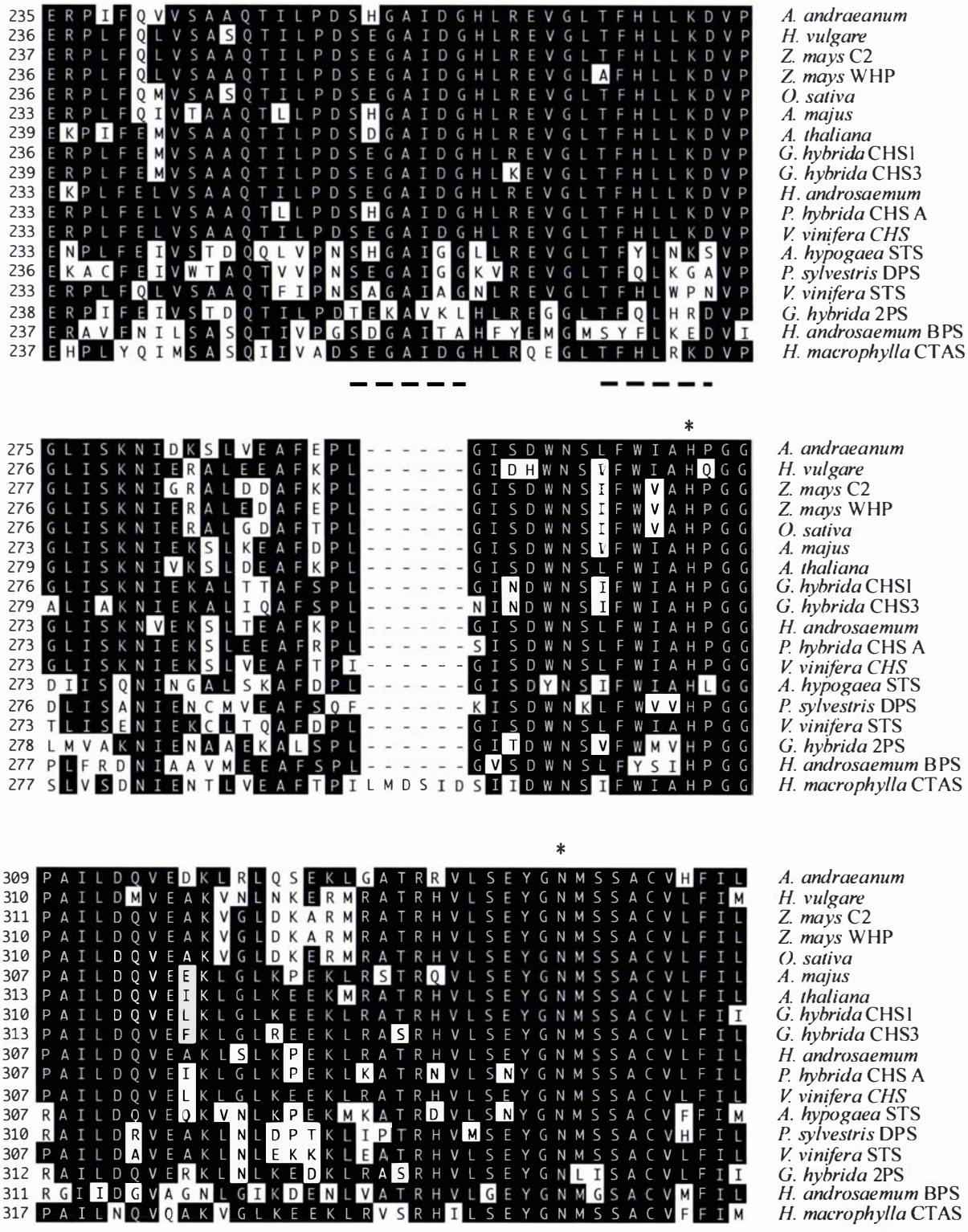

37 TYPDYYFRITNSEHQVELKEKFRRMCEKSMIKKRHMYLTE A. andraeanum
38 DYPDYYFKITKSDHMADLKEKFKRMCDKSQIRKRYMHLTE H. vulgare
39 DYPDYYFRITKSEHLTDLKEKFKRMCDKSMIRKRYMHLTE Z. mays C2
39 DYPDYYFRITKSDHLTDLKEKFKRMCDKSMIRKRYMHLTE Z. mays WHP
38 DYPDYYFRITKSEHMVELKEKFKRMCDKSQIRKRYMHLTE O. sativa
35 TYPDYYFRITNSEHMTELKEKFKRMCDKSNIKRRYMHLTE A. majus
40 EYPDYYFRITNSEHMTDLKEKFKRMCDKSTIRKRHMHLTE A. thaliana
38 DYPDYYFRITKSEHMVDLKEKFKRMCDKSMIRKRYMHYTE G. hybrida CHS1
41 DYPDYYFRVTESEHMVDLKEKFQRMCDKSMIRKRYMHYTE G. hybrida CHS3
35 TYPDYYFRITNSEHKAELKEKFQRMCDKSQIKKRYMYLNE H. androsaemum
35 TYPDYYFRITNSEHKTDLKEKFKRMCEKSMIKKRYMHLTE P. hybrida CHS A
35 DYPDYYFRITNSEHMTELKEKFKRMCEKSMINKRYMHLTE V. vinifera CHS
35 TYADYYFRVTNGEHMTDLKKKFQRVCERTQIKNRHMYLTE A. hypogaea STS
37 TYPDFYFRITGNEHNTELKDKFKRICERSAIKQRYMYLTE P. sylvestris DPS
35 DYADYYFRVTKSEHMTELKKKFNRICDKSMIKKRYIHLTE V. vinifera STS
40 DYADYYFRVTKSEHMVDLKEKFKRICEKTAIKKRYLALTE G. hybrida 2PS
38 DYPDFYFRNTNSEHMTELKEKFKRVCVKSHIRKRHFYLTE H. androsaemum BPS
39 EYPDFYFRVTKSDHLTDLKSKFKRMCERSSIKKRYMHLTE H. macrophylla CTAS

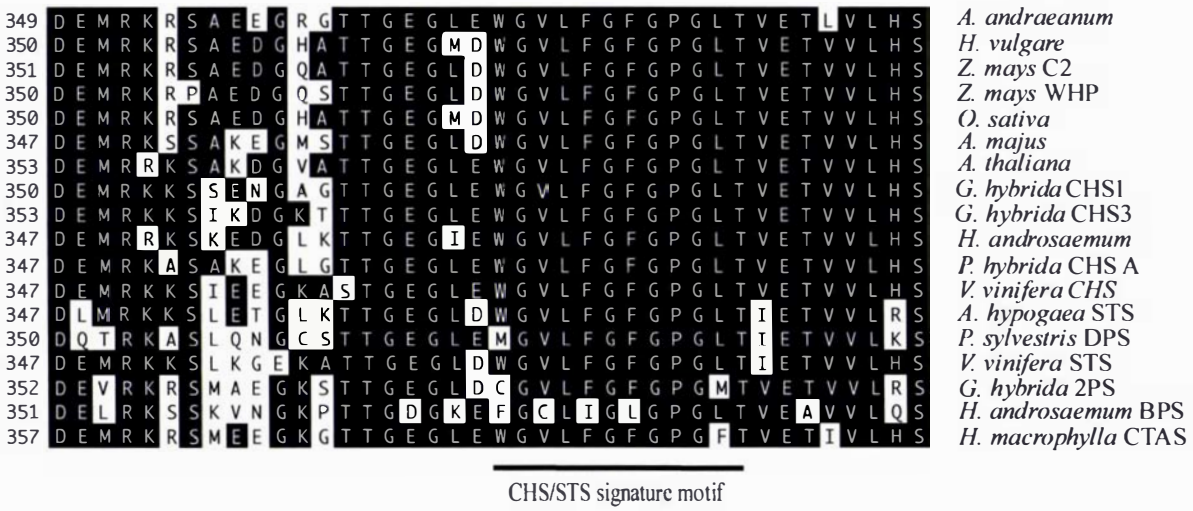

77 EILKENPSMCAYMA-PSLDARQDMVVVEVPRLGKEAATRA A. andraeanum
78 EILEENPNMCAYMA-PSLDARQDIVVVEVPKLGKAAAQKA H. vulgare
79 EFLAENPSMCAYMA-PSLDARQDVVVVEVPKLGKAAAQKA Z. mays C2
79 EFLSENPSMCAYMA-PSLDARQDVVVTEVPKLGKAAAQEA Z. mays WHP
78 EILQENPNMCAYMA-PSLDARQDIVVVEVPKLGKAAAQKA O. sativa
75 EILKENPAMCEYMA-PSLDARQDIVVVEVPRLGKEAAQKA A. majus
80 EFLKENPHMCAYMA-PSLDTRQDIVVVEVPKLGKEAAVKA A. thaliana
78 EYLKQNPNMCAYMA-PSLDVRQDLVVVEVPKLGKEAAMKA G. hybrida CHS1
81 EFLKENPSMCKFMA-PSLDARQDLVVVEVPKLGKEAATKA G. hybrida CHS3
75 EVLKENPNMCAYMA-PSLDARQDIVVVEVPKLGKEAAVKA H. androsaemum
75 EILKENPSMCEYMA-PSLDARQDIVVVEVPKLGKEAAQKA P. hybrida CHS A
75 EILKENPNVCAYMA-PSLDARQDMVVVEVPKLGKEAAVKA V. vinifera CHS
75 EILKENPNMCAYKA-PSLDAREDMMIREVPRVGKEAATKA A. hypogaea STS
77 EILKKNPDVCAFVEVPSLDARQAMLAMEVPRLAKEAAEKA P. sylvestris DPS
75 EMLEEHPNIGAYMA-PSLNIRQEIITAEVPRLGRDAALKA V. vinifera STS
80 DYLQENPTMCEFMA-PSLNARQDLVVTGVPMLGKEAAVKA G. hybrida 2PS
78 EILKENQGIATY-GAGSLDARQRILETEVPKLGQEAALKA H. androsaemum BPS
79 EILEENPNMCTFAA-PSIDGRQDIVVKEIPKLAKEAASKA H. macrophylla CTAS


116 IKEWGQPKSKITHLVFCTTSGVDMPGADYQLAKLLGLRPS A. andraeanum
117 IKEWGQPRSKITHLVFCTTSGVDMPGADYQLTKMLGLRPS H. vulgare
118 IKEWGQPKSRITHLVFCTTSGVDMPGADYQLTKALGLRPS Z. mays C2
118 IKEWGQPKSRITHLVFCTTSGVDMPGADYQLTKALGLRV- Z. mays WHP
117 IKEWGQPRSRITHLVFCTTSGVDMPGADYQLAKMLGLRPN O. sativa
114 IKEWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPS A. majus
119 IKEWGQPKSKITHVVFCTTSGVDMPGADYQLTKLLGLRPS A. thaliana
117 IKEWGHPKSKITHLIFCTTSGVDMPGADYQLTKLLGLRPS G. hybrida CHS1
120 IKEWGFPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPS G. hybrida CHS3
114 IKEWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPS H. androsaemum
114 IKEWGQPKSKITHLVFCTTSGVDMPGCDYQLTKLLGLRPS P. hybrida CHS A
114 IKEWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLKPS V. vinifera CHS
114 IKEWGQPMSKITHLIFCTTSGVALPGVDYELIVLLGLDPS A. hypogaea STS
117 IQEWGQSKSGITHLIFCSTTTPDLPGADFEVAKLLGLHPS P. sylvestris DPS
114 LKEWGQPKSKITHLVFCTTSGVEMPGADYKLANLLGLETS V. vinifera STS
119 IDEWGLPKSKITHLIFCTTAGVDMPGADYQLVKLLGLSPS G. hybrida 2PS
117 IAEWGQPISKITHVVFATTSGFMMPGADYVITRLLGLNRT H. androsaemum BPS
118 IKEWGQPKSNITHLVFCTTSGVDMPGCDYQLTRLLGLRPS H. macrophylla CTAS


156 VKRLMMYQQGCFAGGTVLRLAKDLAENNRGARVLVICSEV A. andraeanum
157 VKRLMMYQQGCFAGGTVLRLAKDLAENNRGARVLVVCSEI H. vulgare
158 VNRLMMYQQGCFAGGTVLRVAKDLAENNRGARVLVVCSEI Z. mays C2
157 VNRLMMYQQGCFAGGTVLRVAKDVAENNRGARVMVVCSEI Z. mays WHP
157 VSRLMMYQQGCFAGGTVLRVAKDLAENNRGARVLAVCSEI O. sativa
154 VKRFMMYQQGCFAGGTVLRMAKDLAENNAGARVLVVCSEI A. majus
159 VKRLMMYQQGCFAGGTVLRIAKDLAENNRGARVLVVCSEI A. thaliana
157 VKRFMMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEI G. hybrida CHS1
160 VKRLMMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEI G. hybrida CHS3
154 VKRLMMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEI H. androsaemum
154 VKRLMMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEI P. hybrida CHS A
154 VKRLMMYQQGCFAGGTVLRLAKDLAENNAGSRVLVVCSEI V. vinifera CHS
154 VKRYMMYHQGCFAGGTVLRLAKDLAENNKDARVLIVCSEN A. hypogaea STS
157 VKRVGVFQHGCFAGGTVLRMAKDLAENNRGARVLVICSET P. sylvestris DPS
154 VRRVMLYHQGCYAGGTVLRTAKDLAENNAGARVLVVCSEI V. vinifera STS
159 VKRYMLYQQGCAAGGTVLRLAKDLAENNKGSRVLIVCSEI G. hybrida 2PS
157 VRRVMLYNQGCFAGGTALRVAKDLAENNEGARVLVVCAEN H. androsaemum BPS
158 IKRLMMYQQGCHAGGTGLRLAKDLAENNKGARVLVVCSEM H. macrophylla CTAS


196 TAVTFRGPSESHLDSLVGQALFGDGASALVVGADPVEGV- A. andraeanum
197 TAVTFRGPHESHLDSLVGQALFGDGAAAVIIGADPDLSV- H. vulgare
198 TAVTFRGPSESHLDSLVGQALFGDGAAAVVVGADPDDRV- Z. mays C2
197 TAVTFRGPSESHVDSLVGQALFGDGAAAGRGGADPDGRV- Z. mays WHP
197 TAVTFRGPSESHLDSMVGQALFGDGAAAVIVGSDPDEAV- O. sativa
194 TAVTFRGPADTHLDSLVGQALFGDGAAAVIVGSDPVVGV- A. majus
199 TAVTFRGPSDTHLDSLVGQALFSDGAAALIVGSDPDTSVG A. thaliana
197 TAVTFRGPNDTHLDSLVGQALFGDGAAAVIVGSDPDLTT- G. hybrida CHS1
200 TAVTFRGPNEGHLDSLVGQALFGDGAAAVIIGSDPDLSV- G. hybrida CHS3
194 TAVTFRGPTDTHLDSLVGQALFGDGAAAIIIGSDPIPEV- H. androsaemum
194 TAVTFRGPNDTHLDSLVGQALFGDGAGAIIIGSDPIPGV- P. hybrida CHS A
194 TAVTFRGPSDTHLDSLVGQALFGDGAAAVIIGADPDTKI- V. vinifera CHS
194 TAVTFRGPNETDMDSLVGQALFADGAAAIIIGSDPVPEV- A. hypogaea STS
197 TAVTFRGPSETHLDSLVGQALFGDGASALIVGADPIPQV- P. sylvestris DPS
194 TVVTFRGPSEDALDSLVGQALFGDGSSAVIVGSDPDVSI- V. vinifera STS
199 TAILFHGPNENHLDSLVAQALFGDGAAALIVGSGPHLAV- G. hybrida 2PS
197 TAMTFHAPNESHLDVIVGQAMFSDGAAALIIGACPDVASG H. androsaemum BPS
198 TVINFRGPSEAHMDSLVGQSLFGDGASAVIVGSDPDLST- H. macrophylla CTAS


235 ERPIFQVVSAAQTILPDSHGAIDGHLREVGLTFHLLKDVP A. andraeanum
236 ERPLFQLVSASQTILPDSEGAIDGHLREVGLTFHLLKDVP H. vulgare
237 ERPLFQLVSAAQTILPDSEGAIDGHLREVGLTFHLLKDVP Z. mays C2
236 ERPLFQLVSAAQTILPDSEGAIDGHLREVGLAFHLLKDVP Z. mays WHP
236 ERPLFQMVSASQTILPDSEGAIDGHLREVGLTFHLLKDVP O. sativa
233 ERPLFQIVTAAQTLLPDSHGAIDGHLREVGLTFHLLKDVP A. majus
239 EKPIFEMVSAAQTILPDSDGAIDGHLREVGLTFHLLKDVP A. thaliana
236 ERPLFEMVSAAQTILPDSEGAIDGHLREVGLTFHLLKDVP G. hybrida CHS1
239 ERPLFEMVSAAQTILPDSEGAIDGHLKEVGLTFHLLKDVP G. hybrida CHS3
233 EKPLFELVSAAQTILPDSEGAIDGHLREVGLTFHLLKDVP H. androsaemum
233 ERPLFELVSAAQTLLPDSHGAIDGHLREVGLTFHLLKDVP P. hybrida CHS A
233 ERPLFELVSAAQTILPDSEGAIDGHLREVGLTFHLLKDVP V. vinifera CHS
233 ENPLFEIVSTDQQLVPNSHGAIGGLLREVGLTFYLNKSVP A. hypogaea STS
236 EKACFEIVWTAQTVVPNSEGAIGGKVREVGLTFQLKGAVP P. sylvestris DPS
233 ERPLFQLVSAAQTFIPNSAGAIAGNLREVGLTFHLWPNVP V. vinifera STS
238 ERPIFEIVSTDQTILPDTEKAVKLHLREGGLTFQLHRDVP G. hybrida 2PS
237 ERAVFNILSASQTIVPGSDGAITAHFYEMGMSYFLKEDVI H. androsaemum BPS
237 EHPLYQIMSASQIIVADSEGAIDGHLRQEGLTFHLRKDVP H. macrophylla CTAS
*
275 GLISKNIDKSLVEAFEPL------GISDWNSLFWIAHPGG A. andraeanum
276 GLISKNIERALEEAFKPL------GIDHWNSVFWIAHQGG H. vulgare
277 GLISKNIGRALDDAFKPL------GISDWNSIFWVAHPGG Z. mays C2
276 GLISKNIERALEDAFEPL------GISDWNSIFWVAHPGG Z. mays WHP
276 GLISKNIERALGDAFTPL------GISDWNSIFWVAHPGG O. sativa
273 GLISKNIEKSLKEAFDPL------GISDWNSVFWIAHPGG A. majus
279 GLISKNIVKSLDEAFKPL------GISDWNSLFWIAHPGG A. thaliana
276 GLISKNIEKALTTAFSPL------GINDWNSIFWIAHPGG G. hybrida CHS1
279 ALIAKNIEKALIQAFSPL------NINDWNSIFWIAHPGG G. hybrida CHS3
273 GLISKNVEKSLTEAFKPL------GISDWNSLFWIAHPGG H. androsaemum
273 GLISKNIEKSLEEAFRPL------SISDWNSLFWIAHPGG P. hybrida CHS A
273 GLISKNIEKSLVEAFTPI------GISDWNSLFWIAHPGG V. vinifera CHS
273 DIISQNINGALSKAFDPL------GISDYNSIFWIAHLGG A. hypogaea STS
276 DLISANIENCMVEAFSQF------KISDWNKLFWVVHPGG P. sylvestris DPS
273 TLISENIEKCLTQAFDPL------GISDWNSLFWIAHPGG V. vinifera STS
278 LMVAKNIENAAEKALSPL------GITDWNSVFWMVHPGG G. hybrida 2PS
277 PLFRDNIAAVMEEAFSPL------GVSDWNSLFYSIHPGG H. androsaemum BPS
277 SLVSDNIENTLVEAFTPILMDSIDSIIDWNSIFWIAHPGG H. macrophylla CTAS
*
309 PAILDQVEDKLRLQSEKLGATRRVLSEYGNMSSACVHFIL A. andraeanum
310 PAILDMVEAKVNLNKERMRATRHVLSEYGNMSSACVLFIM H. vulgare
311 PAILDQVEAKVGLDKARMRATRHVLSEYGNMSSACVLFIL Z. mays C2
310 PAILDQVEAKVGLDKERMRATRHVLSEYGNMSSACVLFIL Z. mays WHP
310 PAILDQVEAKVGLDKERMRATRHVLSEYGNMSSACVLFIL O. sativa
307 PAILDQVEEKLGLKPEKLRSTRQVLSEYGNMSSACVLFIL A. majus
313 PAILDQVEIKLGLKEEKMRATRHVLSEYGNMSSACVLFIL A. thaliana
310 PAILDQVELKLGLKEEKLRATRHVLSEYGNMSSACVLFII G. hybrida CHS1
313 PAILDQVEFKLGLREEKLRASRHVLSEYGNMSSACVLFIL G. hybrida CHS3
307 PAILDQVEAKLSLKPEKLRATRHVLSEYGNMSSACVLFIL H. androsaemum
307 PAILDQVEIKLGLKPEKLKATRNVLSNYGNMSSACVLFIL P. hybrida CHS A
307 PAILDQVELKLGLKEEKLRATRHVLSEYGNMSSACVLFIL V. vinifera CHS
307 RAILDQVEQKVNLKPEKMKATRDVLSNYGNMSSACVFFIM A. hypogaea STS
310 RAILDRVEAKLNLDPTKLIPTRHVMSEYGNMSSACVHFIL P. sylvestris DPS
307 PAILDAVEAKLNLEKKKLEATRHVLSEYGNMSSACVLFIL V. vinifera STS
312 RAILDQVERKLNLKEDKLRASRHVLSEYGNLISACVLFII G. hybrida 2PS
311 RGIIDGVAGNLGIKDENLVATRHVLGEYGNMGSACVMFIL H. androsaemum BPS
317 PAILNQVQAKVGLKEEKLRVSRHILSEYGNMSSACVFFIM H. macrophylla CTAS



CHS/STS signature motif

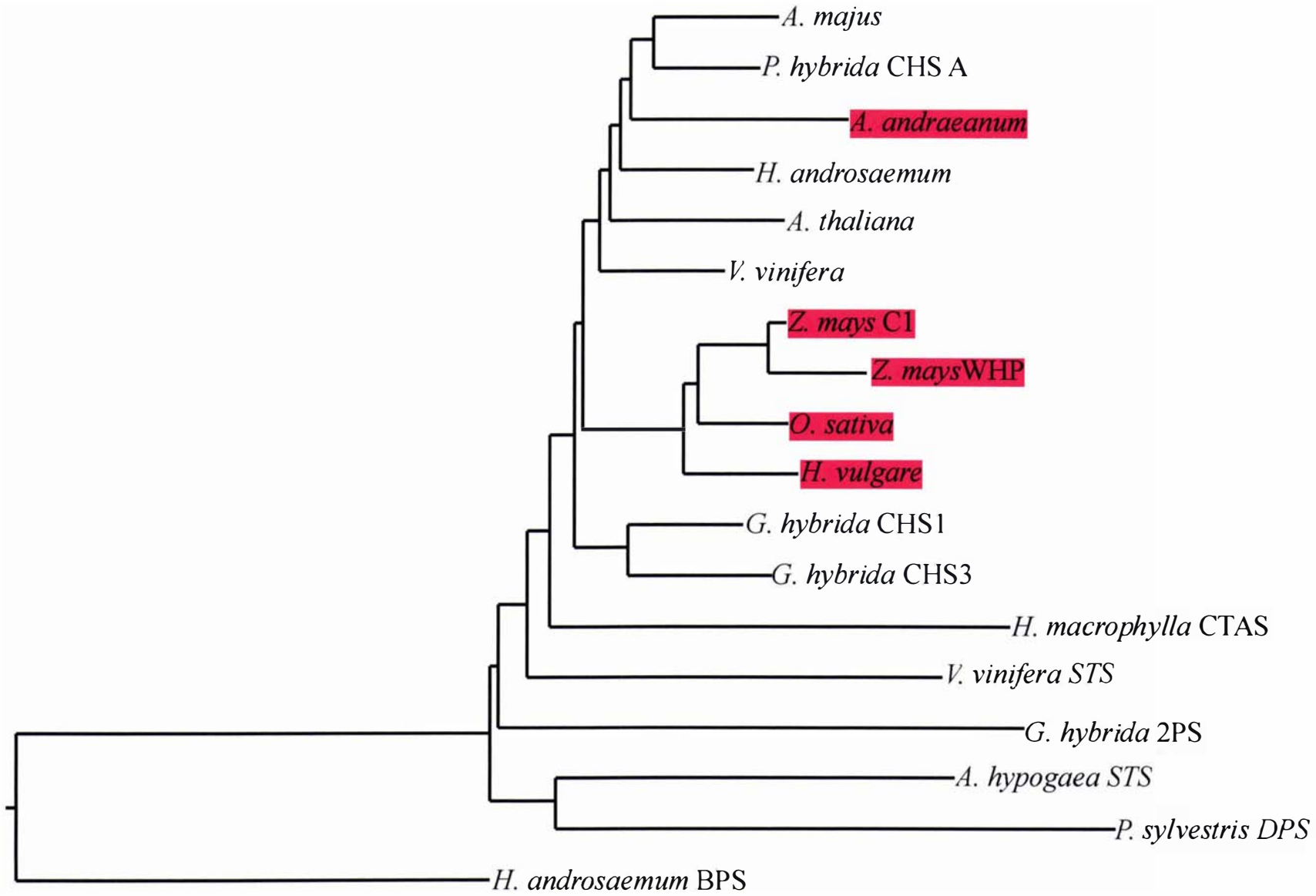
deduced amino acid encoded by *AaCHS1* with functionally defined CHS clones from monocots and dicots, and with other polyketide synthases. From the alignment, AaCHS1 had highest % identity to CHSs, with an average of 83% identity to dicot sequences and 79% to monocot sequences. The % identity falls to an average of 70% for STS and 65.9% for other CHS-related proteins. These averages are supported by the phylogenetic distribution represented in Figure 3.2. AaCHS1 showed a stronger phylogenetic relation to CHS proteins and within this category, it grouped preferentially with dicot sequences. Interestingly, a similar pattern was observed for the gerbera CHS proteins (CHS1 and CHS3), both of which have been functionally defined as CHS enzymes. In addition, AaCHS1, along with the functionally defined CHS proteins, was distinct from the closely related stilbene synthases as well as other CHS-related proteins that are members of the polyketide super family in plants.

3.3 ISOLATION OF AN ANTHURIUM *F3H* cDNA CLONE

Heterologous screening of the anthurium library with an antirrhinum full-length *F3H* cDNA insert gave eight independent positive signals in the primary screen, six of which were followed through to secondary and tertiary screens. Following in vivo excision and plasmid DNA preparation, five of these gave a fragment of the right size to be full-length clones (1.5-1.6 kb). In addition, all had the same restriction pattern suggesting they were derived from the same mRNA transcript (data not shown). These were sequenced with T3 and T7 primers generating sequence data for each clone from the 5' and 3' ends respectively. Upon analysis, three of the clones were found to encode identical putative *F3H* full-length clones based on nucleotide sequence and deduced amino acid sequences. The other two clones, although slightly shorter than full-length, also appeared to have the identical nucleotide sequence to the three full-length clones. One putative full-length cDNA was selected and named *AaF3H1*. The nucleotide sequence for the open reading frame is presented in Appendix III.

The *AaF3H1* full-length cDNA clone was 1.5 kb in length with an estimated G+C content of 59.8%. The cDNA clone contained 217 bp of 5' untranslated sequence, 257 bp of 3' untranslated sequence and a poly $(A)^+$ extension of 85 bp that made sequencing from the 3'

Figure 3.2. Phylogenetic relationship of AaCHS1 with confirmed CHS proteins as well as other polyketide synthases. Comparisons are made with confirmed CHS proteins from dicots and monocots. The monocot species are highlighted in red. The phylogenetic tree is based on alignments of full protein coding sequences. The alignment was done using the MegAlign programme (gap penalty 10, gap length penalty 10) from the DNASTAR sequence analysis software suite. The CHS sequences and their accession numbers are: *Anthurium andraeanum* (this work), *Hordeum vulgare* (P26018), *Oryza sativa* (X89859), *Zea mays* C2 and WHP (P24825, X60204 respectively), *Antirrhinum majus* (P06515), *Arabidopsis thaliana* (P13114), *Gerbera hybrida* CHS1 and CHS3 (Z38096 and Z38098 respectively), *Hypericum androsaemum* (AF315345), *Petunia hybrida* CHS-A (AF233638), *Vitis vinifera* (AB015872). The STS sequences are: *Arachis. hypogaea* STS (A00769), *Pinus sylvestris* DPS- dihydropinosylvin synthase (X60753), *Vitis vinifera* STS (P28343). Other CHS-related proteins: *Gerbera hybrida* 2PS- 2-pyrone synthase (Z38097), *Hypericum androsaemum* BPS- benzophenone synthase (AF352395) and *Hydrangea macrophylla* CTAS-p-coumaroyltriacetic acid synthase.


A. majus
P. hybrida CHS A
A. andraeanum
H. androsaemum
A. thaliana
V. vinifera
Z. mays C1
Z. maysWHP
O. sativa
H. vulgare
G. hybrida CHS1
G. hybrida CHS3
H. macrophylla CTAS
V. vinifera STS
G. hybrida 2PS
A. hypogaea STS
P. sylvestris DPS
H. androsaemum BPS

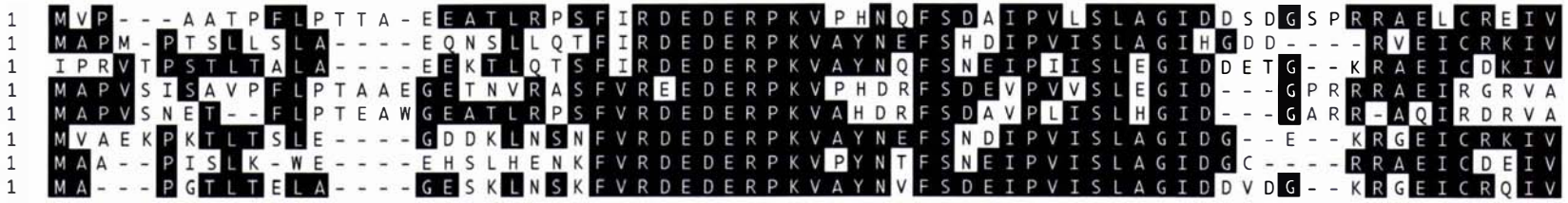
terminus difficult. Two conserved polyadenylation signals (AATAAT, Joshi, 1987b) were located downstream of the stop codon at positions 1154 and 1263 of the full-length sequence.

The longest open reading frame was 1115 bp and encoded a 370 amino acid deduced polypeptide with a predicted mass of 40.9 kD and an estimated isoelectric point of 5.3. The open reading frame began with the first ATG start codon at position 217 and terminated at codon TGA at position 1327.

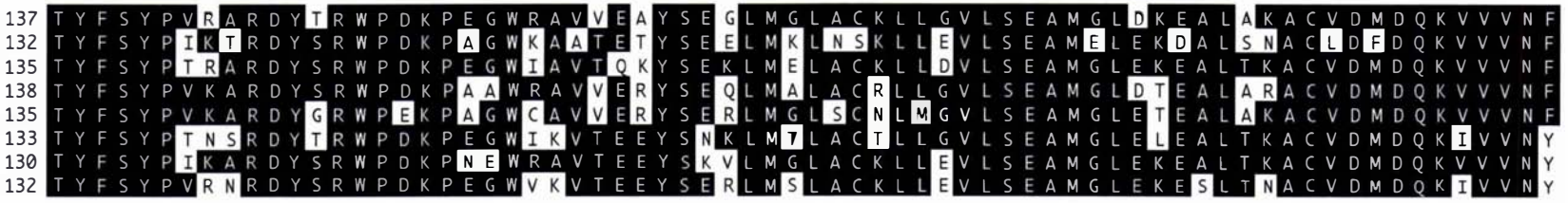
The *AaF3H1* cDNA clone hybridised strongly to a heterologous *F3H* probe and when BLAST searches were performed with either the cDNA sequence or the deduced polypeptide sequence, the best hits were always for F3H proteins, with significantly reduced BLAST scores for other related proteins. When the amino acid sequence of AaF3H1 was compared with the sequences of confirmed F3H clones, high % identities were observed including maize (76 %), barley (74%) and arabidopsis (77%). The amino and carboxyl terminals show some variation in length and composition whereas very high similarities were observed for residues 25-324 (Figure 3.3).

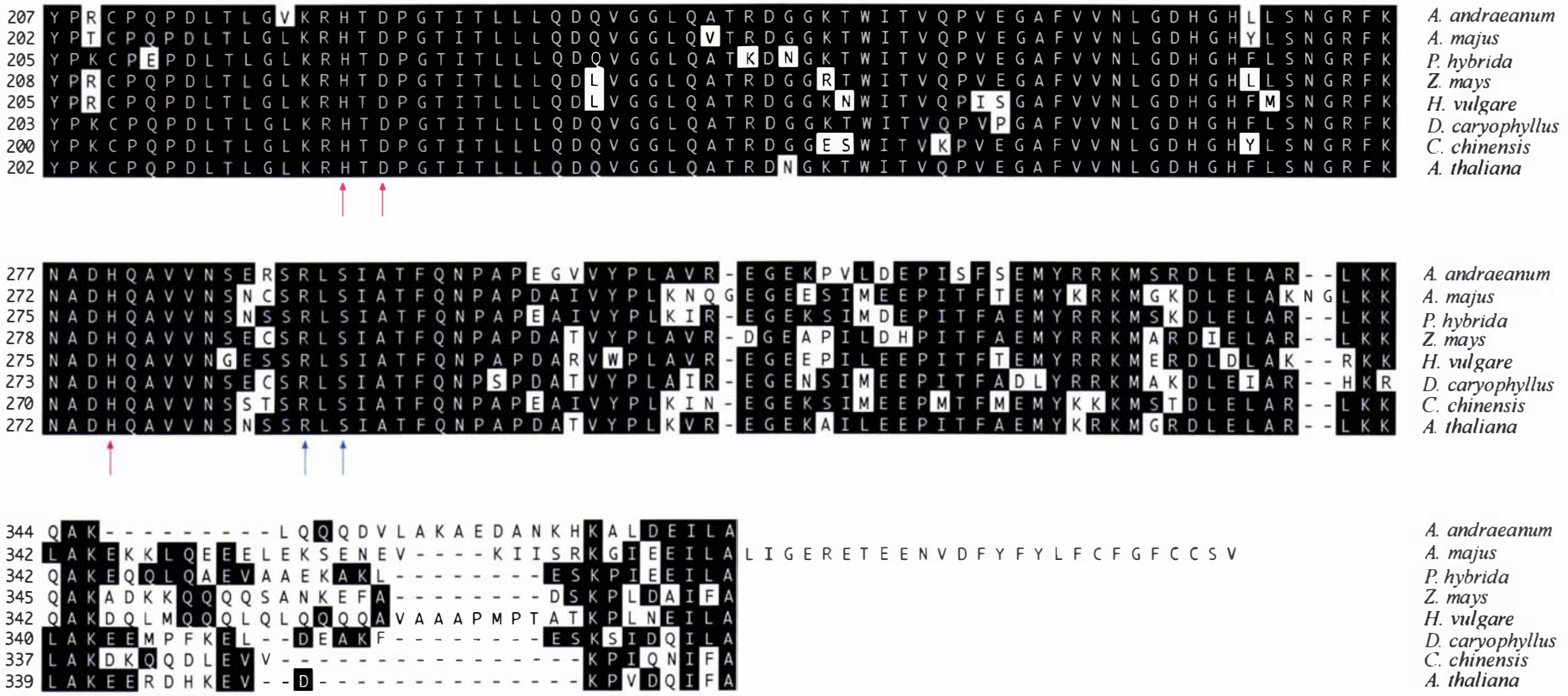
With the exception of antirrhinum F3H, the translational stop site for all F3H clones in the alignment occurred at the same position, suggesting a very conserved C-terminal region. (Figure 3.3). Several structural motifs of functional significance were also conserved in AaF3H1. Conserved histidine and aspartate residues have been observed in F3H proteins and in other members of the 2-oxoglutarate family. Britsch et al. (1993) have identified two histidines and one aspartate as part of the putative iron-binding site. An arginine and a serine have been implicated as part of the 2-oxoglutarate binding site (Lukacin and Britsch 1997; Lukacin et al., 2000). All these residues were found conserved in AaF3H1. The arginine-Xaa-Serine (R-X-S) motif has also been suggested to be essential for the binding site of 2-oxoglutarate enzymes (Britsch et al., 1993). This motif was also found in AaF3H1 (Figure 3.3).

Figure 3.3. Multiple alignment of the deduced amino acid sequence of AaF3H1 with those of confirmed F3H proteins. Functionality was determined either by in vitro expression or northern analysis of wild type and F3H mutants. The alignment was done using the Clustal Method (Higgins and Sharp, 1989). Black shading indicates identical amino acids. Dashes indicate gaps in the sequence for alignments. The conserved leucine rich motif is marked and labeled. Arrows point to amino acid residues that form the conserved active sites. The putative iron binding sites are marked with red arrows while the putative 2-oxoglutarate binding site (R-X-S motif) is marked with blue arrows. F3H amino acid sequences in the alignment sequences in the alignment and their accession numbers: *Zea mays* (UO4434), *Hordeum vulgare* (P28038), *Anthurium andraeanum* (this work), *Dianthus caryophyllus* (Q05964), *Arabidopis thaliana* (U33932), *Petunia hybrida* (Q07353), *Callistephus chinensis* Q05963) and *Antirrhinum majus* (Martin et al., 1991).



conserved leucine rich region suggested to form an alpha-helix




207 YPRCPQPDLTLGVKRHTDPGTITLLLQDQVGGLQATRDGGKTWITVQPVEGAFVVNLGDHGHLLSNGRFK
202 YPTCPQPDLTLGLKRHTDPGTITLLLQDQVGGLQVTRDGGKTWITVQPVEGAFVVNLGDHGHYLSNGRFK
205 YPKCPEPDLTLGLKRHTDPGTITLLLQDQVGGLQATKDNGKTWITVQPVEGAFVVNLGDHGHFLSNGRFK
208 YPRCPQPDLTLGLKRHTDPGTITLLLQDLVGGLQATRDGGRTWITVQPVEGAFVVNLGDHGHLLSNGRFK
205 YPRCPQPDLTLGLKRHTDPGTITLLLQDLVGGLQATRDGGKNWITVQPISGAFVVNLGDHGHFMSNGRFK
203 YPKCPQPDLTLGLKRHTDPGTITLLLQDQVGGLQATRDGGKTWITVQPVPGAFVVNLGDHGHFLSNGRFK
200 YPKCPQPDLTLGLKRHTDPGTITLLLQDQVGGLQATRDGGESWITVKPVEGAFVVNLGDHGHYLSNGRFK
202 YPKCPQPDLTLGLKRHTDPGTITLLLQDQVGGLQATRDNGKTWITVQPVEGAFVVNLGDHGHFLSNGRFK
A. andraeanum
A. majus
P. hybrida
Z. mays
H. vulgare
D. caryophyllus
C. chinensis
A. thaliana
277 NADHQAVVNSERSRLSIATFQNPAPEGVVYPLAVR-EGEKPVLDEPISFSEMYRRKMSRDLELAR--LKK
272 NADHQAVVNSNCSRLSIATFQNPAPDAIVYPLKNQGEGEESIMEEPITFTEMYKRKMGKDLELAKNGLKK
275 NADHQAVVNSNSSRLSIATFQNPAPEAIVYPLKIR-EGEKSIMDEPITFAEMYRRKMSKDLELAR--LKK
278 NADHQAVVNSECSRLSIATFQNPAPDATVYPLAVR-DGEAPILDHPITFAEMYRRKMARDIELAR--LKK
275 NADHQAVVNGESSRLSIATFQNPAPDARVWPLAVR-EGEEPILEEPITFTEMYRRKMERDLDLAK--RKK
273 NADHQAVVNSECSRLSIATFQNPSPDATVYPLAIR-EGENSIMEEPITFADLYRRKMAKDLEIAR--HKR
270 NADHQAVVNSSTSRLSIATFQNPAPEAIVYPLKIN-EGEKSIMEEPMTFMEMYKKKMSTDLELAR--LKK
272 NADHQAVVNSNSSRLSIATFQNPAPDATVYPLKVR-EGEKAILEEPITFAEMYKRKMGRDLELAR--LKK
A. andraeanum
A. majus
P. hybrida
Z. mays
H. vulgare
D. caryophyllus
C. chinensis
A. thaliana
344 QAK----------LQQQDVLAKAEDANKHKALDEILA
342 LAKEKKLQEELEKSENEV----KIISRKGIEEILA
342 QAKEQQLQAEVAAEKAKL---------ESKPIEEILA
345 QAKADKKQQQQSANKEFA---------DSKPLDAIFA
342 QAKDQLMQQQLQLQQQQAVAAAPMPTATKPLNEILA
340 LAKEEMPFKEL--DEAKF---------ESKSIDQILA
337 LAKDKQQDLEV----------------KPIQNIFA
339 LAKEERDHKEV--D----------------KPVDQIFA
LIGERETEENVDFYFYLFCFGFCCSV

F3H proteins also have a conserved leucine rich region defined as LMXLACKLLGVLSEAMGLDKEAL. Britsch et al. (1993) suggested that this region might be involved in forming a α-helix essential for F3H function. This essential amino acid motif was also found in AaF3H1 (Figure 3.3).

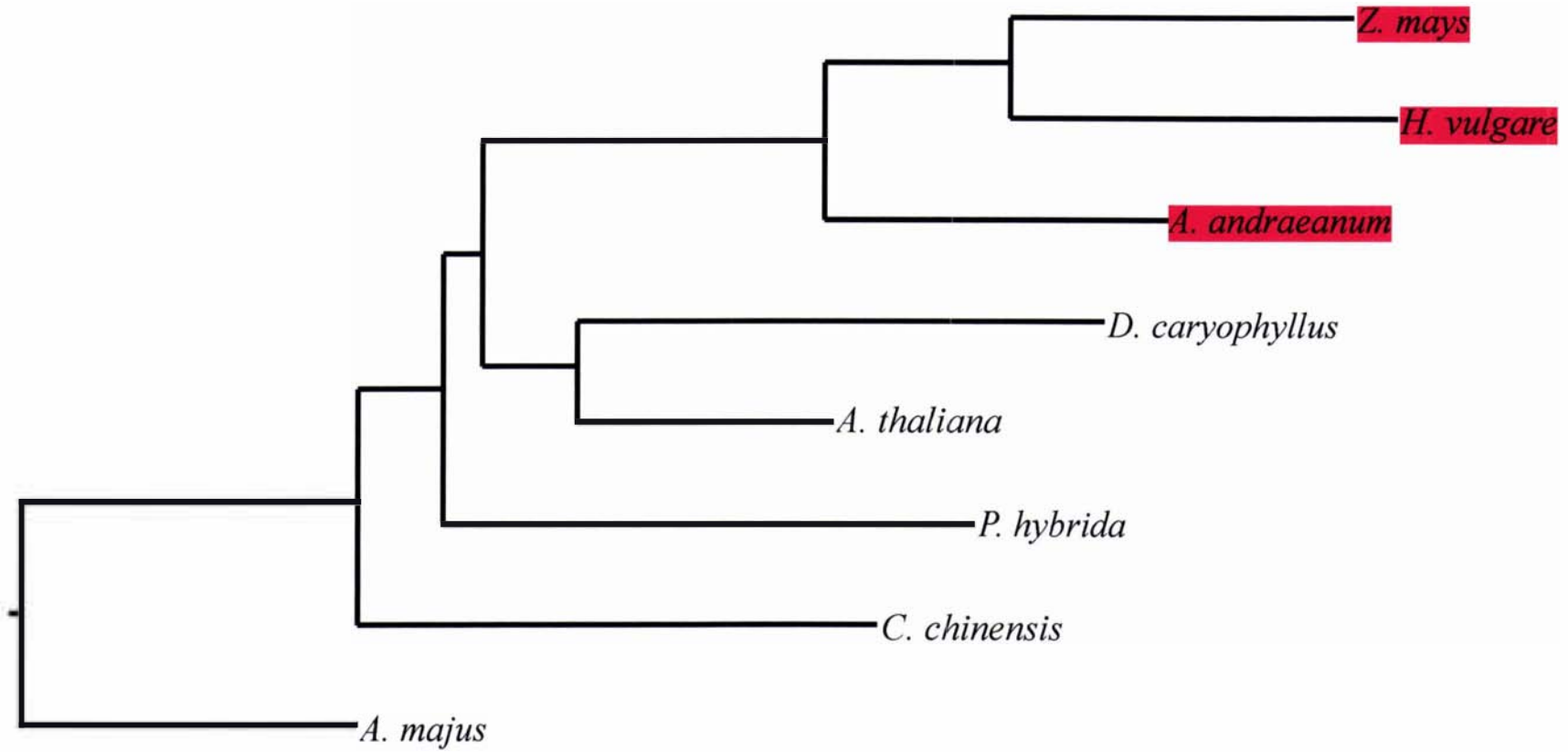
In addition, to the structural comparisons, the phylogenetic tree constructed based on multiple alignments among F3H sequences (Figure 3.4) showed that the anthurium F3H was more similar to monocot sequences than to dicot sequences.

3.4 ISOLATION OF AN ANTHURIUM *DFR* cDNA CLONE

Heterologous low stringency screening of the cDNA library with a full-length antirrhinum *DFR* cDNA insert resulted in six independent positive clones being isolated. All inserts were approximately 1.2 kb in length and appeared to be derived from the same transcript as evidenced by restriction mapping (data not shown). Four clones were selected and sequenced at the 5' and 3' ends with the T3 and T7 sequencing primers. Sequence analysis revealed that all the clones represent the same gene and only differed in length. One cDNA clone (DFR 7A) was just short of the putative start codon by 13 nucleotides. The remaining clones appeared to encode full-length sequences. One cDNA was chosen and named *AaDFR1*. The nucleotide sequence for the open reading frame is presented in Appendix IV.

AaDFR1 had an unusual 5' untranslated region of 155 bp that consisted of a -CACA- stretch of 40 bp beginning at position 40 in the clone, followed by a 12 bp -TATA- stretch and a 16 bp -GAGA- stretch. This was not a cloning or sequencing artifact, as it is was found in the 5' sequence data of four independent cDNA isolates that came from a non-amplified cDNA library. The open reading frame consisted of 1041 bp with the ATG start codon at position 156 and the stop codon, TGG, at position 1202. Interestingly, there were two successive ATG codons at the start of this clone. According to the sequence there were no further ATG sites upstream from this point in the clone that represented possible translational start sites. Joshi (1987b) proposed that one of two plant consensus sequences (AAAGAAATG or TAAACAATG) flank the ATG start codon. For *AaDFR1*, neither of these two sequences was found next to the two ATG start site. However, a -TAACA- site

Figure 3.4. Phylogenetic relationship of AaF3H1 to confirmed F3H proteins from other species. The monocot species are highlighted in red. The phylogenetic tree is based on alignments of the full protein coding sequences. The alignment was done using the MegAlign programme (gap penalty 10, gap length penalty 10) from the DNASTAR sequence analysis suite of software. Protein sequences used: *Zea mays* (UO4434), *Hordeum vulgare* (P28038), *Anthurium andraeanum* (this work), *Dianthus caryophyllus* (Q05964), *Arabidopis thaliana* (U33932), *Petunia hybrida* (Q07353), *Callistephus chinensis* (Q05963) and *Antirrhinum majus* (Martin et al., 1991).


Z. mays
H. vulgare
A. andraeanum
D. caryophyllus
A. thaliana
P. hybrida
C. chinensis
A. majus

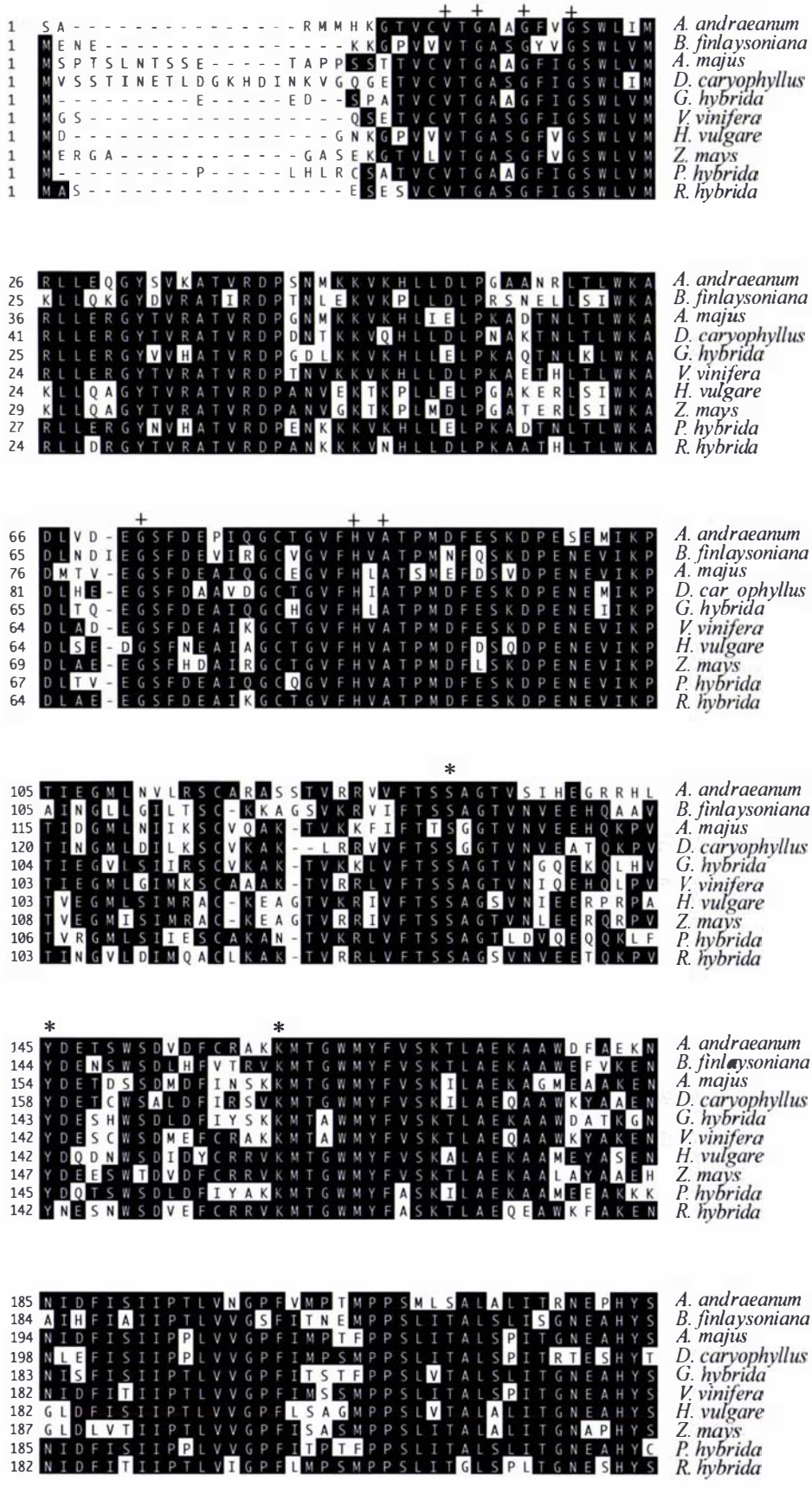
was located just 27 bp upstream from the putative ATG codons. A putative plant polyadenylation signal, AATAAA (Joshi, 1987b), was located downstream of the termination codon at position 1328 of the cDNA clone. The 3' untranslated region consisted of 183 nucleotides and a stretch of 18 adenylic acids constituted the poly $(A)^+$ tail. The G+C content of the clone was 55.6%.

The open reading frame encoded a deduced 347 amino acid polypeptide with predicted mass of 38.7 kD and an isoelectric point of 5.6. When the deduced amino acid sequence of AaDFR1 was compared to those of other species, high levels of conservation were found throughout the length of the polypeptide sequences, although conservation was higher towards the N-terminal end of the clones with significant deviations at the C-terminal ends (Figure 3.5). AaDFR1 shared an average sequence identity of 60%, with slightly higher overall levels of identity being shared with DFR from dicots (data not shown). While the phylogenetic tree shown in Figure 3.6, shows AaDFR1 having a common ancestor with monocot DFR the AaDFR1 polypeptide appears to have accumulated sufficient differences to distinguish it from the other monocot sequences available.

Residues that are conserved in the short chain dehydrogenase/reductase (SDR) family, to which DFR belongs (Baker and Blasco, 1992), were also found in anthurium DFR along with residues that are suggested to define the catalytic site for DFR (Figure 3.5).

Related (but unpublished) work in this thesis was conducted to determine if AaDFR1 could catalyse DHM (data not shown). Particle bombardment was used in an effort to complement DFR activity in a petunia line, BR140, that is dominant for *F3'5'H* and recessive for *F3'H* (supplied by Dr. Filippa Brugliera, Florigene, Australia). Successful complementation would have resulted in the accumulation of delphinidins. However, technical difficulties were encountered, that seemed tissue related, preventing any meaningful results from being obtained. Internal positive controls with GFP, as well as positive controls with a lisianthus *DFR* construct, were also unsuccessful.

Figure 3.5. Multiple alignment of the deduced amino acid sequence of AaDFR1 with those of confirmed DFR proteins. (N.B only DFR from B. *finlaysoniana* has not been confirmed). The alignment was done using the Clustal Method (Higgins and Sharp, 1989). Black shading indicates identical amino acids. Dashes indicate gaps in the sequence for alignments. The asterisks identify putative residues that form part of the active site of the DFR short chain dehydrogenase reductase (SDR) enzyme family to which DFR belongs. Protein residues marked with + are conserved in all SDR proteins. The amino acid sequences and their accession numbers are: *Anthurium andraeanum* (this work), *Bromheadia finlaysoniana* (AF007096), *Antirrhinum majus* (P14721), *Dianthus caryophyllus* (P51104), *Gerbera hybrida* (P51105), *Vitis vinifera* (P51110), *Hordeum vulgare* (P51106), *Zea mays* (P51108), *Petunia hybrida* (P14720) and *Rosa hybrida* (BAA12723).


A. andraeanum
B. finlaysoniana
A. majus
D. caryophyllus
G. hybrida
V. vinifera
H. vulgare
Z. mays
P. hybrida
R. hybrida

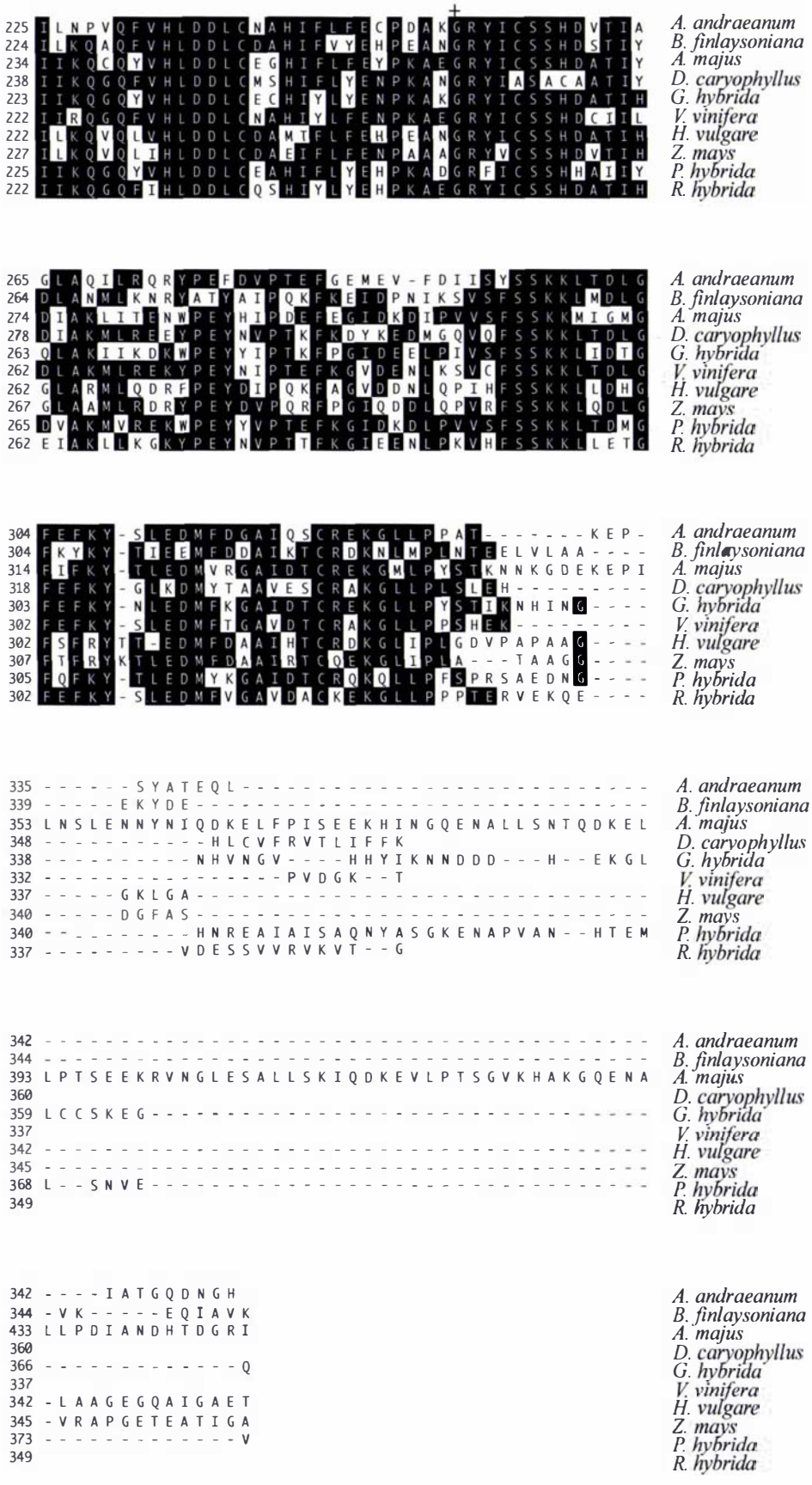
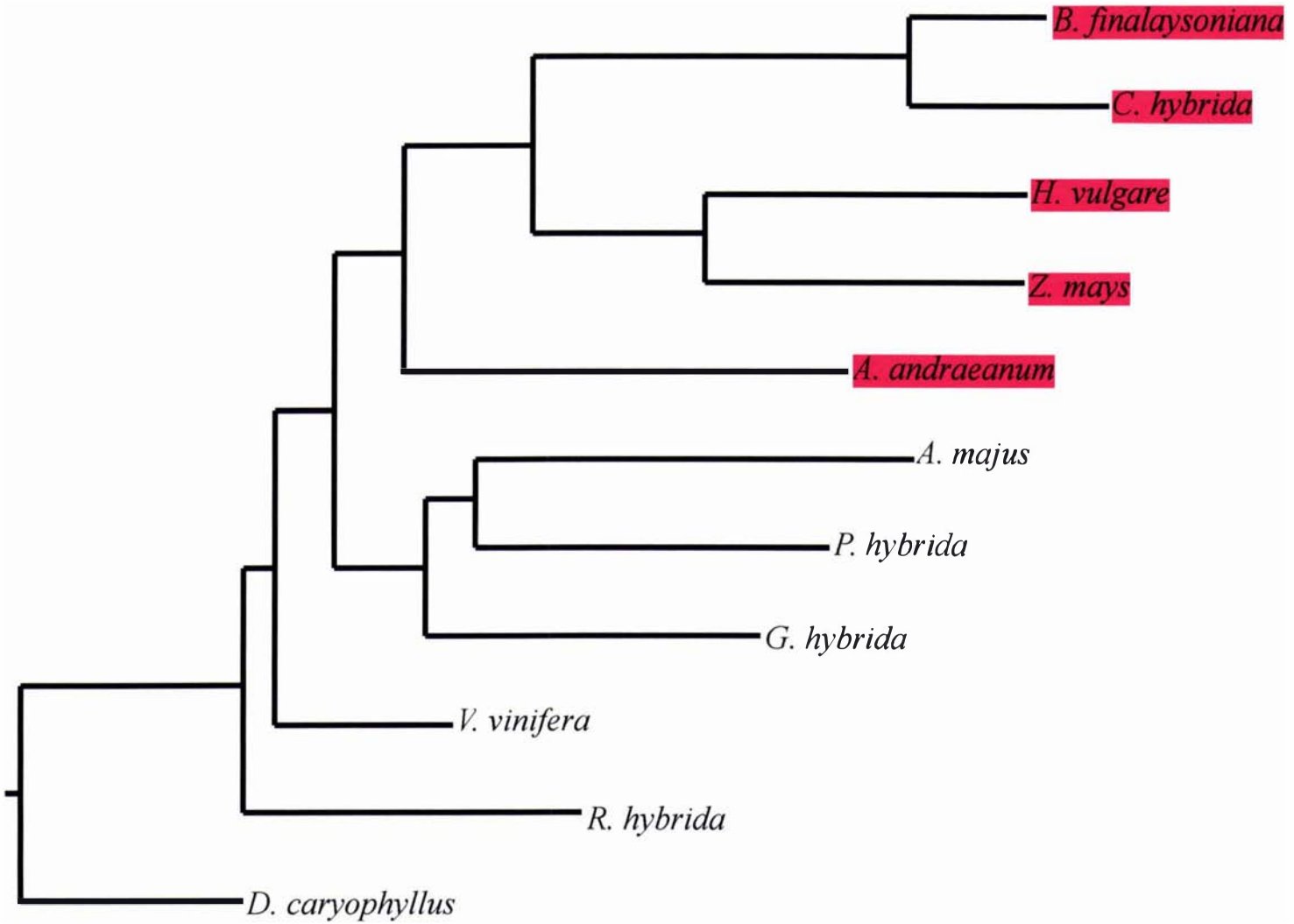

A. andraeanum
B. finlaysoniana
A. majus
D. caryophyllus
G. hybrida
V. vinifera
H. vulgare
Z. mays
P. hybrida
R. hybrida

Figure 3.6. Phylogenetic relationship of AaDFR1 to confirmed DFR proteins from other species. (N.B only DFR from B. *finlaysoniana* has not been confirmed). Comparisons are made between monocot and dicot species. The monocot species are highlighted in red. The phylogenetic tree is based on alignments of full protein coding sequences. The alignment was done using the MegAlign programme (gap penalty 10, gap length penalty 10) from the DNASTAR sequence analysis suite of software. The sequences and their accession numbers are: *Bromheadia finlaysoniana* (AF007096), *Cymbidium hybrida* (AF017451) *Hordeum vulgare* (P51106), *Zea mays* (P51108), *Anthurium andraeanum* (this work), *Antirrhinum majus* (P14721), *Petunia hybrida* (P14720), *Gerbera hybrida* (P51105), *Vitis vinifera* (P51110), *Rosa hybrida* (BAA12723), and *Dianthus caryophyllus* (P51104).


B. finalaysoniana
C. hybrida
H. vulgare
Z. mays
A. andraeanum
A. majus
P. hybrida
G. hybrida
V. vinifera
R. hybrida
D. caryophyllus

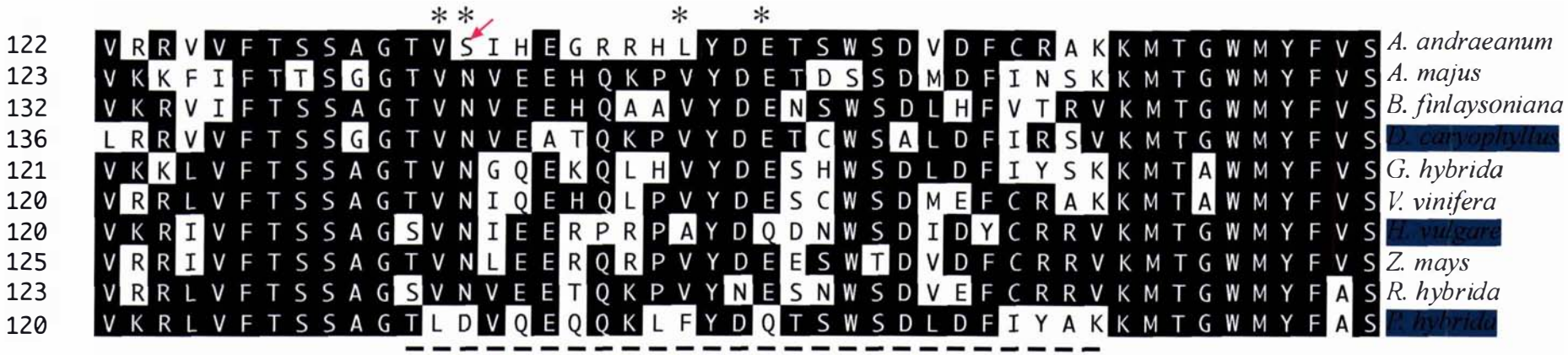
The residues that determine substrate specificity in DFR (Beld et al., 1989; Johnson et al., 2001) are shown in Figure 3.7. The four residues found primarily in DHK-accepting DFR but not in petunia DFR, which cannot utilise DHK, are marked with an asterisk. Of these four residues only two are conserved in AaDFR1. The glutamic acid-145 (E145) from gerbera, which was shown to be essential for DFR activity, is also conserved in AaDFR1 along with valine-133 (V133), the function of which has not been investigated. The remaining two residues are both conservative substitutions. In AaDFR1, a serine substitutes the conserved gerbera asparagine-134 (N134) and a leucine substitutes for valine-142 (V142) (Figure 3.7).

3.5 ISOLATION OF AN ANTHURIUM *ANS* cDNA CLONE

A full-length apple *ANS* cDNA probe was used for heterologous low stringency screening of the cDNA primary library. The primary screen gave seven independent positives. Of these, only two (*ANS 2* and *ANS 4*) gave positives through to secondary and tertiary screens. Both inserts were approximately 1.6 kb in length and had similar restriction patterns suggesting that they were derived from the same transcript (data not shown). Both clones were sequenced at their 5' and 3' ends using M13 reverse and M13 forward primers respectively. The sequence data showed that the clones were identical. One cDNA clone was chosen and named *AaANS1*. The nucleotide sequence for the open reading frame is presented in Appendix V.

The full-length *AaANS1* cDNA was 1.5 kb in length with a G-C content of 61% and 63 bp of 5' untranslated sequence. There were two ATG codons 12 bp apart at the start of the clone. The first ATG was at position 64 and the second at position 79. A stop codon (TAG) was located 1156 bp into the clone and marks the end of the longest open reading frame of 1092 bp. The 3' untranslated region was 328 bp in length and two putative polyadenylation signals were noted (Joshi, 1987b). The first, AATAAA, was 191 bp downstream of the stop codon, and the second was 278 bp downstream from the stop codon. A poly $(A)^+$ tail of 17 adenylic acid residues marked the end of the cDNA clone.

Figure 3.7. Amino acid alignment showing substrate specificity region for DFR. This region of DFR proteins is seen sandwiched between two highly conserved domains. The four residues, which are absent in DHK-accepting DFR, are marked by asterisks. In anthurium a serine residue, marked with a red arrow, replaces the N134, which is believed to determine substrate specificity in gerbera DFR. No residues specific to DHM-accepting DFRs, which are shaded in blue, are observed in this region. The sequences and their accession numbers are: *Anthurium andraeanum* (this work), *Antirrhinum majus* (P14721), *Bromheadia finlaysoniana* (AF007096), *Dianthus caryophyllus* (P51104), *Gerbera hybrida* (P51105), *Vitis vinifera* (P51110), *Hordeum vulgare* (P51106), *Zea mays* (P51108), *Petunia hybrida* (P14720) and *Rosa hybrida* (BAA12723).



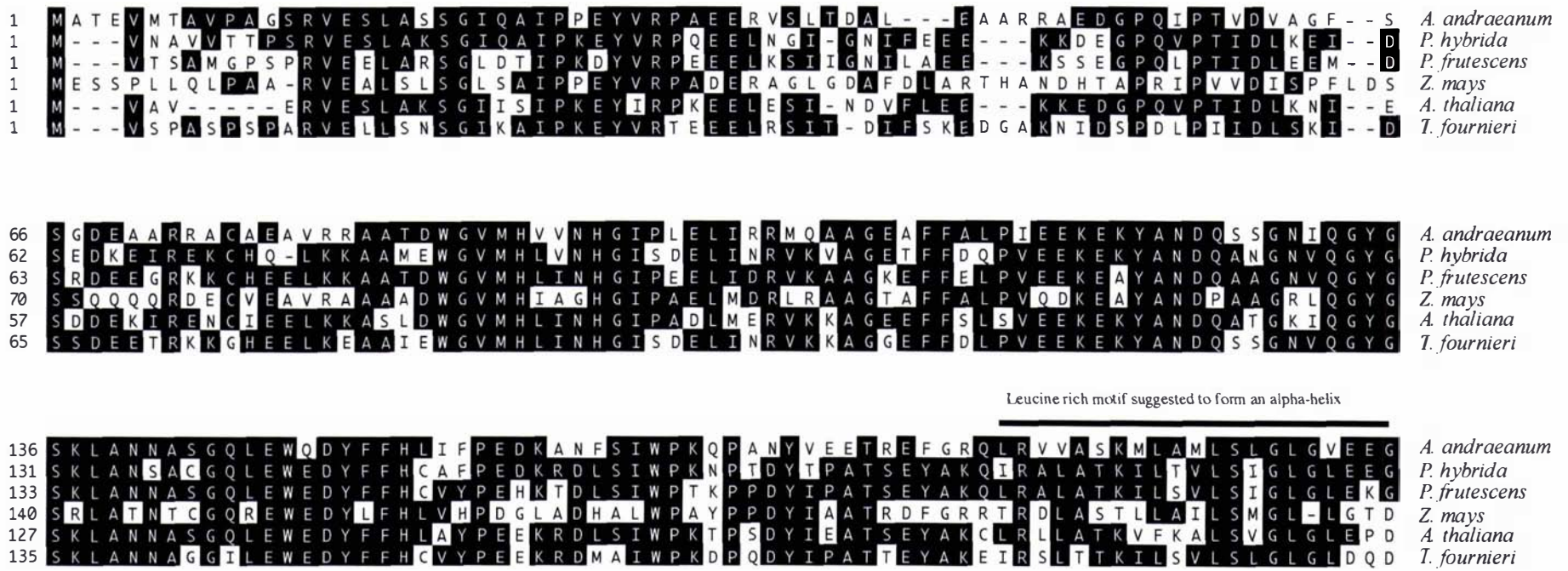
Proposed DFR substrate specificty region

The 1092 bp open reading frame encoded a 364 deduced polypeptide with a predicted mass of 41.7 kD and an isoelectric point of 5.7. The *AaANS1* cDNA clone hybridised strongly to a heterologous ANS probe and its expression pattern supported its involvement in anthocyanin biosynthesis (discussed in more detail in Chapter 4). In addition, when a BLAST search was performed with either the cDNA sequence or the deduced polypeptide sequence, the best hits were always for ANS proteins with significantly reduced BLAST scores for other related proteins. All the residues that define the highly conserved catalytic sites for the 2-oxoglutarate- dependent dioxygenase enzyme family, to which ANS belongs, were also present in the putative anthurium ANS clone. From the protein alignment with other deduced ANS polypeptide sequences from various plant species (Figure 3.8), AaANS1 averaged 60% identity. Surprisingly the lowest % identity was with the monocot ANS sequence from maize (55%). Consequently, the phylogenetic analysis did not show a very strong grouping of AaANS1 with maize (Figure 3.9).

3.6 SCREENING FOR AN ANTHURIUM *FNS* cDNA CLONE

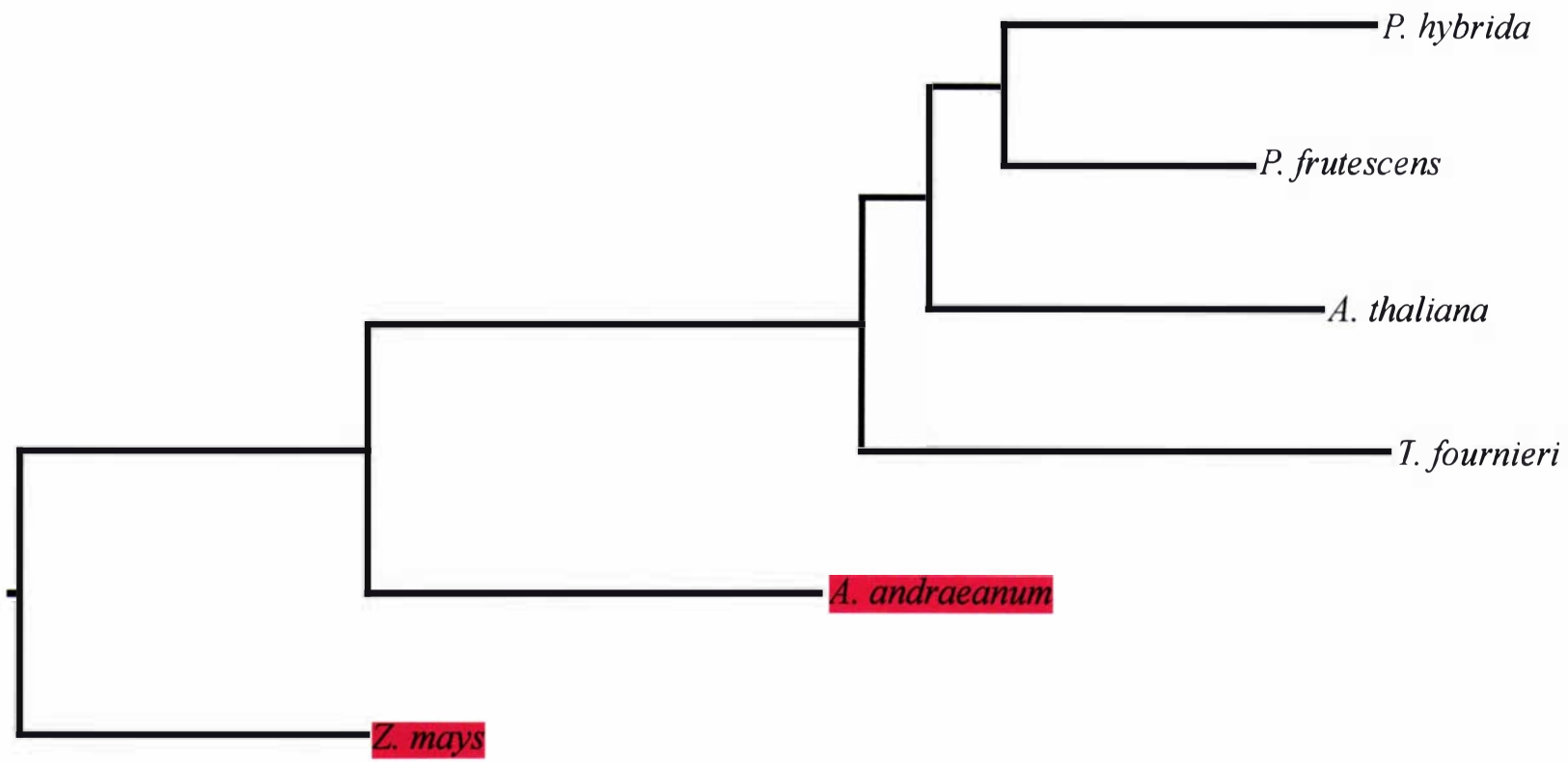
Flavone *C*-glycosides are the primary non-anthocyanin flavonoids accumulating in anthurium spathe tissue. Therefore, isolating a cDNA clone for *FNS* was an obvious target *to assist characterisation of flower colour development in anthurium. An antirrhinum FNS II* clone was isolated in our laboratory (Dr. Kevin Davies, Crop & Food Research, New Zealand, personal communication). The full-length antirrhinum *FNS* insert was used as a probe to screen an aliquot of 20,000 recombinant phage particles at 55 °C. The primary screen gave four independent positive signals of which only three (*FNS 2, 3* and *8*) went through as positive secondary and tertiary signals to give single plaque picks for in vivo excision and subsequent plasmid DNA extraction. An *Eco* RI/*Xho* I digest of the three clones followed by gel electrophoresis on a 1% (w/v) agarose gel revealed a 1.6 kb band for *FNS 3*, whereas both *FNS 2* and *8* showed an identical restriction pattern, migrating as a doublet of approximately 1.8 kb (data not shown). All three clones were sequenced with M13 reverse and forward primers.

Figure 3.8. Multiple alignment of the deduced amino acid sequence of AaANS1 with those of confirmed ANS proteins. All clones used for comparison with AaANS1 have been functionally confirmed either by in vitro expression in yeast or by northern analysis of wild type plants. The CLUSTAL Method by Higgins and Smart (1989) was used to perform the alignment. Black shading indicates identical amino acids. Dashes indicate gaps in the sequence for alignments. The conserved leucine rich motif is marked and labeled. Arrows point to amino acid residues that form the conserved active sites. Histidine and aspartate (marked with red arrows) are required for ferrous ion binding while arginine and serine (marked with blue arrows) are probable binding sites for 2-oxoglutarate. Proteins and their accession numbers: *Anthurium andraeanum* (this work), *Petunia hybrida* (P51092), *Perilla fructescens* (AB003779), *Zea mays* (P41213), *Arabidopsis thaliana* (U70478) and *Torenia fournieri* (AB044091).


1 MATEVMTAVPAGSRVESLASSGIQAIPPEYVRPAEERVSLTDAL---EAARRAEDGPQIPTVDVAGF--S A. andraeanum
1 M---VNAVVTTPSRVESLAKSGIQAIPKEYVRPQEELNGI-GNIFEEE---KKDEGPQVPTIDLKEI--D P. hybrida
1 M---VTSAMGPSPRVEELARSGLDTIPKDYVRPEEELKSIIGNILAEE---KSSEGPQLPTIDLEEM--D P. frutescens
1 MESSPLLQLPAA-RVEALSLSGLSAIPPEYVRPADERAGLGDAFDLARTHANDHTAPRIPVVDISPFLDS Z. mays
1 M---VAV-----ERVESLAKSGIISIPKEYIRPKEELESI-NDVFLEE---KKEDGPQVPTIDLKNI--E A. thaliana
1 M---VSPASPSPARVELLSNSGIKAIPKEYVRTEEELRSIT-DIFSKEDGAKNIDSPDLPIIDLSKI--D T. fournieri
66 SGDEAARRACAEAVRRAATDWGVMHVVNHGIPLELIRRMQAAGEAFFALPIEEKEKYANDQSSGNIQGYG A. andraeanum
62 SEDKEIREKCHQ-LKKAAMEWGVMHLVNHGISDELINRVKVAGETFFDQPVEEKEKYANDQANGNVQGYG P. hybrida
63 SRDEEGRKKCHEELKKAATDWGVMHLINHGIPEELIDRVKAAGKEFFELPVEEKEAYANDQAAGNVQGYG P. frutescens
70 SSQQQQRDECVEAVRAAAADWGVMHIAGHGIPAELMDRLRAAGTAFFALPVQDKEAYANDPAAGRLQGYG Z. mays
57 SDDEKIRENCIEELKKASLDWGVMHLINHGIPADLMERVKKAGEEFFSLSVEEKEKYANDQATGKIQGYG A. thaliana
65 SSDEETRKKGHEELKEAAIEWGVMHLINHGISDELINRVKKAGGEFFDLPVEEKEKYANDQSSGNVQGYG T. fournieri
Leucine rich motif suggested to form an alpha-helix
136 SKLANNASGQLEWQDYFFHLIFPEDKANFSIWPKQPANYVEETREFGRQLRVVASKMLAMLSLGLGVEEG A. andraeanum
131 SKLANSACGQLEWEDYFFHCAFPEDKRDLSIWPKNPTDYTPATSEYAKQIRALATKILTVLSIGLGLEEG P. hybrida
133 SKLANNASGQLEWEDYFFHCVYPEHKTDLSIWPTKPPDYIPATSEYAKQLRALATKILSVLSIGLGLEKG P. frutescens
140 SRLATNTCGQREWEDYLFHLVHPDGLADHALWPAYPPDYIAATRDFGRRTRDLASTLLAILSMGL-LGTD Z. mays
127 SKLANNASGQLEWEDYFFHLAYPEEKRDLSIWPKTPSDYIEATSEYAKCLRLLATKVFKALSVGLGLEPD A. thaliana
135 SKLANNAGGILEWEDYFFHCVYPEEKRDMAIWPKDPQDYIPATTEYAKEIRSLTTKILSVLSLGLGLDQD T. fournieri


206 K - - - L E A E V G G M - - - - - - - - E D L L L Q M K I N Y Y P R C P Q P E L A V G V E A H T D V S A L S F I L H N M V P G L Q V H N E G Q
201 R - - - L E K E V G G M - - - - - - - - E D L L L Q M K I N Y Y P K C P Q P E L A L G V E A H T D V S A L T F I L H N M V P G L Q L F Y E G Q
203 R - - - L E K E V G G A - - - - - - - - E D L I V Q M K I N F Y P K C P Q P E L A L G W E A H T D V S A L T F I L H N M V P G L Q L F Y E D K
209 R G D A L E K A L T T T T T R T A A D D D L L L Q L K I N Y Y P R C P Q P E L A V G V E A H T D V S A L S F I L H N G V P G L Q V L H G A R
197 R - - - L E K E V G G L - - - - - - - - E E L L L Q M K I N Y Y P K C P Q P E L A L G V E A H T D V S A L T F I L H N M V P G L Q L F Y E G K
205 R - - - L E K E V G G K - - - - - - - - D D L T L Q M K I N Y Y P K C P Q P E L A L G V E A H T D V S A L T F I L H N M V P G L Q V L Y K G E
A. andraeanum
P. hybrida
P. frutescens
Z. mays
A. thaliana
T. fournieri
266 W V T A R C V P D S I I M H V G D V V E I L S N G Q Y K S I L H R G L V N K E K V R V S W A V F C E P P R D K I L L Q P L P D I L G Q G H P
261 W V T A K C V P N S I I M H I G D T I E I L S N G K Y K S I L H R G V V N K E K V R F S W A I F C E P P K E K I I L K P L P E T V T E A E P
263 W V T A K C V P N S I I M H I G D T L E I L S N G K Y K S I L H R G L V N K E K V R I S W A V F C E P P K E K I V L Q P L P E T V S E V E P
279 W V T A R H E P G T I I V H V G D A L E I L S N G R Y T S V L H R G L V N R E A V R I S W V V F C E P P P D S V L L H P L P E L V T E G H P
257 W V T A K C V P D S I V M H I G D T L E I L S N G K Y K S I L H R G L V N K E K V R I S W A V F C E P P K D K I V L K P L P E M V S V E S P
265 W V T A S C V P D S I I L H V G D T I E I L S N G M Y K S V L H R G L V N R E R V R V S W A V F C E P P K E K I V L K P L P E T V G E D R P
A. andraeanum
P. hybrida
P. frutescens
Z. mays
A. thaliana
T. fournieri
336 A Q F P P R T F S Q H I Q H K L F R K T Q G D - - - - - - - F -
331 P R F P P R T F A Q H M A H K L F R K D D K D A A V E H K V F N E D E L D T A A E H K V L K K D N Q D A V A E N K D I K E D E Q C G P A E H
333 P R F P P R T F A Q H L K H K L F R K T D - - - - - - - - - - - - - G D L D -
349 A R F T P R T F K Q H L D R K L F K K K Q - Q H
327 A K F P P R T F A Q H I E H K L F G K E Q -
335 A L F P P R T F A Q H M K H K L F K K S D D E - V H D E E G V V V V E H
A. andraeanum
P. hybrida
P. frutescens
Z. mays
A. thaliana
T. fournieri
360 - A T P N
401 K D I K E D G Q G A A A E N K V F K E N N Q D V A A E E S K
358 - - - - - - - - - - - - - - E K P T Y
372 K A K A E E E D G G N G D - - - - - - H H R H E P P P Q T N
348 - E E L V S E K N D
370 D D - K S V E
A. andraeanum
P. hybrida
P. frutescens
Z. mays
A. thaliana
T. fournieri

Figure 3.9. Phylogenetic relationship of AaANS1 to confirmed ANS proteins from other species. Comparisons are made between ANS proteins from monocot and dicot species. The monocot species are highlighted in red. The phylogenetic tree is based on alignments of full protein coding sequences. The alignment was done using the MegAlign programme (gap penalty 10, gap length penalty 10) from the DNASTAR sequence analysis suite of software. The protein sequences and their accession numbers are: *Petunia hybrida* (P51092), *Perilla fructescens* (AB003779), *Arabidopis thaliana* (U70478), *Torenia fournieri* (AB044091), *Anthurium andraeanum* (this work) and *Zea mays* (P41213).


P. hybrida
P. frutescens
A. thaliana
T. fournieri
A. andraeanum
Z. mays

A BLAST search using the sequence *FNS 3* showed that neither the DNA nor the possible reading frames had any significant homology to known *FNS II* clones previously isolated or to P450 proteins, which is the enzyme family to which *FNS II* proteins belong.

The similarity of *FNS 2* and *FNS 8* was confirmed with the sequence data obtained. However, the 5' terminal reaction failed because of a 70 bp -CACA- stretch in the 5' untranslated region of both clones. In addition, no significant information could be gained from the 3' terminal reaction because of the presence of long poly $(A)^+$ tails. Thus BLAST searches against the database were not possible.

3.7 SCREENING FOR AN ANTHURIUM *CHI* cDNA CLONE

The first approach was through a library screen, because of the ready availability of cDNA clones in our laboratory. Using a lisianthus full-length *CHI* insert, an aliquot of 20,000 plaques was screened at 55 °C as previously described. Attempts were also made with antirrhinum and petunia *CHI* full-length clones. In addition, to the library screens, RT-PCR was performed using degenerate primers designed to conserved regions from *CHI* amino acid alignments. The details of the primers used are in Appendix I.

In the library screen, petunia *CHI* gave no positives in the primary screen and antirrhinum *CHI* gave a single positive, which did not carry through to the secondary screen. However, seven independent positives were obtained with the lisianthus *CHI* and all these gave positive signals in the secondary screen. The excised phagemid from positive single plaque picks from the secondary screen were recovered by in vivo excision and plasmid DNA extracted. Of the seven positives from the secondary screen only three (CHI 8, 9 and *11*) gave bands of the expected size (approximately 0.8-1.0 kb) with restriction analysis (data not shown). The clones were sequenced using M13 reverse and M13 forward primers, respectively. Sequence results showed that *CHI 8* and *11* were identical to each other, but along with *CHI 9*, did not show any homology to *CHI* cDNA sequences in the database.

A range of RT-PCR reactions were carried out. Only two primers, (*CHI 1* and *CHI 1-1*, see Appendix I) gave products of the expected estimated size. However, all primers gave a

band of expected size with the antirrhinum positive control first strand cDNA. Sequence analysis of the anthurium PCR products and subsequent BLAST searches proved that the PCR products were not *CHI*.

3.8 SCREENING FOR AN ANTHURIUM *F3'H* cDNA CLONE

The cDNA library was screened with a full-length antirrhinum *F3'H* cDNA which gave only a single positive signal, which did not follow through to secondary screens. Subsequent probings with 50,000 plaques, and in one instance with as many as one million plaques, proved equally unproductive.

Attempts to isolate a clone using degenerate primers (*F3'5'H-L* and *F3'5'H-R*, see Appendix I) that had been successfully employed in isolating an orchid *F3'H* cDNA (Professor Heidi Kuehnle, University of Hawaii, personal communication) gave the same outcome as the library screen.

3.9 SCREENING FOR AN ANTHURIUM *UF3GT* cDNA CLONE

In the flavonoid pathway, various sugar moieties are added to the flavonoid ring of flavones, flavonols and the anthocyanidins. These glycosylation reactions are catalysed by enzymes belonging to the family of glycosyltransferases (GTs). Anthurium species do not accumulate flavonols, the flavones are *C*-glycosylated and the anthocyanins are diglycosylated at the 3-position with glucose and rhamnose to form the anthocyanin rutinoside (Iwata et al., 1979). This suggests that both the glucosyl transferase (UF3GT) and the rhamnosyl transferase (3RT) are active in anthurium spathe. There are several flavonol and anthocyanin GT cDNA sequences published but to date there has been no report in the literature of the cloning of a flavone *C*-GT cDNA. Consequently, anthocyanin GT cDNA clones were the primary targets of this isolation effort. Their availability would be useful in the subsequent analysis of gene expression patterns for anthurium anthocyanin biosynthetic genes as well as in future flower colour manipulation efforts.

Both RT-PCR and library screening techniques were used in attempts to secure an anthurium *UF3GT*. The initial set of experiments employed 3' RACE technology (Frohman et al., 1988) using as templates, a range of first strand cDNA stocks made from: a) RNA from all six stages pooled together, b) RNA from Stages 3-5 and c) RNA from each individual flower stage. With oligo(dT_{17}) as the anchor primer, PCR reactions were carried out using primers designed to the highly conserved FV(L)THCGWNS motif of glycosyltransferases (Vogt and Jones, 2000). The primers had been used to isolate glycosyltransferase cDNAs from a diverse range of plant species including petunia, antirrhinum and sinningia (Dr. Chris Winefield, Crop & Food Research, New Zealand, personal communication). Several primers were used but only *UFGT 3-1* extended (primer details are provided in Appendix I) gave a band of the expected size. PCR reactions were with the high fidelity Platinum *Pfx* Polymerase (Section 2.9.2).

The expected band size using the *UFGT* primers was 500 bp and was obtained for all positive controls used in the PCR reactions. However, with the cDNA from anthurium a band size of approximately 550 bp was consistently obtained. There was no observed difference in the product obtained using the different first strand cDNA stocks. Although slightly larger than expected, the 550 bp product was cloned and sequenced. However, upon analysis, the fragment did not encode any region of homology to known glycosyltransferase sequences.

Library screens were performed, first with a partial sinningia cDNA clone and an antirrhinum *UF3GT* full-length cDNA. Both screens were done at 55 °C. Aliquots from the primary library equivalent to 20,000 plaques were screened. In addition, larger screens were performed with 100,000 to one million plaques from an amplified library stock.

The sinningia cDNA used to screen the library was thought to correspond to a UDP flavonol glycosyltransferase. At 55 °C several positives were obtained and these were taken through to secondary and tertiary screens. The phagemids from the positive single plaque picks were rescued by in vivo excision and plasmid DNA purified and sequenced as

previously described. However, again the sequence data proved negative, with clones having no significant homology to *UF3GT*'s in the database.

Duplicate lifts were performed with the antirrhinum *UF3GT Eco* RI full-length insert probe, and three independent positive signals (*UF3GT 6, 9 and 1*1) were obtained which matched perfectly on the first and second lift of the duplicate membranes. However, there were no positive signals in the secondary screen (which was with a new probe) except for very faint signals from *UF3GT 11*. This did not follow through to the tertiary screen. Nonetheless, a single plaque pick from the secondary screen of *UF3GT 11* was selected and in vivo excision performed and plasmid DNA prepared for sequencing. Prior to sequencing the clone, it was tested with PCR using *UFGT 3-1* and *UFGT 3-1* extended and the dT_{17} as the 3' anchor primer. No product of the desirable size was obtained and the sequence data were consistent with this result, confirming that the clone recovered by the screen of the library with antirrhinum *UF3GT* was not an anthurium anthocyanin glycosyltransferase.

3.10 DISCUSSION

Having sequences for *CHS, F3H, DFR* and *ANS* clones available from several other plant species on the database enabled extensive comparisons to be made with the anthurium cDNA candidates obtained by the library screen. With all the alignments shown in Figures 3.1 to 3.8 very significant levels of % identities are evident between the anthurium clones and those of the relevant predicted amino acid sequences of genes from other plant species. Such analyses, outside functional characterisation, are sufficient to conclude, with a high likelihood, that *AaCHS1, AaF3H 1, AaDFR1* and *AaANS1* correspond to mRNAs from genes for *CHS, F3H, DFR* and *ANS* respectively.

While it was expected that the deduced amino acid sequences of the four anthurium cDNA clones would show a closer phylogenetic relationship with related clones from other monocots, this was observed only for AaF3H1 and AaDFR1. The grouping of the predicted AaDFR1 with DFR proteins from other monocots (Figure 3.6) supports the idea developed in Liew et al. (1998) for the pattern of DFR segregation into monocots and dicots. However, even among the monocots, AaDFR1 shows significant divergence (Figure 3.6).

Both AaANS1 and AaCHS1 group with dicots suggesting that they have diverged from the other monocot genes or came from a separate ancestor, shared with dicots. There appears to be no consistently strong correlation in phylogeny between the anthurium biosynthetic genes investigated and those of other monocots.

3.10.1 *AaCHS1*

CHS belongs to the group of plant specific polyketide synthases, a burgeoning family of enzymes, that shares greater than 65% identity and includes: STS, 2-pyrone synthase, p-coumaroyl acetic acid synthase, phloroisovalerophenone synthase and benzophenone synthase (reviewed in Schröder, 1999, 2000). CHS uses 4-coumaroyl-CoA as a starter molecule and catalyses the stepwise decarboxylative condensation of acetate units from malonyl-CoA to form a tetraketide. Cyclisation and aromatisation of the enzyme-bound tetraketide intermediate leads to formation of naringenin chalcone (Kreuzaler and Hahlbrock, 1975).

The CHS active site consists of two functionally independent units, the coumaroyl-binding pocket and the cyclisation pocket, defined by four residues conserved in all the known CHS-related enzymes. These residues are also conserved in AaCHS1 and correspond to cysteine-166, phenlalanine-217, histidine-305 and asparagine-338 (Figure 3.1). In fact all the essential structural features of the deduced polypeptide sequences that have been shown to be essential for the function of this family of enzymes are conserved in AaCHS1, confirming that this clone is a polyketide synthase.

However, because these structural features are not unique to CHS proteins, they cannot be used to determine if the putative protein product encoded by *AaCHS1* is in fact the anthurium *CHS*. Therefore, there are limitations to the use of homology searches for ascribing identity to putative CHS proteins (Schröder, 1997). Inferences on the functional identity of AaCHS1 can best be made with comparisons of % identity with functionally defined CHS proteins along with other members of the polyketide family, as well as from phylogenetic analysis. The amino acid sequence of AaCHS1 has highest % identity to functionally defined CHS proteins with an average of 83% identity. This is markedly

different to the low values for the CHS-related proteins such as 58% for benzophenone synthase from *Hypericum androsaemum* (Accession number AF352395). The levels of % identity of AaCHS1, with the known CHS proteins in Figure 3.1, is well within the range obtained when these proteins are compared among themselves. Similar % identities are also observed when AaCHS1 and functionally defined CHS proteins are compared to polyketide synthases that are not CHS.

The level of divergence seen in the primary protein structure was useful in distinguishing the functional identity of the gerbera CHS family. Gerbera CHS2 (GCHS2) showed only 73% identity to the two CHS proteins CHS1 and CHS3. In addition, *GCHS2* was expressed in almost all tissues and its expression profile did not correlate with flavonoid production (Helariutta et al., 1995). The identity of this protein was later confirmed as 2-pyrone synthase (Eckermann et al., 1998). By comparison, the *AaCHS1* expression pattern showed a strong correlation with flavonoid and anthocyanin production in anthurium spathe (discussed in detail in Chapters 4 and 5). This provides additional favourable evidence for the functional identity of *AaCHS1*. However, it would be useful to determine if *AaCHS1* transcripts are detected in tissues that do not accumulate anthocyanins. There is a limitation to this approach as CHS activity is necessary for the production of other flavonoids besides anthocyanins. In this regard, information on the number of CHS copies in the anthurium genome would be useful, as it cannot be stated for certain that *AaCHS1* represents the only *CHS* mRNA in anthurium tissue.

Of the polyketide family, CHS is by far the most widely distributed enzyme, being present in all angiosperms. The other related proteins are more restricted. STS, though found in a limited number of plants, is one of the more closely related enzymes to CHS with greater than 65% identity. While it is not known whether anthurium tissues accumulate stilbenes (e.g resveratrol), they were recently described in monocots, where they were not previously suspected (Powell et al., 1994). Therefore, it is useful to be able to distinguish AaCHS1 from this category of polyketide synthases. This is almost impossible by comparisons of the amino acid sequences and while phylogenetic analyses may show strong and distinct groupings (Figure 3.2), this cannot be taken as proof of function. However, careful

examination of specific residues adjacent to the active site may provide some point of difference between CHS and STS proteins (Ferrer et al., 1999).

It has been shown that the specific structure of the active sites of these two enzymes governs the stereochemistry of the cyclisation reaction. Although the active sites for the condensing reactions are in the same position (the conserved cysteine in Figure 3.1) and there are no consistent differences in the residues lining the active sites of the CHS and STS enzymes, sequence variability occurs in the solvent exposed residues of the strands lining the cyclisation pocket (residues 253-259) and (262-268) (Ferrer et al., 1999). In this region AaCHS1 has the residues that are conserved in CHS enzymes and does not exhibit the variability seen in STS proteins.

Clearly the evidence from the analysis of the primary structure of AaCHS1 seems to suggest that it encodes an anthurium CHS protein, however, definitive proof of function requires genetic evidence or heterologous expression in a suitable host to assay for CHS activity.

3.10.2 *AaF3H1* and *AaANS1*

Both *AaF3H1* and *AaANS1* hybridised strongly to their respective heterologous probes and, as will be discussed in Chapters 4 and 5, their gene expression patterns correlate strongly with anthocyanin expression. In addition, BLAST searches with the DNA sequence or the amino acid sequence of AaF3H1 or AaANS1 gave best hits for the respective proteins.

F3H and ANS belong to the dioxygenase family of proteins. There are four structural characteristics that define the identity of this family of proteins and that are of mechanistic significance (Britsch et al., 1993; Lukacin and Britsch, 1997; Lukacin et al., 2000). Firstly, conserved histidine and aspartic acid residues have been identified as part of the putative iron-binding site (Britsch et al., 1993; Lukacin and Britsch, 1997), and arginine and serine residues (the RXS motif, where X could be any of the amino acid residues) as part of the 2-oxoglutarate binding site (Lukacin and Britsch, 1997; Lukacin et al., 2000). These residues are conserved in AaF3H1 and AaANS1 (Figures 3.3 and 3.8, respectively). Similarly,

proline residues that have been suggested to play a role in the folding of the polypeptide (Britsch et al., 1993) are also conserved in identity and spacing in AaF3H1 and AaANS1.

Finally, the leucine rich motif identified in Figure 3.3 is suggested to form an α-helix, which may be critical for function (Britsch et al., 1993). This, too, is conserved in both AaF3H1 and AaANS1. Together, this information strongly indicates that *AaF3H1* and *AaANS1* do encode the anthurium *F3H* and *ANS*, respectively.

3.10.3 *AaDFR1*

The *AaDFR1* cDNA clone gave strong hybridisation signals with its heterologous probe. In addition the temporal transcript abundance profile of the gene suggest that its expression is concomitant with anthocyanin production in the spathe (details are provided in Chapter 4). Further to this, BLAST searches with the DNA sequence or the deduced amino acid sequence gave best hits for the DFR proteins and strengthen the case for *AaDFR1* being derived from the endogenous *DFR* transcript in anthurium spathe tissue.

There are several structural features of AaDFR1 that are highly conserved in functionally characterised DFR proteins. These conservations are at specific points of homology within the family of proteins to which DFR belongs and at the proposed catalytic site for DFR.

While the particle bombardment experiments with petunia BR140 would have confirmed functionality of AaDFR1, the results are also critical to future flower colour modification efforts in anthurium aimed at developing blue flowers. Difficulties in using these mutants for transformation or transient expression experiments have been reported (Filippa Brugliera, personal communication). However, time did not allow for continued work in this regard, but efforts to optimise the transient expression protocol in this petunia line with the aim of determining the ability of AaDFR1 to catalyse DHM may prove useful.

Barring functional characterisation, information from the analysis of the primary structure of AaDFR1, strongly suggests that the cDNA clone for *AaDFR1* obtained from the spathe library corresponds to mRNA from the anthurium *DFR* gene.

DFR is known to display substrate specificity in several species such as petunia (Gerats et al., 1982), *C. hybrida* (Johnson et al., 1999), mathiola (Hellar et al., 1985) and carnation (Stich et al., 1992). Anthurium accumulates both pelargonidin and cyanidin glycosides (Iwata et al., 1979, 1985) suggesting that this DFR can work with either DHK or DHQ, like the maize and gerbera DFR (Meyer et al., 1987; Elomaa et al., 1995). The motif, along with the four amino acid residues within the motif that determine substrate specificity have been identified (Beld et al., 1989; Johnson et al., 2001). Two of these residues are conserved in AaDFR1 while the remaining two residues are conservative substitutions (Figure 3.7).

These two conservative substitutions in the proposed substrate specificity region suggest that it may be the class of amino acid and not the amino acid itself that is important in determining the substrate specificity. As seen with gerbera, substituting a hydrophilic residue (N) for the hydrophobic leucine changes the substrate specificity of the DFR (Johnson et al., 2001). However, an anthurium DFR carrying a hydrophilic substitution of serine for asparagine can catalyse both DHK and DHQ. It may be that the serine substitution in AaDFR1 introduces a greater affinity for DHQ explaining the higher levels of cyanidin-based anthocyanins compared to pelargonidin-based ones in the red anthurium spathe. However, pelargonidin dominates in orange anthurium spathe, and it may very well be that there is no substrate specificity in anthurium DFR and that greater levels of cyanidin based anthocyanin in the red line is simply due to higher activity of F3'H.

That the substitution of serine for asparagine-134 (N134) in anthurium DFR may not be significant finds additional support from the proposed biochemical model explaining the preference shown by DFR to different dihydroflavonol substrates. It has been suggested that differential hydrogen bonding between the dihydroflavonol substrate and DFR may be the factor that distinguishes the three substrates (Johnson et al., 2001), as the dihydroflavonols differ only in the degree of B-ring hydroxylation. Therefore, the adjusted

substrate specificity observed in gerbera was probably due to the substitution of a hydrophobic residue (leucine) for the wild type hydrophilic amino acid. Such a change theoretically changes the putative hydrogen bonding potential of the mutated DFR. Serine contains an aliphatic chain with a hydroxyl group that makes the amino acid highly reactive as it readily forms hydrogen bonds. Therefore, the serine substitution in AaDFR1 should provide the equivalent hydrogen bonding as asparagine.

Interestingly, Johnson et al. (2001) did not include the barley DFR in their alignments. Like the petunia DFR, the barley clone has a glutamine substitution (Figure 3.7) for the corresponding, and highly conserved glutamic acid, yet the clone is functional (Kristiansen and Rohde, 1991). Again it appears that it is the class of residue rather than a specific residue that is important, as both glutamic acid and glutamine are hydrophilic amino acids. Therefore, DFR functionality probably requires a hydrophilic residue in this position.

The three DFR proteins in the alignment (Figure 3.7) that are known to catalyse DHM are carnation, petunia and barley (Gerats et al., 1982; Kristiansen and Rohde, 1991; Stich et al., 1992). However, from Figure 3.7, there are no residues within the substrate specifying motif that are specific only to these three DHM-accepting DFR proteins. However, it is not known if the absence of delphinidin-derived anthocyanins in the other species in the alignment (Figure 3.7) is due to DFR specificity or the absence of F3'5'H. Further work is needed, therefore, to identify the residues that will confer substrate specificity to DHM. However, the implication that substrate specificity can be altered by a single amino acid change is useful in regards to devising strategies to engineer blue coloured anthurium. However, in practical terms it would be more convenient to introduce a DFR clone that catalyses DHM as part of the strategy.

Taking the investigation one step further, to infer into the nature of the catalytic site, a serine residue and a tyrosine-X_3 lysine couplet in the short chain dehydrogenase/reductase (SDR) family has been shown to be critical for catalytic activity (Johnson et al., 2001). These residues are absolutely conserved in all DFR proteins including AaDFR1.

Taken together, both the molecular and the phylogenic analysis suggest that AaDFR1 was derived from a DFR transcript in the anthurium spathe and does encode an anthocyanin specific DFR protein.

3.10.4 Attempts to isolate anthurium *FNS*

The availability of a cDNA clone controlling flavone production in anthurium spathe might be useful in transgenic approaches to manipulate flower colour as well as influencing other economic factors such resistance to diseases. Three routes are proposed for the formation of flavones. FNSI (a 2-oxoglutarate type enzyme), FNS II (a P450 enzyme) and F2H (a P450 enzyme) for the flavone *C*-glycoside. Of these, FNS II appears to be the more common route, for in comparison to F2H, which has only been reported in one plant species (Akashi et al., 1998), FNS II has been cloned from gerbera (Martens and Forkmann, 1999), torenia (Akashi et al., 1999) and antirrhinum (Kevin Davies, personal communication). On this basis, antirrhinum *FNS II* as opposed to *FNS I* was used a probe to screen the library.

The antirrhinum *FNS II* cDNA probe, based on the primary, secondary and tertiary screens seemed like the ideal probe, as independent picks in the primary screen were consistently detected in the secondary and tertiary screens. However, the inability to clone an *FNS II* or at least a P450 gene was quite surprising. Given the inability to isolate an *FNS* cDNA clone for anthurium and the presence of flavone *C*-glycosides in this species, it may be that flavones are produced through the alternate pathway catalysed by F2H (Figure 1.4). However, should this be the case one would have expected the *FNS II* probe to cross hybridise to F2H given the reported 54% homology of the gerbera *FNS II* to the *F2H* of *G. echinata* (Martens and Forkmann, 1999).

The results from the screening are unexpected for other reasons as well. Firstly, the quantity of flavones and the times in which they are produced (as will be discussed in Chapter 4), suggest that a cDNA for *FNS* should be well represented in the cDNA pool of the library.

In light of the interference caused by -CACA- stretch in the 5' untranslated region of FNS 2 and FNS 8 it may be worthwhile performing selected restriction digest reactions to remove this region. The clone can then be re-ligated and internal sequence information obtained. Alternatively, using degenerate primers designed to the P450 binding domain may prove to be a successful approach, as it was for gerbera (Martens and Forkmann, 1999).

3.10.5 Attempts to isolate an anthurium *CHI*

It has been suggested that the isomerisation of chalcones into the corresponding flavanones could be non-enzymatic (Stafford, 1991), and spontaneous isomerisation has been documented in vitro (Moustafa and Wong, 1967; Mol et al., 1985). Against this backdrop, and based on biochemical flavonoid data for anthurium, a functional CHI in spathe tissue was inferred, and, hence targeted. The inability to secure a cDNA clone using heterologous low stringency screening with lisianthus, antirrhinum and petunia cDNAs was surprising. However, the petunia *CHI* cDNA was not able to recognise maize CHI in low stringency Southern or northern blots (Grotewold and Peterson, 1994). On this basis maize *CHI* was isolated using degenerate primers designed to conserved regions from petunia, barley and antirrhinum *CHI* (Grotewold and Peterson, 1994).

Once cloned, the maize CHI was shown to share 58% identity with petunia CHI-A, 55% with CHI-B, 50% *Phaseolus vulgaris* (bean) CHI and 59% with antirrhinum CHI. If 55-59% identity could not detect maize *CHI* by northern or Southern, then the inability to isolate an anthurium *CHI* by heterologous screening may be due to low levels of homology between the probe and the target sequence.

It may be suggested that in selecting the size range for my library, small fragments, such as those encoding for a CHI protein, were excluded. However, in screening for other genes, cDNA fragments ranging from 0.5-0.6 kb were obtained suggesting that the estimated size range of *CHI* cDNA was included and should have been represented in the library population.

Furthermore, it seems unlikely that anthurium *CHI* gene expression levels are so low that they cannot be detected by library screen or even PCR techniques, as there has been no report of these techniques failing outright in other plant species. However, in reviewing the literature, it is clear that PCR primers of lower degeneracy could have been designed, so increasing the possibility of recognising and amplifying the desired target fragment.

3.10.6 Attempts to clone an anthurium *F3'H*

Flavonoid 3' hydroxylase belongs to the P450 monooxygenase class of enzymes and it is known that genes for these enzymes can be expressed at very low levels. However, there is no clear explanation for the lack of success in isolating a cDNA clone as heterologous screening has been successful for *F3'H* in other studies (Brugliera et al., 1999).

The results, however, of the screening efforts with the anthurium library fit well with the biochemical data discussed in Chapter 5 showing that no microsomal F3'H protein activity was detected in anthurium spathe tissue. It is possible that anthurium has developed an alternative route for producing 3' hydroxylated flavonoids but this would require further investigation to confirm.

3.10.7 Attempts to clone an anthurium *UF3GT*

Heterologous screening has been used successfully to isolate GT cDNA clones in other studies (Martin et al., 1991; Yamazaki et al., 2002). This suggests, that anthocyanin GTs are expressed at high enough levels to be detected in library screens and that there is sufficient homology between anthocyanin GTs from different plant species for recognition and isolation of the clone by heterologous screening of cDNA libraries. Hence, screening the anthurium cDNA library with the antirrhinum *UF3GT* or the sinningia GT fragment was a reasonable approach. However, the similarity amongst GTs is restricted to the proposed plant secondary product GT consensus sequence (called the PSPG box) the proposed binding site of the UDP-sugar moiety (Bairoch, 1991; Hughes and Hughes, 1994). Outside this conserved domain homology levels are generally very low.

Bearing this in mind, it may be that the level of divergence in the anthurium clone does not support heterologous screening at 55 °C. Better results may be obtained with a lower stringency screen. In addition, the high G+C content of maize was believed to contribute to the difficulties in isolating antirrhinum *UF3GT* using the *Bronze 1* cDNA (Dr. Cathie Martin, John Innes Centre, United Kingdom, personal communication). Clones such as *ANS and CHS* isolated from anthurium have G+C content in excess of 60% compared to 40-45% for the same genes in other plant species. A similar situation may exist with anthurium *UF3GT* and screening the library with the maize GT may be more profitable. The codon bias in anthurium may also explain the inability of the *UFGT* primers to recognise and amplify the anthurium *UF3GT*.

In summary, heterologous screening of an anthurium spathe cDNA library yielded four full-length cDNA clones *AaCHS1*, *AaF3H1*, *AaDFR1* and *AaANS1*. Detailed sequence analysis with functionally defined orthologues from other species, coupled with BLAST homology searches and phylogenetic comparisons, suggests that the four clones encode anthurium *CHS, F3H, DFR* and *ANS*. The temporal and spatial expression patterns of these genes is presented and discussed in the following chapter.

Chapter 4

Temporal and spatial expression patterns of flavonoid biosynthetic genes during the development of anthurium flowers

4.1 INTRODUCTION

Characterising the expression pattern of genes involved in flavonoid biosynthesis in anthurium is fundamental to developing a complete understanding of the molecular biology of flower colour development in this ornamental species. One of the aims of the work presented here was to determine whether the suite of flavonoid biosynthetic genes in anthurium was coordinately regulated as subsets, as in antirrhinum and petunia, or as one entire group, as with maize. This would allow inferences to be made on the number of different regulatory mechanisms operating in the control of gene expression in this metabolic pathway. The link between the expression pattern and the regulatory mechanism is based on the well documented fact that the expression of genes involved in anthocyanin biosynthesis is controlled primarily through the regulation of transcription (Cone et al., 1986; Almeida et al., 1989; Martin et al., 1991). The expression studies may also indicate whether *AaCHS1* and *AaF3H1* are for genes involved in anthocyanin biosynthesis.

4.2 MATERIALS AND METHODS

4.2.1 Plant material used for northern analysis

The range of plant tissue used in this study is described in Table 2.1 of Chapter 2. Tissue was collected from the spathes of Altar, Atlanta, Lido and Acropolis and from spadix tissue of Montana. The specific lines used for a particular experiment are discussed in the following sections.

4.2.1.1 Temporal expression of flavonoid biosynthetic genes

As described in Chapter 2, flower development was divided into six developmental stages that were consistent across all the lines being investigated. Spathe tissue was collected for RNA isolation from all or some of these stages and used in northern analysis to determine the temporal pattern of expression for anthurium flavonoid biosynthetic genes. The extraction protocol is described in Section 2.2.1.

4.2.1.2 Tissue specific expression patterns of flavonoid biosynthetic genes

The expression pattern of anthurium *CHS, F3H, DFR* and *ANS* was investigated in a young leaf from Altar and from spadix tissue of Altar and Montana (Figure 4.1 and 4.2). Montana is a lavender coloured miniature *A. amnicola* hybrid. This species has been used in several interspecific crosses because of the ease with which it can be crossed to other species. Hybrids with *A. andraeanum*, *A. antioquiens*, *A. formosum* and other species have been developed (Kamemoto and Kuehnle, 1996). Although it is not known for certain the origin of Montana, it resembles closely the hybrid of *A. amnicola* crossed with *A. lindenianum* in Kamemoto and Kuehnle (1996). Montana has a pink spathe and spadix (Figure 4.1). The expression pattern of the four genes in this spadix was compared with that in the bright yellow coloured spadix of Altar, that accumulates a small amount of anthocyanin in a ring of cells within each floret. The stages used for the spadix correspond to the stages defined for the development of the flower. As with flavonoid extractions on the spadix, total RNA was isolated from the layer of coloured cells by the method described in Section 2.2.2.

4.2.1.3 Investigating diurnal rhythms in anthurium flavonoid biosynthetic gene expression

To investigate the possibility of fluctuations in transcript levels in spathe tissue during the day, the original Stage 4 was further subdivided into five distinct growing stages, namely Stage 4-1, 4-2, 4-3, 4-4 and 4-5, as shown in Figure 4.3. Stage 4, was chosen to be subdivided because it was the only developmental stage that allowed such separation into smaller distinct phases. The description of the stages is as follows:

- Stage 4-1: newly extended peduncle with an average peduncle length of 3 to 4 cm.
- Stages 4-2, 4-3 and 4-4 are 1 to 2 weeks apart in growth. Peduncle lengths are 10 to 15 cm, 20 to 25 cm and 30 to 35 cm, respectively. For these stages the spathe is fully coloured and tightly folded around the developing spadix within.
- Stage 4-5: The peduncle is fully extended with a length of 40 to 43 cm. The spathe is still tightly wrapped around the spadix. In addition, small protuberances of the true flowers are just evident on the base of the spadix. This is notably absent from any of the previous stages.

Figure 4.1. The coloured lines used to examine transcript abundance in the spadix. Annotations are: Sp- spathe and Spx- spadix.

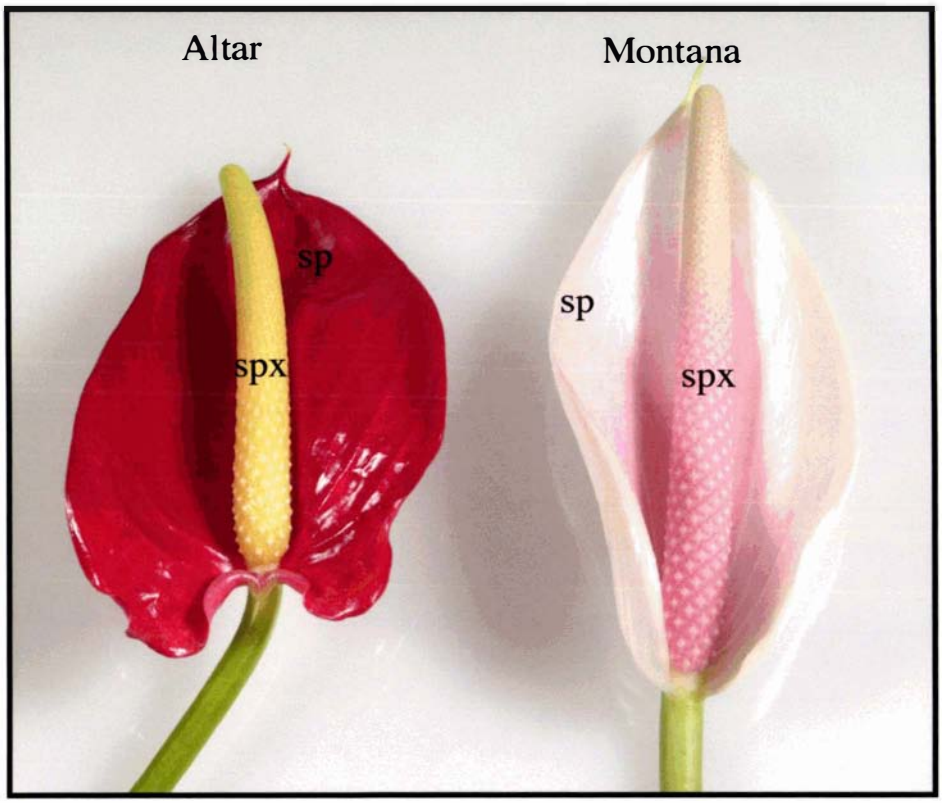

Altar
Montana
sp
spx
sp
spx

Figure 4.2. High magnification photo of the spadix in Altar. Only very small amounts of anthocyanin are produced and their deposition appears confined to a ring of cells within each floret (see arrows). The boundary of a single floret is marked out in black. Red spathe tissue is the background.

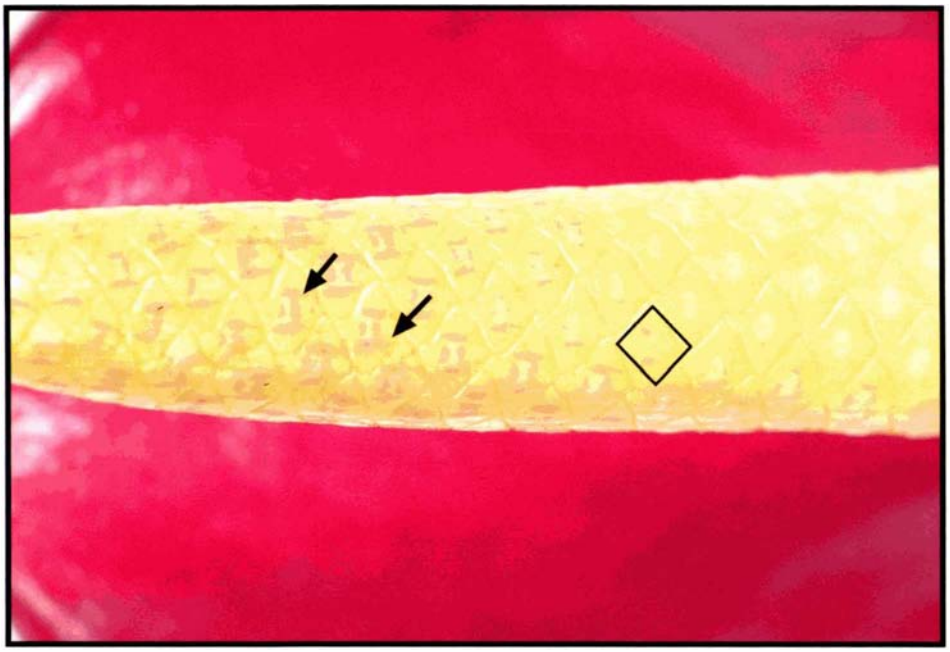
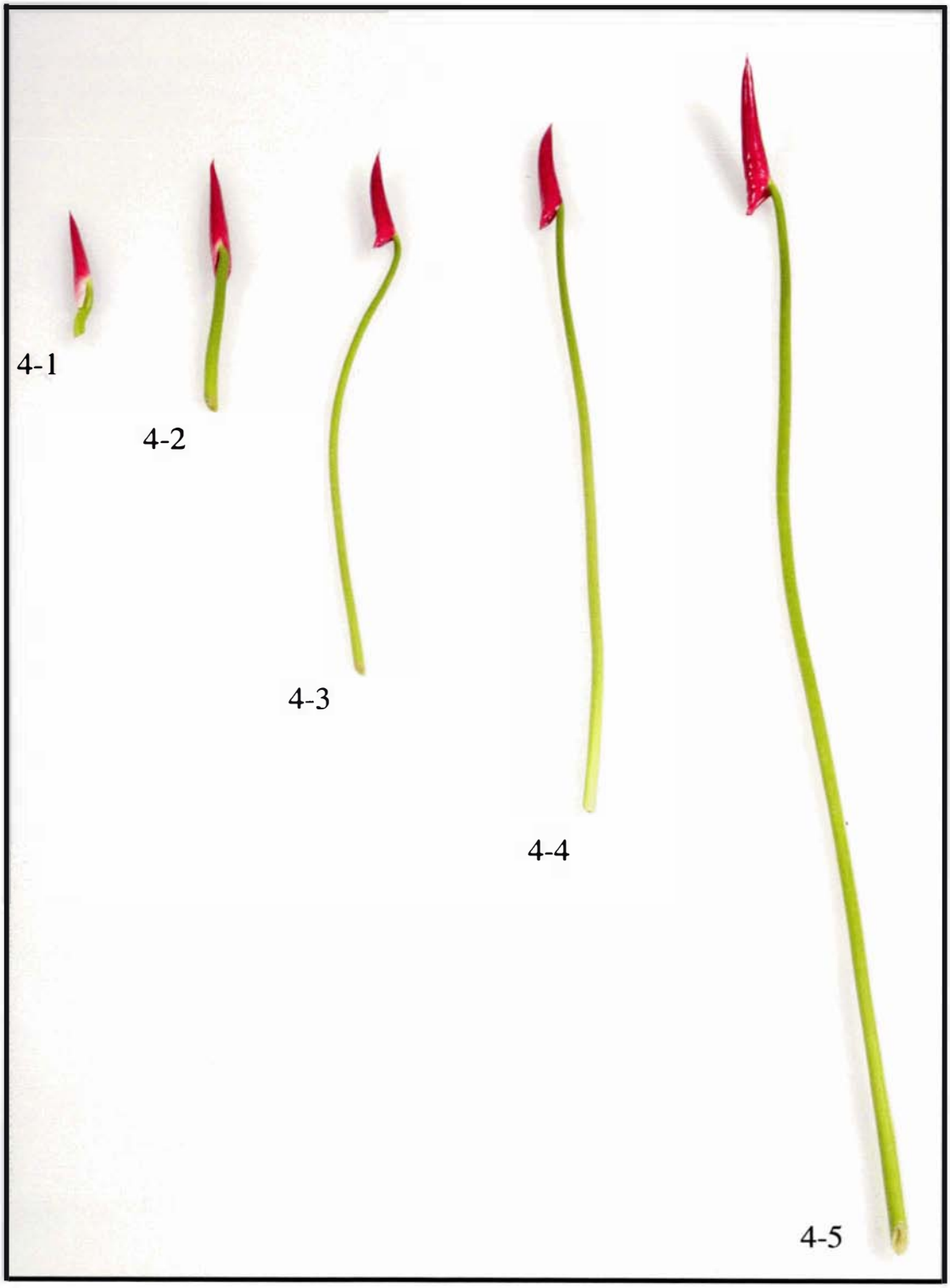


Figure 4.3. Stage 4 subdivisions for diurnal gene expression analysis. These five subdivisions of Stage 4 in Altar enabled the investigation of transcript abundance of *CHS, F3H, DFR* and *ANS* over shorter times and within shorter developmental intervals.


4-1
4-2
4-3
4-4
4-5

Spathe samples for each stage were collected at three distinct time points during the day, at dawn, noon and dusk over the period extending from mid August (winter) to mid October (spring). Such an extended collection period was necessary because of the slow growth of the tissue. Collected samples were immediately stored in liquid N_2 or crushed in liquid N_2 for immediate RNA isolation by the method described in Section 2.2.1.

4.2.2 Biochemical analyses of anthurium spathe and spadix

4.2.2.1 Extraction and quantification of flavonoid and anthocyanin levels

Anthocyanin and flavonoid production profiles were done for both spathe and spadix from three cyanic anthurium lines, Montana, Altar and Atlanta. The method used for extracting from both tissue types is described below. Whenever flavonoids compounds were being extracted from spadix tissue, a razor was used to separate the layer of coloured cells.

Three separate samples (200 mg FW) for each line were weighed into individual microcentrifuge tubes. An overnight extraction was then set up with 1 mL of 70% (v/v) methanol (with 10% (v/v) acetic acid in water). The following day, tubes were centrifuged for 3 min at maximum speed in a bench top microcentrifuge and the supernatant transferred to a new microcentrifuge tube. The pellets were then resuspended in 90% (v/v) methanol (with 10% (v/v) acetic acid in water) for another overnight extraction. Once completed, the tubes were centrifuged as before and the supernatants from both rounds of extraction combined to give the crude extract. The extracts were dried in vacuo on a Savant SC210 Speedvac to near dryness and made up to a final volume of 1 mL in 80% (v/v) methanol (with 10% (v/v) acetic acid in water). These were then concentrated to 200 µL by drying in the Speedvac at room temperature. An aliquot of 800 µL of methanol was added to each tube and the contents mixed and centrifuged as above to pellet out any solid matter. Total flavonoid and anthocyanin levels were estimated by spectrophotometry using a Jasco V-530 UV/VIS spectrophotometer (Jasco, Japan) and extinction coefficients ($E^{1\%}_{1cm}$) of 14,300 and 35,000 for flavonoids and anthocyanins respectively. Total flavonoid levels were estimated by reading the absorption of 10-20 µL of extract at a wavelength of 350 nm in 1 mL of methanol in a quartz cuvette. To measure total anthocyanin levels, the

absorbance of a similar quantity of extract (10-20 μL) was measured in 0.1 M HCl (in methanol) at 530 nm (the wavelength of maximum absorption). Values of 0.01-0.02 were regarded as representing the absence of anthocyanin. Results are presented as the mean of the three replicates with their respective error measurements.

4.3 RESULTS

4.3.1 Anthocyanin and flavonoid production profiles for different coloured lines

The more common anthurium spathe colours include red/pink, orange and white. Of growing popularity are those with green, brown and mottled spathe colours (refer to Figure 1.3). Visually, anthocyanin levels increase in the spathe as it develops. However, there are observable differences between species. A biochemical profile, such as described in this section provides more detailed information of anthocyanin and flavonoid production in the spathe.

Both Altar (red) and Atlanta (orange) share very similar anthocyanin production profiles. Anthocyanin levels increase throughout the development of the spathe with a peak at Stage 6. However, the most rapid increase in production occurs between Stages 3 and 4 (Figure 4.4). While anthocyanin levels do increase throughout development in the pink line, Lido, the most rapid increase in production occurs between Stages 4 and 6 (Figure 4.4). This peak in anthocyanin, occurring in the latter stages of spathe development in Lido, is expressed as a blush of colour that develops only in the fully expanded spathe. Total flavonoid levels (Figure 4.5), investigated only in Altar, follow a different pattern to the anthocyanin profiles, with peak levels occurring very early in spathe development (Stage 2).

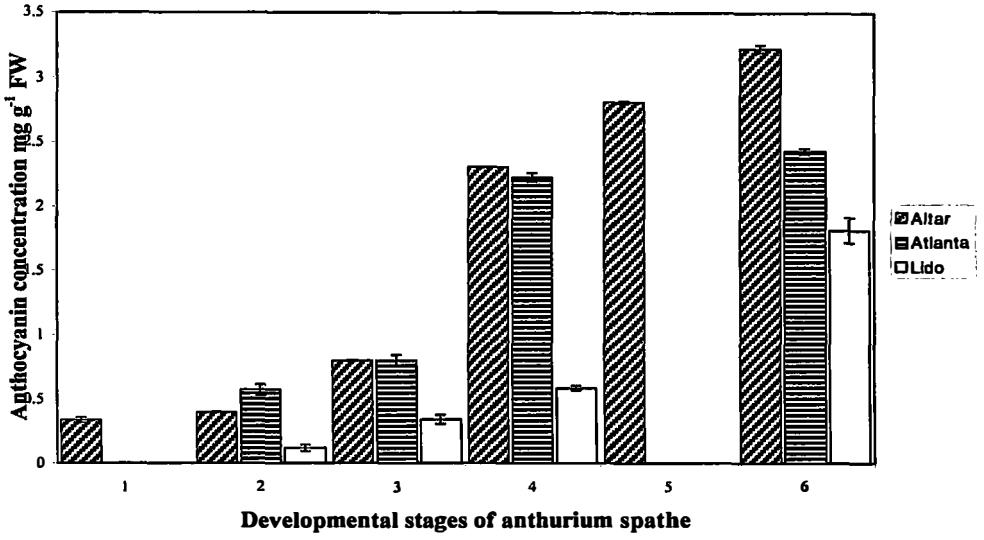


Figure 4.4. Anthocyanin production profile for Altar, Atlanta and Montana monitored over the development stages of the spathe. Only Stages 2, 3, 4 and 6 were used for Atlanta and Montana. The data points are mean values of the three replicates used for each line. The error bars are standard deviations of the mean values.

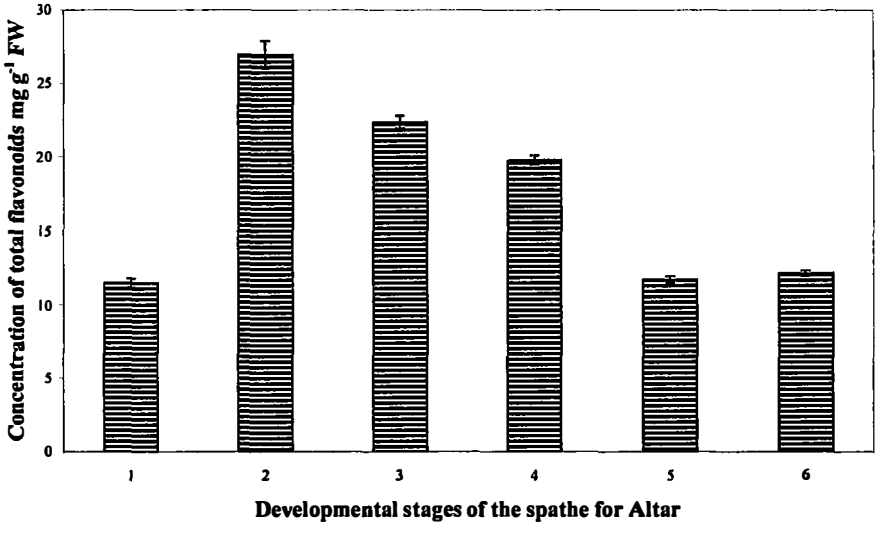


Figure 4.5. The Developmental profile of total flavonoids in the spathe of Altar. The data points are mean values of the three replicates used for each line. The error bars are standard deviations of the mean values.

4.3.2 Temporal gene expression patterns for *CHS, F3H, DFR* and *ANS* in the spathe of Altar

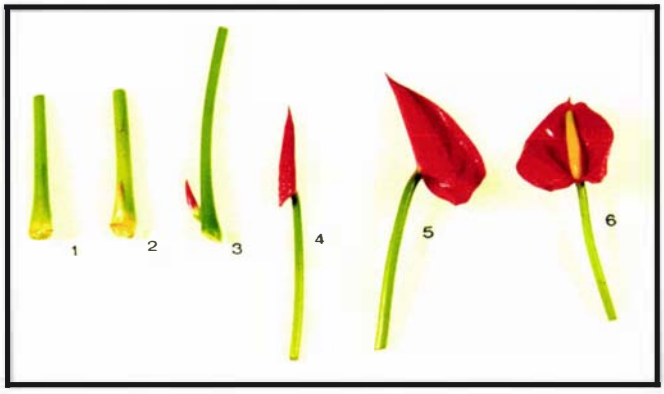
Transcripts of the four genes examined could be detected in all the stages of spathe development investigated (Figure 4.6). Transcripts for *CHS, F3H,* and *ANS* appear to be constitutively expressed over all six stages, although slight increases can be seen between Stages 2 and 3 for *CHS* and *ANS*. However, *DFR* shows a distinct gene expression pattern, with a marked increase in transcript level at Stage 3 of spathe development followed by a significant decline between Stages 4 and 6. The peak in *DFR* levels correlates well with the anthocyanin production pattern, which increases most rapidly between Stages 3 and 4.

The presence of transcripts for *CHS, F3H, DFR* and *ANS* in the first stage of flower development corresponds to the early accumulation of anthocyanin in this anthurium line. Anthocyanin levels are low between Stages 1 and 3 of the spathe and increase sharply at Stage 4, with gradual increases through to Stage 6 (Figure 4.4). The anthurium flower at Stage 1 is still enclosed in the leaf sheaths at the swollen base of the petiole (Figure 4.6) and so is not yet visible. However, if the sheaths are removed and the immature flower exposed, the spathe is usually already pigmented in a pattern that originates from the junction of the spathe and spadix moving outward towards the tip of the spathe. This pattern holds until Stage 4 when anthocyanin levels and colour have increased resulting in pigmentation of the entire spathe. The detection of transcripts at the very first stage of flower development fits with this early pigmentation pattern. Until Stage 4 of flower development, total RNA would have been extracted from tissue that was part pigmented and part acyanic.

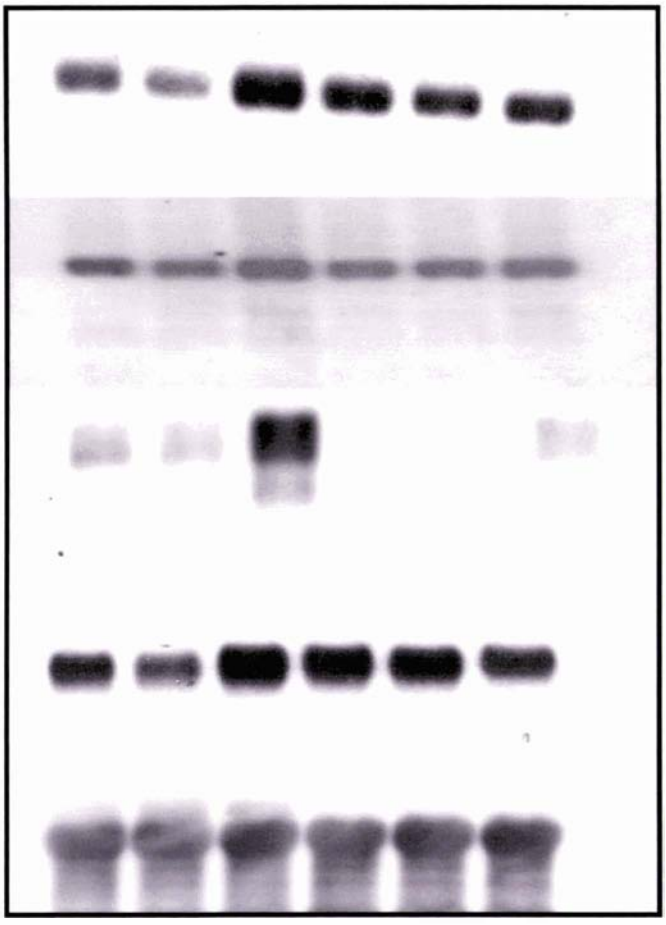
4.3.3 Temporal expression patterns of *CHS, F3H, DFR* and *ANS* in four anthurium lines

Transcript levels for *CHS, F3H* and *ANS* were comparable between Atlanta and Altar (Figure 4.7) with transcript levels for *DFR* again showing the most variability. Interestingly, there were variations seen in these Altar samples compared to those used for Figure 4.6. For example in Figure tissue 4.7, the peak in *DFR* transcript levels occurs

Figure 4.6. Flavonoid biosynthetic gene expression in Altar spathe. Transcript abundance and temporal distribution for four flavonoid biosynthetic genes (*CHS, F3H, DFR* and *ANS*) over the six stages of flower development for the red anthurium line was determined by loading an aliquot of 20 μg of total RNA for each stage, and a 25/26S ribosomal RNA probe (pTIP6) from *Asparagus officinalis* (asparagus) (King and Davies, 1992) was used as the loading control. Radiolabeled full-length cDNA clones for each gene were used as probes. The developmental stages of the flower from which spathe samples were taken are also shown.



Stages 1 2 3 4 5 6



CHS

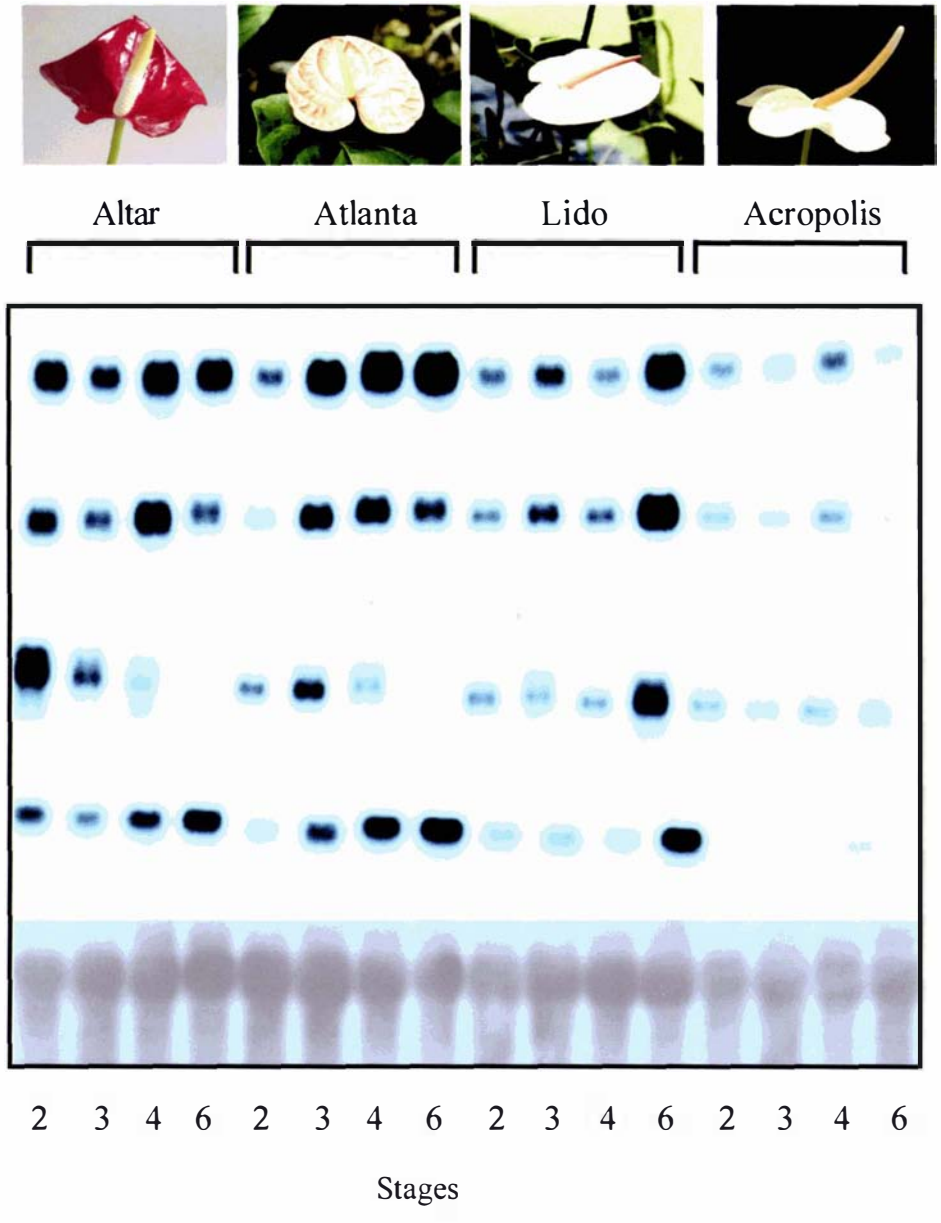
F3H

DFR

ANS

pTIP6

Figure 4.7. Comparison of flavonoid biosynthetic gene expression in four anthurium lines. 20 μg RNA was used for each of the lines and was isolated from four stages of spathe development (Stages 2, 3, 4 and 6). Radiolabeled full-length cDNA clones for each gene were used as probes. A 25/26S ribosomal RNA probe (pTIP6) from asparagus (King and Davies, 1992) was used as the loading control.


Altar
Atlanta
Lido
Acropolis
CHS
F3H
DFR
ANS
pTIP6
2 3 4 6 2 3 4 6 2 3 4 6 2 3 4 6
Stages

earlier in the spathe development, at Stage 2, compared to Stage 3 for the previous analysis of Altar (Figure 4.6). The *DFR* patterns in Atlanta matched the previous results for Altar (this can be observed by comparing both Figures 4.6 and 4.7).

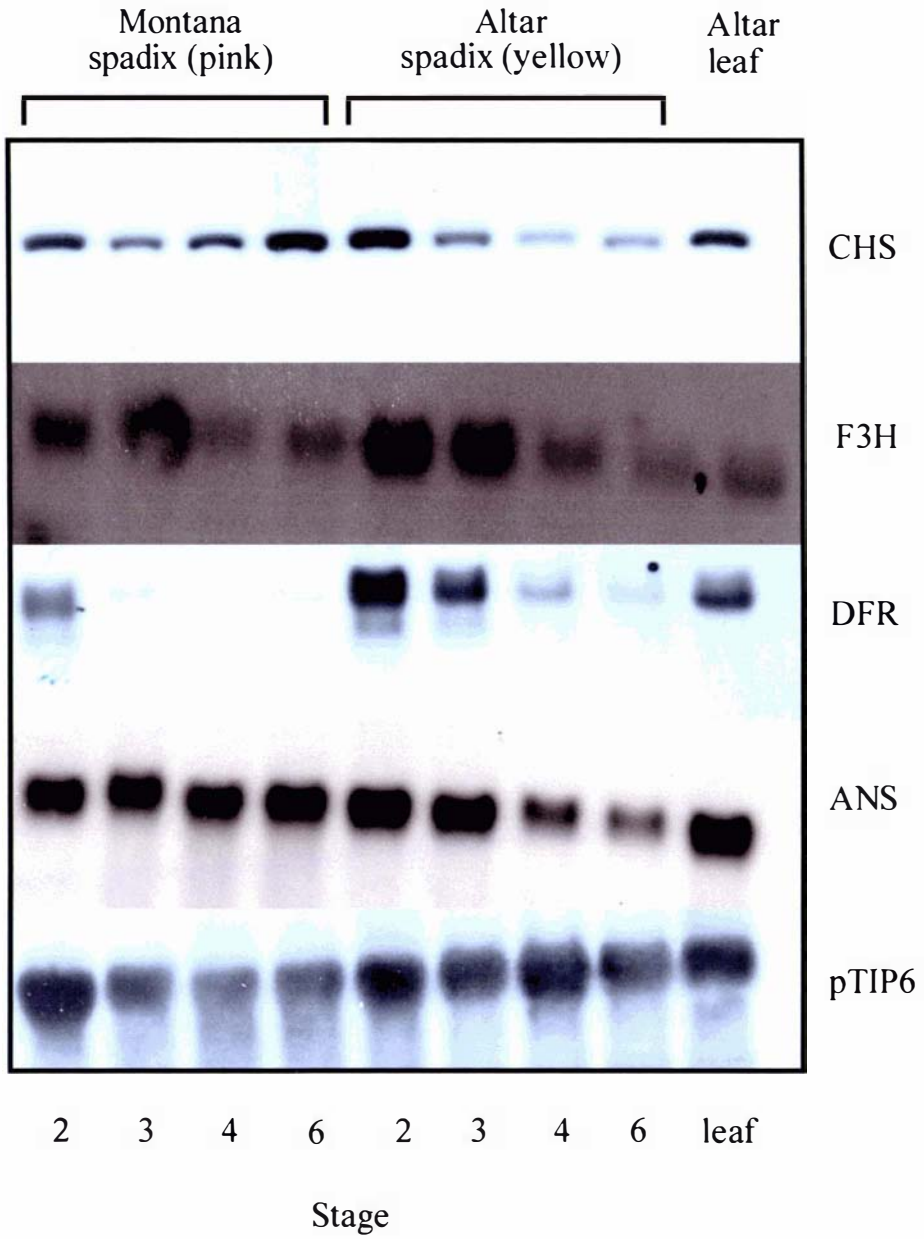
In Acropolis, all four flavonoid biosynthetic genes had equally reduced transcript levels when compared to the three cyanic lines (Figure 4.7) while the expression pattern of the four biosynthetic genes in Lido followed closely the anthocyanin accumulation profile for that particular line (Figure 4.4). In this pink line, anthocyanin peaks late into flower development, with full colour developing only after the flower is fully opened. Even at this stage it is the adaxial surface of the spathe that is mostly coloured. This late stage production of anthocyanin fitted well with the late peak of transcript accumulation for all four genes, at Stage 6 (Figure 4.7).

4.3.4 Tissue specific expression of anthurium flavonoid biosynthetic genes

In addition to the spathe, transcripts of all four genes were detected in leaf and spadix, of the lines Altar and Montana (Figure 4.8). Anthurium leaves, especially Altar and Atlanta, accumulate anthocyanins, but this is usually only noticeable during the very early stages of leaf growth prior to the accumulation of chlorophyll.

Altar has a red spathe and a yellow spadix whereas Montana has a pink spathe and pink/purple spadix. Although the primary colour pigment in the yellow spadix of Altar is carotenoid (David Lewis, personal communication), microscopic examination showed that anthocyanins are found in small amounts in the ring of cells surrounding each floret (Figure 4.2). In contrast, cyanidin glycosides predominate in Montana spadix tissue (David Lewis, personal communication), giving the pink/purple colour to the spadix. Irrespective of this difference in the quantity of anthocyanin present in the two tissues, similar levels of transcripts for all four biosynthetic genes were detected in the yellow spadix of Altar compared to the pink spadix of Montana. It is not known if non-anthocyanin flavonoids are produced in the spadix of either line.

Figure 4.8. Flavonoid biosynthetic gene expression in spadix and leaf. Transcript abundance for *CHS, F3H, DFR* and *ANS* are compared in the leaf of Altar and the spadix tissue of Altar and Montana. 20 μg of total RNA was loaded for each lane and a 25/26S ribosomal RNA probe (pTIP6) from asparagus (King and Davies, 1992) was used as the loading control. Radiolabeled full-length cDNA clones for each gene were used as probes.


Montana
spadix (pink)
Altar
spadix (yellow)
Altar
leaf
CHS
F3H
DFR
ANS
pTIP6
2
3
4
6
2
3
4
6
leaf
Stage

In the spadix tissue of Altar (Figure 4.8), *DFR* does not exhibit an expression pattern that is distinct from the other biosynthetic genes examined. In Altar spadix, transcript for all the genes, decline as the spadix matures.

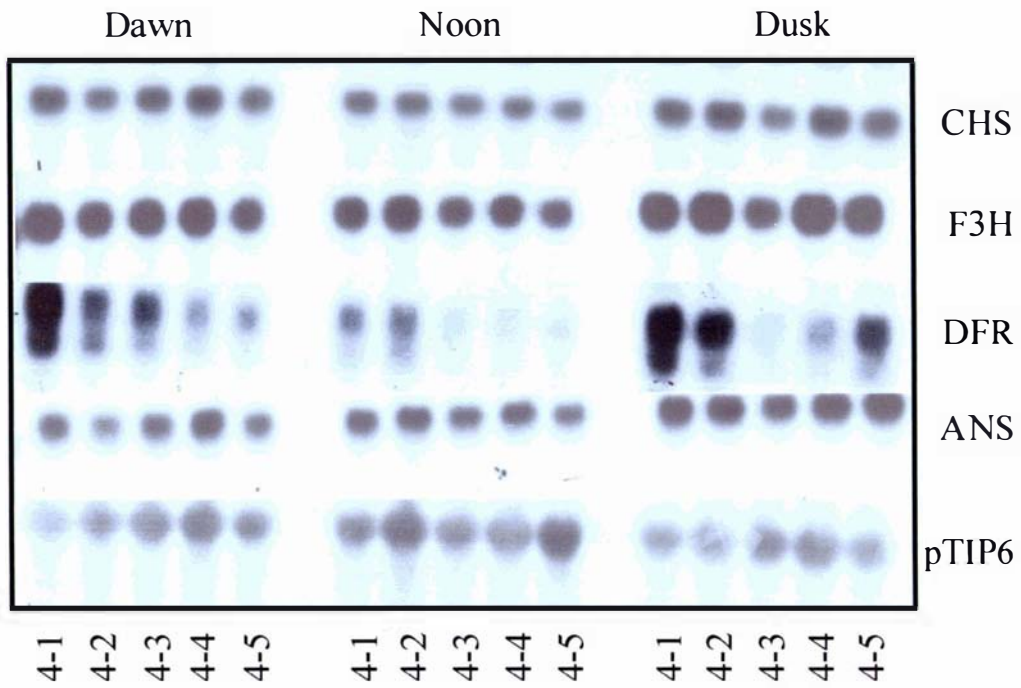
4.3.5 A diurnal rhythm to the temporal expression of DFR transcript levels

There is a lengthy time interval of 2-3 weeks between each of the spathe developmental stages identified. With this slow pattern of development for the anthurium flower, it is possible that the stages used encompassed too many developmental changes, resulting in interesting gene expression patterns going undetected. In addition, when the transcript abundance for *DFR* in Altar shown in Figure 4.6 is compared to that in Figure 4.7 some clear inconsistencies are observed that cannot be accounted for by simple variations in total RNA loadings. As a result, a sampling variation was inferred. Moreover, experiments were set up as described in Section 4.3.3 to investigate the possibility of a diurnal rhythm to the steady state levels of *DFR* transcripts. Subdividing Stage 4, as shown in Figure 4.3, facilitated the enquiry into rhythmic patterns of *DFR* expression while at the same time allowing a more detailed investigation of the expression pattern of anthurium flavonoid biosynthetic genes over a shorter time period.

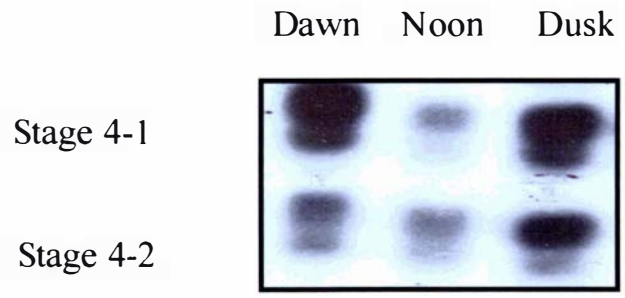
Transcript levels for *CHS, F3H* and *ANS* showed no variation in abundance for all the stages investigated and during the three time points studied. However, *DFR* showed a different expression pattern with a distinct variation in transcript levels occurring during the day (Figure 4.9). *DFR* transcript levels were high at dawn but decreased during the day to a low level at noon and then were high again at dusk. This pattern was consistent for the five subdivisions of Stage 4.

Figure 4.9. Diurnal rhythm of *DFR* gene expression. Panel A shows transcript levels for *CHS, F3H, DFR ANS* at dawn, noon and dusk across five developmental subdivisions of Stage 4 in Altar. Panel B is extracted from panel A and highlights the diurnal changes in *DFR* transcript levels over the time frame investigated. Note the marked change in transcript levels for Stages 4-1 and 4-2 during the three time periods of the day. 20 μg of total RNA was loaded for each lane and a 25/26S ribosomal RNA probe (pT1P6) from asparagus (King and Davies, 1992) was used as the loading control. Radiolabeled full-length cDNA clones for each gene were used as probes.

A


Dawn
Noon
Dusk
CHS
F3H
DFR
ANS
pTIP6
4-1
4-2
4-3
4-4
4-5
4-1
4-2
4-3
4-4
4-5
4-1
4-2
4-3
4-4
4-5

B


Dawn
Noon
Dusk
Stage 4-1
Stage 4-2

4.4 DISCUSSION

4.4.1 Temporal gene expression pattern

The temporal expression patterns of flavonoid biosynthetic genes have been documented for several plant species such as antirrhinum (Martin et al., 1991; Jackson et al., 1992), petunia (Quattrocchio et al., 1993), maize (Dooner and Nelson, 1977; Ludwig and Wessler, 1990), lisianthus (Davies et al., 1993; Oren-Shamir et al., 1999), arabidopsis (Shirley et al., 1995; Pelletier and Shirley, 1996; Pelletier et al., 1999), grape (Boss et al., 1996) and perilla (Gong et al., 1997). With the exception of maize these are all dicot species. In many dicots the expression of genes involved in anthocyanin biosynthesis appears to be separated into two subsets that represent the division of the pathway into two independently regulated units referred to as the early and late biosynthetic genes (EBGs and LBGs respectively). For example, in antirrhinum *CHS* and *CHI* comprise EBGs while *F3H, DFR* and *ANS* comprise the LBGs. In these plants, anthocyanin expression is concomitant with LBG expression only.

However, in grape berries there is no EBG/LBG division, but rather that anthocyanin production is concomitant only with *UFGT* expression (Boss et al., 1996). Furthermore, while it is thought that the absence of *UFGT* expression is responsible for white grapes, it is not known if this is due to mutations(s) in the structural or regulatory gene(s) (Boss et al., 1996). Similarly, in perilla, no distinct subdivision into LBGs and EBGs is observed, as the expression of all structural genes examined was coordinately regulated (Gong et al., 1997). Anthurium seems to represent another divergence from the previously studied species.

As mentioned above, the only monocot in which similar research has been conducted is maize. In this case all the structural genes involved in anthocyanin biosynthesis are coordinately induced. Therefore, the transcripts of all the structural genes involved in anthocyanin production in the maize kernel have the same temporal distribution pattern (reviewed in Mol et al., 1998).

Although anthurium is a monocot, it differs from the maize in having at least two subgroups with different regulatory patterns. It is clear from northern analysis on Altar that the temporal expression patterns of *CHS, F3H* and *ANS* are temporally coordinated while that of *DFR* appears to be governed by a different temporal control mechanism. The simplest model to explain coordinate regulation of genes is that the coordinately regulated genes all respond to the same transcription factor(s), by virtue of common regulatory sequences in their promoters. Therefore, the same regulatory proteins could be involved in the coordinate expression of *CHS, F3H* and *ANS*. However, given the sharp peak of *DFR* transcript levels from Stage 2 to 3 followed by the rapid decline by Stage 4, the existence of another set of regulatory proteins acting in anthocyanin biosynthesis in anthurium can be inferred.

The modular control of the pathway is thought to be an evolutionary feature that enables flowers to develop separate regulatory elements controlling the production of UV-protectant flavones and flavonols independent of anthocyanin production (Martin et al., 1991). The role of a similar mechanism in anthurium spathe tissue is crucial given the extensive damage usually seen if the plant is exposed to direct sunlight. However, while modular control is suggested based on the results, it is not immediately clear how coordinate regulation of *CHS, F3H* and *ANS* coupled with the suggested selective regulation of *DFR* relates to this objective.

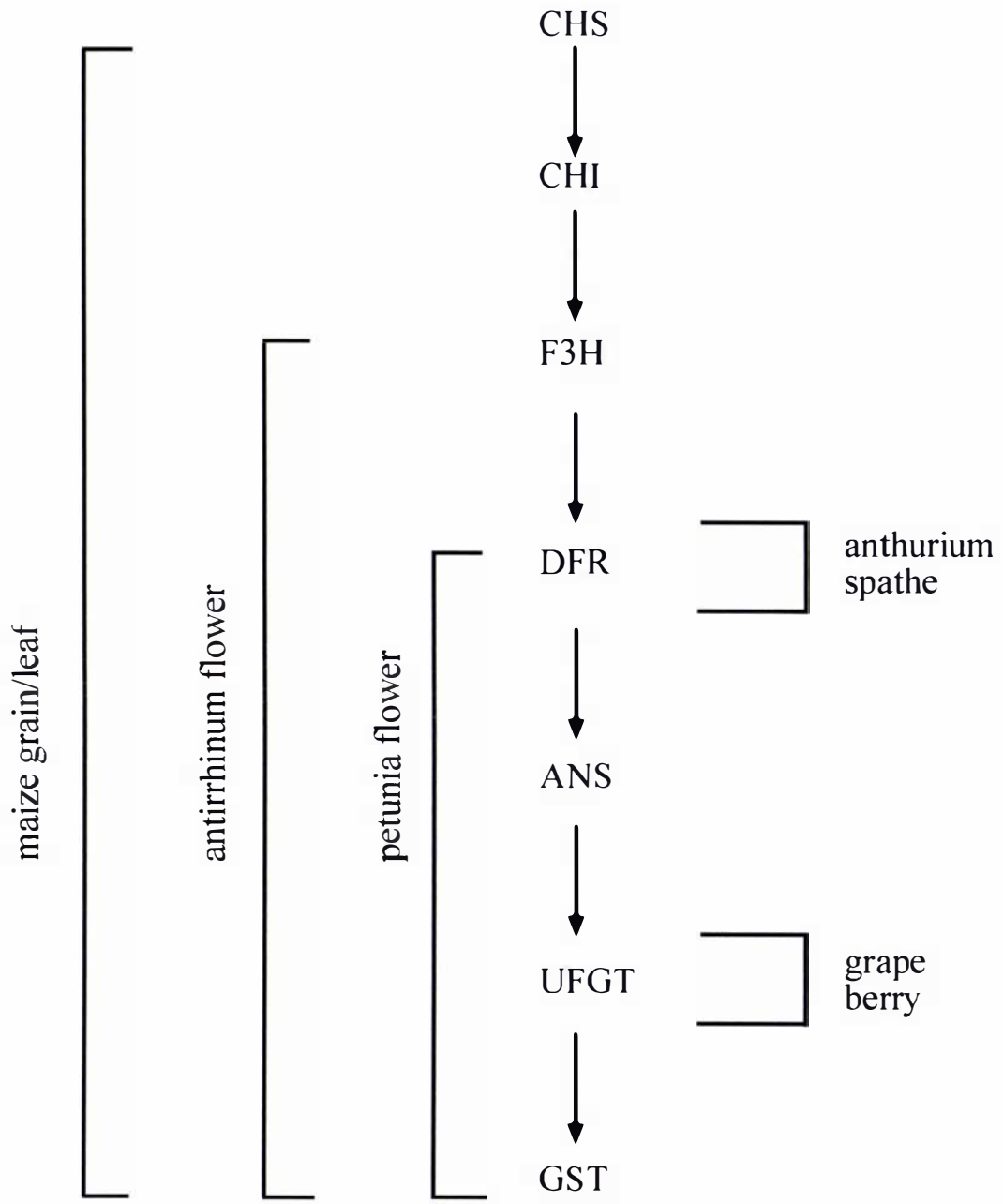
The selective regulation of *DFR* transcript levels is supported by the expression pattern observed in Atlanta (the orange line), where *DFR* transcript accumulation is again at variance with that shown by *CHS, F3H* and *ANS* (Figure 4.7). This trend did not occur in Lido. This may be because colour develops in this line only when the spathe is fully expanded coinciding with peak transcript levels for all four genes at Stage 6 (Figure 4.7). To make comments about the specific regulation of *DFR* in this line, studies of spathe development may need to be investigated beyond Stage 6. However, it can be said that as with other cyanic lines, anthocyanin production is correlated with increased *DFR* transcript levels in the spathe.

The results from these experiments suggest that in the anthurium spathe, *DFR* is a prime target for regulatory control for anthocyanin biosynthesis and, like grape, anthurium may have a single point of regulatory control. The proposed model for the regulation of anthocyanin biosynthesis in anthurium spathe is presented in Figure 4.10. However, a more complete picture will emerge once clones for *CHI, FNS, F3'H* and *UFGT* are obtained.

The transcript abundance profile for anthurium *CHS* and *F3H* strongly supports their involvement in anthocyanin biosynthesis (Figure 4.7). However the presence of their (*CHS* and *F3H)* transcripts, particularly in early flower development, may not only be related to anthocyanins. As mentioned previously anthurium accumulates large quantities of flavones, particularly flavone-*C*-glycosides (Williams et al., 1981). Therefore, the biochemical profile for total flavonoids in anthurium (Figure 4.5) would largely reflect the production patterns for flavone-*C*-glycosides. Peak production levels occur in the first two stages of spathe development. This would require the activation of *CHS* and *CHI* (but not *F3H*) gene activity in the early stages of flower development. This could imply the existence of a separate regulatory protein that shares both *CHS* and *CHI* targets with specific anthurium anthocyanin regulators. This would resemble the situation in maize, where phlobaphene biosynthesis branches off from anthocyanin biosynthesis. In maize floral organs, the Myb gene *P* controls phlobaphene pigmentation by directly activating *CHS, CHI* and *DFR* (Grotewold et al., 1994).

An alternative to this scenario is the existence of more than one copy of *CHS* and *CHI* in the genome of anthurium, similar to that seen in petunia. In petunia, *CHS* exists as a large multi-gene family with eight-ten members (Koes et al., 1987), only two of which are related to anthocyanin biosynthesis (Koes et al., 1989). If a similar situation applies in anthurium spathe tissue, then there exists the possibility of a *CHS* or *CHI* gene specific for flavone synthesis. While attempts were made in this study to determine if gene families existed for the flavonoid structural genes the data were inconclusive (data not shown). However, should multiple copies of *CHS* or *CHI* exist, the northern data presented may not have been able to distinguish them, as full-length clones were used as cDNA probes.

Figure 4.10. The proposed model for the regulation of anthocyanin biosynthesis in anthurium spathe. *DFR* serves as a key regulatory target and *CHS, F3H* and *ANS* are all coordinately regulated. For maize, antirrhinum and petunia, genes in bracket are coordinately regulated. For grape regulatory control appears to be vested in *UFGT* only.


CHS
CHI
F3H
DFR
ANS
UFGT
GST
maize grain/leaf
antirrhinum flower
petunia flower
anthurium spathe
grape berry

Attempts to isolate a *CHI* cDNA to allow studies of the *CHI* gene expression pattern proved unsuccessful. However, given the fact that flavones accumulate, it can be inferred that the CHI protein is active in anthurium tissues (unless the reaction is spontaneous in anthurium), and one would predict its gene expression pattern might resemble that of *CHS*.

4.4.2 Spatial expression pattern

The expression patterns of *CHS, F3H, DFR* and *ANS* in the spadix reveal some interesting deviations from those seen in the spathe. From a biochemical perspective, this comes as no surprise, as spadices have a different biochemical profile to spathes. For example, the spadices of Altar and Atlanta are bright yellow in colour and accumulate carotenoids, which are not present in the spathes (data not shown).

While it is not known if flavonoids accumulate in these spadices, they have been detected in those of Lido and Montana (data not shown). However, in the Altar spadix (yellow) the regulation of flavonoid structural genes appears very different to that in the spathe. *DFR* exhibits a less distinct expression pattern in comparison to that of *CHS, F3H* and *ANS*, as all four genes have high expression levels at Stages 2 and 3 (except for *CHS* where transcript levels fall off in Stage 2) followed by marked declines as the spadix matures (Figure 4.8). This decline in transcript levels for the biosynthetic genes as the tissue develops, is also at variance with what was observed in the spathe. In the latter, expression levels of *CHS, F3H* and *DFR* remained relatively consistent as the flower matured. Only with *DFR* were changes in transcript levels observed as the spathe matured (Figures 4.8 and 4.9).

In Montana spadix, (the *A. amnicola* hybrid with pink coloured spadix), anthocyanins are the primary pigments and the expression pattern of the four genes (especially *DFR*) in this tissue closely resembles that seen in the spathe of Altar. However, *DFR* seems to have a unique expression compared to that for the other three genes, in spadix tissue of Montana. Thus, in a different tissue, and even in a different species (*A. amnicola*), *DFR* again emerges as a prime regulatory target possibly with its own regulatory machinery.

Another point of interest is that the yellow spadix produces only a small amount of anthocyanin (Figure 4.2), yet the transcript levels in this tissue are comparable to, and in some cases, exceed that observed in the pink coloured spadix, which produces much more anthocyanin pigment. This is not consistent with *DFR* serving as the point of flux for anthocyanin production. Such high levels of *DFR* and *ANS* expression usually accompany a quantitatively significant amount of anthocyanin. The fact that little anthocyanin is being produced with abundant transcript for *CHS, F3H, DFR* and *ANS*, suggests that in the spadix of Altar, the anthocyanin pathway is blocked at a structural gene level acting downstream from *ANS*. This implicates *UFGT* activity in a pattern very similar to that of Shiraz grape berries (Boss et al., 1996). In young berry skins, *CHS, CHI, F3H, DFR* and *ANS* are all expressed, yet no anthocyanins are detected because the structural gene *UFGT* is not induced.

The simplest explanation for limited anthocyanin pigment in Altar spadix is a structural gene mutation. The reduction of *CHS, F3H and ANS* transcript levels in Altar spadix as it matures is in striking contrast to the expression pattern exhibited by the same genes in the spathe. Therefore the regulatory model suggested for the spathe of Altar may not apply for its spadix. Indeed, the expression pattern of anthurium *CHS, F3H, DFR* and *ANS* in Altar spadix, does support the previous suggestion that other regulatory targets for anthocyanin biosynthesis exist in anthurium and that there may be tissue specific elements to their expression. Such situations have been found for other species. In maize, for example, whereas R along with C1 regulates the entire suite of flavonoid genes in the kernel, the gene expression patterns in the anthers appear to be more complex. In anther tissue, the *R* allele appears to be required for *CHS, F3H* and *DFR* gene expression pattern in a stage or cell specific manner but not for *CHI* (Deboo et al., 1995). The possibility of a similar phenomenon occurring in anthurium spadix tissue is supported by the genetic studies of Kamemoto et al. (1988). They showed that the genetics of spadix colour inheritance is more complex than that suggested for the spathe, with several additional but undefined genetic factors determining the final colour achieved. Interestingly, they suggested that two different genetic factors (*C* and *R*) in the spadix, as opposed to *M* and *O* in the spathe, (discussed in detail in Chapter 5) control spadix colour. In combination, *C* and *R* produce

red to orange spadices and *C* produces orange to coral spadices. When both are lacking, spadix tissue is yellow and this was attributed to flavones (Kamemoto et al., 1988). Related work contradicts this report, as yellow coloured spadices of Altar and Atlanta are due to carotenoids (David Lewis, personal communication). Also analysis of transcript profiles confirms that anthocyanin and anthocyanin-specific genes are expressed in yellow spadix tissue of Atlanta. Therefore, the regulatory protein that activates these genes in the spadix is present and functional and suggests that this is the unlikely identity for either *C* or *R*. In addition, given that a structural gene mutation is suggested for the low levels of pigment in Altar spadix tissue it is possible that *C* and *R* encode structural genes such as *UFGT* or *GST*. It does seem as though the regulatory pattern suggested for the control of anthocyanin biosynthesis in the red spathe of Altar may be different for its yellow coloured spadix and that there may be a basis for the suggestion of two separate genetic factors (*M* and *O* versus *C* and *R*) controlling flavonoid production in each tissue type (Kamemoto et al., 1988).

4.4.3 Diurnal regulation of DFR expression and its possible implications

Diurnal changes in plant responses (whether anatomical, physiological or molecular) describe plant responses that occur daily or over a day/night cycle. When the observed behaviour is inherent and independent of external periodicity then it is described as a circadian rhythm. The experiment performed on anthurium spathe in this regard examined diurnal changes in transcript levels of anthocyanin biosynthetic genes.

Diurnal and circadian regulation of gene expression is very common in plants and is necessary for the efficient temporal regulation of many physiological processes, such as stomatal aperture, leaf movement, ion uptake, nitrogen assimilation, carbon fixation and photosynthesis (reviewed in Kreps and Kay, 1997; Millar and Kay, 1997; McClung, 2000, 2001). However, despite these examples, little is known about the rhythmic control of the expression of genes involved in anthocyanin biosynthesis.

Northern analysis of tissue sampled at three time points during the day detected a definite pattern of fluctuation in *DFR* transcript levels. *DFR* mRNA steady state levels oscillate,

with high levels at dawn, reduced levels at noon, and high levels again at dusk (Figure 4.9). The fact that no such pattern is shown for the expression of *CHS, F3H or ANS*, reinforces the suggestion that *DFR* is a prime regulatory target for anthocyanin biosynthesis in the spathe of Altar. The changes in *DFR* transcript levels are most noticeable for Stages 4-1 and 4-2 (Figure 4.9B). The distinctiveness of the rhythm suggests that this phenomenon could have developmental significance.

For experiments represented in Figure 4.6 and 4.7, samples were collected between mid morning and early afternoon but never at dawn or dusk. It is possible that had another sampling time been chosen for the collection of tissues to perform the experiments that gave the results in Figure 4.6, the sharp decline at Stage 4 may not have been as distinct.

The peak at dusk, relative to the levels at noon, is consistent with the reported "anticipation" of the dark to light transition (Kreps and Kay, 1997), hinting that the diurnal rhythm of *DFR* transcript levels may be regulated by the circadian clock.

A global examination of clock-controlled genes, was carried out in arabidopsis by Harmer et al. (2000). They used a highly reproducible oligonucleotide-based array to determine the steady state mRNA levels in arabidopsis at four-hour intervals during the 'subjective' day and night. They found that 6% of the more than 8000 genes examined exhibited circadian changes in steady state mRNA levels. Davis and Millar (2001) obtained similar results. Interestingly, this circadian pattern was found for genes involved in the phenylpropanoid biosynthetic pathway. Several of the structural genes of this pathway, including DFR, were shown to peak before 'subjective' dawn. This may correspond to the high DFR transcript levels observed at dawn (Figure 4.9). In addition, some transcripts encoding transcription factors were found to oscillate in phase with the transcripts of their own target genes (Harmer et al., 2000). Moreover, they were able to accurately predict, purely by a bioinformatic approach, a gene regulatory sequence. By performing a computational analysis of cycling genes, a highly conserved promoter motif (AAAATATCT), referred to as the 'evening element' was identified in 31 genes that have a clock-controlled gene expression pattern (Harmer et al., 2000). This element, when fused to the luciferase reporter

gene, conferred rhythmic luciferase expression (Harmer et al., 2000). Similar experiments by Michael and McClung (2002) confirmed the function of this motif in the clock-regulated expression of the *CATALASE 3* gene in arabidopsis. Furthermore, luciferase reporter gene fusion studies showed that the arabidopsis *CHS* promoter is controlled at the transcription level by a distinct circadian clock (Thain et al., 2002). However, the presence of the evening element was not reported or investigated in that study. The cloning of a portion of the anthurium *DFR* promoter is discussed in Chapter 7. Analysis of this showed that similar (but not identical) motifs at different positions in the fragment. However, specific promoter studies would have to be performed to determine the functional significance of these motifs in anthurium *DFR*. In addition a genomic fragment with the entire promoter region may contain this motif.

It is not known if the rhythm observed here for *DFR* expression is under the control of a circadian clock in anthurium tissue. To investigate for circadian rhythms the experiments would have to be repeated with plants grown in continuous light or darkness for 24 hours. However, the results do suggest an interaction between *DFR* gene expression and light input pathways.

Although *DFR* is an anthocyanin specific gene, it seems unlikely that its apparent diurnal fluctuations could be related to anthocyanin production. Therefore, assuming that the result is not an artifact, the question arises as to the adaptive significance, if there is one, in anthurium of diurnally controlled *DFR* expression. What benefit could the plant derive from changing DFR transcript levels, almost on an hourly basis? This becomes contextually significant when we consider that anthurium does not produce or accumulate flavonols (Williams et al., 1981).

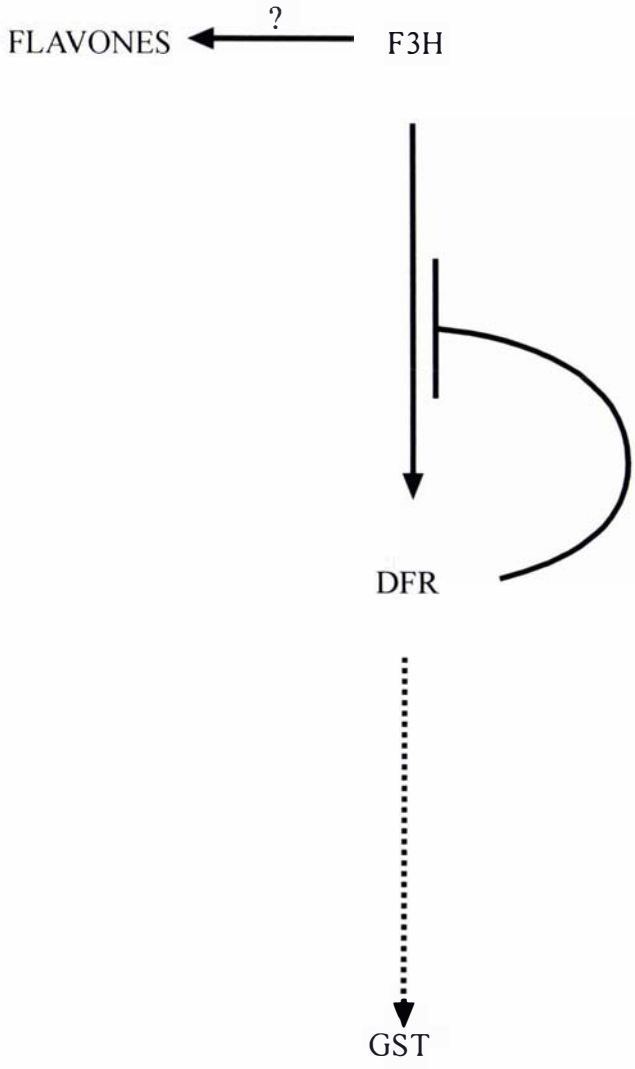
One hypothesis involves the production and accumulation of UV-B absorbing flavone compounds. Anthurium is an obligate shade plant showing high levels of plant tissue damage when exposed to direct sunlight. Flavonoids, including those of the flavone types have been implicated in protecting plants from damaging UV light (Landry et al., 1995; Markham et al., 1998). The high levels of flavone-*C*-glycosides that anthurium accumulates

suggest that these may be the principal UV-B screening chemicals in anthurium spathe tissue. However, if that were so, one would expect that *F3H* would be the primary target for regulatory control and not *DFR*, as the former is a direct competitor for naringenin, the precursor of all flavones. The model in Figure 4.11 provides one explanation for a link between diurnal regulation of *DFR* transcript levels and UV-B protection.

According to the model, low *DFR* transcript levels at noon result in the accumulation of F3H product (dihydroflavonols) causing a feedback inhibition of F3H protein activity thus allowing flavones to be produced via the anthurium *FNS* or *F2H* equivalent. Securing an anthurium *FNS* or its equivalent and performing experiments to determine if its transcript levels are adjusted at noon would test this model. In addition, the periodic arrest of *DFR* transcript levels and its proposed effect on flavone production at a certain point during the day, suggests a turnover of flavones in anthurium tissue. This phenomenon will have to be investigated further to determine the validity of this model.

Another useful additional study would be to use the 3' end of the cDNA clones as probes for northern analyses reported to ensure specificity of hybridisation. While it is not known if multigene families for the cDNA clones isolated exist in anthurium, there is the possibility that the signals from the northerns are the result of probes binding to closely related members of a gene family.

Figure 4.11. Proposed functional significance of diurnal expression of *DFR*. The model suggests a link between the rhythmic regulation of *DFR* transcript levels in anthurium spathe and its role in the accumulation of UV-B protective flavones.


?
FLAVONES
F3H
DFR
GST

Chapter 5

Molecular testing of the genetic model for flower colour inheritance in anthurium

5.1 INTRODUCTION

As discussed in Chapter 1, the genetic model put forward for the inheritance of flower colour in anthurium is one based on recessive epistasis. Two genes are suggested to control the pathway in spathe tissue, genes *M* and *O*. The *O* locus controls the production of pelargonidin-based anthocyanins, while *M* locus controls the production of cyanidin-based anthocyanins. When the *O* locus is homozygous and recessive, it has an epistatic effect over the *M* locus, and is responsible for the white phenotype (Kamemoto et al., 1988). The validity of this model and the identity or mode of action of these genes has not been determined.

Genetic changes that result in a mutant phenotype, such as a white anthurium, could occur in a regulatory or structural gene. If the mutation event prevents transcript from being produced then the molecular effect can be scored by northern analysis. This is probably the simplest tool that could give a reliable indicator as to the nature of the mutation in white anthurium. Given that anthocyanin biosynthetic genes are typically transcriptionally controlled (Martin et al., 1991; Sainz et al., 1997a), then a mutation in a regulatory gene would, depending on the nature of the mutation, affect the transcription of the target genes. This would be reflected as the down-regulation of transcript levels for several genes at once.

According to Kamemoto's model, once the *O* locus carries a dominant allele, then it is the nature of the *M* locus that determines the difference both genotypically and phenotypically between orange and red spathe colours. If *M* is dominant and present in the heterozygous or homozygous state then cyanidin-glycosides accumulate. However, if *M* is homozygous recessive then pelargonidin-glycosides accumulate, resulting in orange spathe colour (Kamemoto et al., 1988). It is not known what is the direct effect of *M* on flavone synthesis, as all white lines, whether homozygous or heterozygous at the *M* locus, continue to accumulate flavones and the impact on flavone hydroxylation has not been reported.

One molecular model that fits these data is that the *M* locus encodes F3'H, which catalyses the conversion of naringenin or dihydrokaempferol to eriodictyol or dihydroquercetin

respectively (refer to Figure 1.4). Alternatively, *M* could encode a gene specifically regulating *F3'H* activity. In either case, orange/coral anthurium lines should lack F3'H activity, thereby allowing dihydrokaempferol to accumulate resulting in the production of pelargonidin based anthocyanins.

The biochemical fact that a mixture of cyanidin and pelargonidin glycosides exist in some red and orange lines, that were not genetically classified (Iwata et al., 1985), does not compromise this model. It simply implies that for such red lines, there is only a partial reduction in F3'H activity. Thus, the orange and red lines that accumulate both precursors for pelargonidin and cyanidin based anthocyanins are possibly heterozygous at the *M* locus, or the mutant allele in these lines allows a reduced but not complete lack of F3'H activity in spathe. In this regard, cloning an *F3'H* genomic sequence from an *O-Mm* and an *O-mm* orange line and comparing these genomic sequences may provide links between gene changes and function.

An attempt was made to assess the validity of the genetic model using cultivars of defined genotype. The ability to work with genotypically defined anthurium lines allowed tentative conclusions to be made from this molecular assessment of the genetic model. Four lines with white spathes were used. Three (1244, 1250 and 1349) were genotypically defined and obtained with kind permission from Professor Heidi Kuehnle. No genetic data were available for the fourth line, Acropolis, which was obtained from Rainbow Park Nurseries in Auckland (refer to Table 1 in Chapter 2 for a description of the plant material).

The genetically defined lines were all of the genotype *oom-*. Stocks of other genotypes were not available within the short time frame of collection. UH1244 (White lady) originated from a cross between progenies from [Uniwai (white) x *A. lindenianum* (white)] and [*A. andraeanum* (white) x *A. antioquiense* (white)] (Kuehnle et al., 2000, 2002). UH1349 and UH1250 originated from a cross between Tropic Mistl (cream coloured) and a selection from the cross between *A. antioquiense* x Marian Seefurth, (pink coloured), (Kuehnle et al., 2000, 2002).

5.2 MATERIAL AND METHODS

5.2.1 Northern analysis

The procedure for northern blotting is described in Chapter 2 (Section 2.14.1 and 2.14.2). RNA from the three genetically defined lines was isolated in Professor Kuehnle's laboratory during a two-week visit. Because of the limited supply of tissue, samples could not be obtained for all stages in each cultivar. Northern analysis was performed using the full-length anthurium cDNA clones for *CHS, F3H, DFR* and *ANS*. Blots were compared with separate blots of RNA from Altar spathe tissue.

5.2.2 F3'H assay procedure

Due to difficulties encountered in isolating a *F3'H* cDNA clone, biochemical assays were conducted to determine if F3'H protein activity is present in anthurium spathe. The assay is based on the procedure for the rapid isolation of microsomal fraction using Mg^{2+} precipitation (Diesperger et al., 1974).

This assay was performed with 1 g of freshly picked tissue. For anthurium (Altar), this comprised a mix of spathe tissue from Stages 2 to 4. For petunia, the positive control, flower limb tissue at Stages 2 and 3 was used. The petunia flower stages were as described in Froemel et al. (1985). The petunia cultivar used was BR140 described in Chapter 3, as being recessive for *F3'H* but dominant for *F3'5'H*. Petunia F3'5'H was shown to work with naringenin as substrate in a similar assay (Menting et al., 1994) and so was considered a valid positive control, as evidenced by the results.

Each extraction was performed with 10 mL of extraction buffer comprised of the following ingredients:

Extraction Buffer
0.1 M Tris-HCl pH 7.5
0.02 M 2β ME

5 μg mL^{-1} pepstain A
1 mM PMSF
0.25 M sucrose
0.25 M mannitol
0.001 M EDTA.
0.5% (w/v) BSA

A bulk stock of 200 mL of extraction buffer (without the 2β ME, PMSF and pepstain A) was made up and stored at room temperature. The remaining ingredients were added at the time when the extraction was being performed. Once the buffer was made, 0.1 g of polyclar AT and 0.5 g of acid-washed sand were added to a pre-cooled (4 °C) mortar. The tissue was added (1 g), along with 5 mL of extraction buffer, and crushed into a homogenous suspension. The homogenate was then filtered through one layer of Mira cloth into a sterile 15 mL Corex tube. This was followed by a 5 min spin at 5,000 rpm in a Sorval SS-34 rotor. The resulting supernatant was transferred to a fresh Corex tube and centrifuged again at 10,000 rpm for 20 min. Once the spin was completed, the supernatant was transferred to a Corex tube and 1M $MgCl_2$ was added to a final concentration of 30 mM to precipitate the microsomal fraction. Tubes were left on ice for 10 min, and then centrifuged at 12,000 rpm for 20 min. The microsomal pellet was then resuspended in 200 μL of assay buffer.

Assay buffer:

0.1 M Tris-HCl
0.0001 M PMSF
0.001 M 2β ME

Four replicate assay reactions were set up for each species, using the fresh extract, and having the necessary negative and positive controls. The assay was set up as follows:

100 μL of resuspended microsomal pellet
5 μL naringenin (1 mg mL^{-1})
5 μL 50 mM NADPH (resuspended in assay buffer)
90 μL assay buffer

The reaction was allowed to run for 2 h and in some instances, overnight, at room temperature. To terminate the reaction two sequential ethyl acetate (500 μL) extractions were performed. The extracts were then dried under O_2-free N_2 and the pellet resuspended in 10 μL of 80% (v/v) methanol. The entire sample was then loaded onto TLC cellulose plates, 2 cm from the base and 3 cm apart from the next sample. Once loaded, samples were dried and ascending chromatography against naringenin and eriodictyol standards was performed with chloroform/acetic acid/water (CAW) 10:9:1 (v/v/v).

5.3 RESULTS

5.3.1 Northern analysis of genotypically defined white lines

Although transcripts for all four genes were detected in the white lines, for two lines they were at reduced levels compared to the cyanic cultivar Altar (Figure 5.1). This pattern was also observed for Acropolis (Figure 4.7). Cultivar 1250 deviates from this pattern with transcript levels that are similar to those of Altar for *CHS, F3H* and *DFR*. *ANS* transcript levels were lower than Altar but higher than in 1244 or 1349.

5.3.2 F3'H assay

Whereas F3'H activity was consistently observed in the petunia positive control, with the conversion of naringenin to eriodictyol, no such activity was observed for the anthurium (Altar) microsomal fraction, even though the procedure was repeated several times with different stages of spathe tissue (Figure 5.2). In addition, no naringenin was detected in the anthurium assays and neither was any seen in the control assay with denatured petunia protein. Additional control experiments were performed without substrate and without protein, both of which had negative results. Finally, an assay was set up with a mixture of protein extracts from petunia and antirrhinum and this efficiently converted naringenin to eriodictyol (data not shown).

Figure 5.1. Northern analysis for testing genetic model. The blots show a comparison of transcript levels in genotypically defined white anthurium cultivars (1349, 1244 and 1250 each with a reported genotype *oom*-) with that Altar. Membranes were blotted separately but probed simultaneously in the same probing solution. Radiolabeled full-length cDNA clones were used as probes. 20 μg total RNA was loaded for each lane and a 25/26S ribosomal RNA probe (pTIP6) from asparagus (King and Davies, 1992) was used as the loading control.

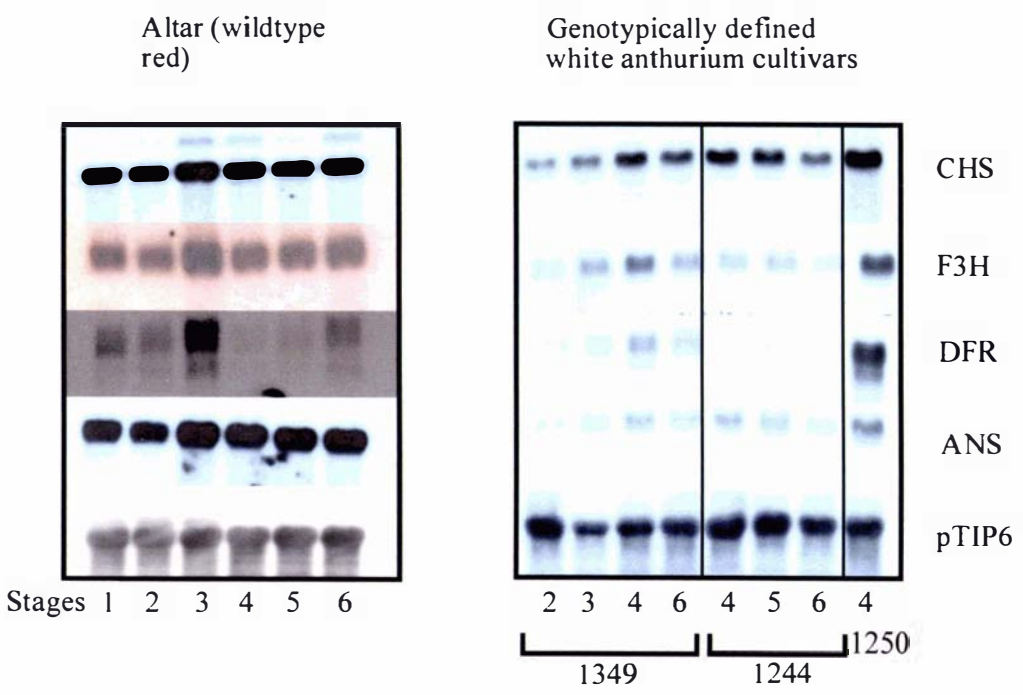
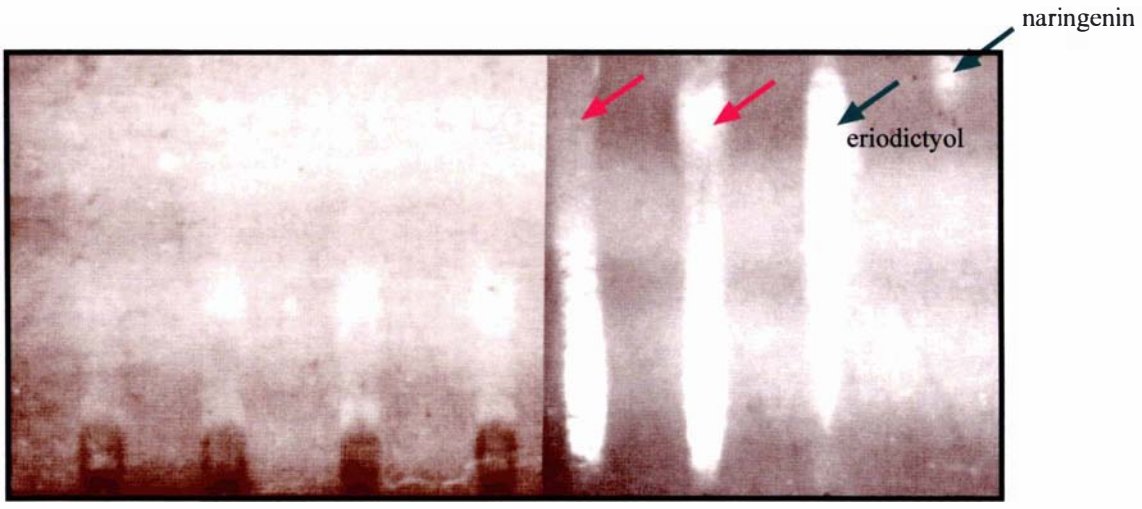

Altar (wildtype red)
Genotypically defined white anthurium cultivars
CHS
F3H
DFR
ANS
pTIP6
Stages 1 2 3 4 5 6
2 3 4 6 4 5 6 4
1250
1349
1244

Figure 5.2. Chromatogram for F3'H assay. A1-A4 are the anthurium replicates (using Altar spathe tissue), while P1 and P2 are the petunia positive controls. The positions of the two flavonoid standards S1 (eriodictyol) and S2 (naringenin) are indicated by green arrows and red arrows point to the eriodictyol formed in the petunia assay.



A1 A2 A3 A4 P1 P2 S1 S2

5.4 DISCUSSION

5.4.1 Investigating the nature of the *O* locus in anthurium

There was a marked decline in transcripts for all four genes (*CHS, F3H, DFR* and *ANS*) over the various stages of spathe development in Acropolis (Figure 4.7) and cultivars 1349 and 1244 (Figure 5.1) when compared with Altar, although significant levels of *CHS* transcripts are still present. Therefore, the white phenotypes observed in Acropolis, 1349 and 1244 are most likely due to a mutation in a regulatory gene for anthocyanin biosynthesis. Given that 1349 and 1244 are recessive for the *O* locus, the data suggest that this locus could correspond to the affected regulatory gene. Although 1349 and 1244 may be homozygous recessive at the *M* locus, *mm* does not eliminate pigment formation, as cultivars that are *O-mm* have orange/coral spathe (Kamemoto et al., 1988). Therefore, the results suggests that the *oo* genotype in white lines results in reduced expression levels of several flavonoid genes in the spathe, and is possibly the main reason for the absence of pigment in these lines. However, this is contradicted by the 1250 data, if the genotyping is correct.

If *O* is a regulatory gene controlling the expression of all the flavonoid biosynthetic genes in the anthurium spathe, then its mutation should result in a reduction in flavone levels in white anthurium. Interestingly, the absence of anthocyanins does not result in either a decrease or a greater production of flavones in white lines, because in related work, similar quantities of flavones were detected in anthurium whites as in coloured lines (David Lewis, personal communication). Although the genetic make up of the lines tested was not known, one interpretation of the data is that for anthurium the regulation of flavone biosynthesis is separate from that of anthocyanins. Therefore, the existence of more than one regulatory gene controlling flavonoid accumulation in anthurium is likely. In fact, in other species the regulators of anthocyanin biosynthesis do not regulate other branches of the pathway. As a result the regulation of flavonoid biosynthesis in anthurium may have similarities to species such as maize (Grotewold et al., 1994) and antirrhinum (Moyano et al., 1996) with independent regulators for the different branches of the flavonoid pathway.

Should a separate regulator exist in anthurium spathe for flavones, it may target *CHS* and *CHI* directly. The fact that the *CHS* transcript levels are reduced in the spathe along with *DFR* and *ANS* suggests the existence of common regulatory elements. However, significant levels of *CHS* transcripts are still produced, which may indicate the activity of other regulatory proteins. The availability of an anthurium *CHI* cDNA clone would be useful in exploring this possibility. From this preliminary effort, it appears that the actual genetics of flower colour inheritance in anthurium tissue appears more complex than the dihybrid inheritance model suggested by Kamemoto et al. (1988).

Higher transcript levels are detected in line 1250 for *CHS, F3H and DFR* than in lines 1349 and 1244 (Figure 5.1), and only *ANS* transcript levels appear to be markedly reduced compared to Altar. It would be useful to monitor transcript levels in more than one stage for 1250. However, the high transcript levels of *CHS, F3H* and *DFR* seems to suggest that this cultivar may be a structural mutant in one of the genes acting downstream from *DFR* such as *ANS, UFGT* or *GST*.

It may also be possible that the genotype for 1250 is incorrect. However, if 1250 is in fact *oo* homozygous recessive, then it seems unlikely that *O* can encode a regulatory factor necessary for *CHS, F3H* or *DFR* expression. Thus, the regulatory mutation in some white cultivars may be in a locus distinct from *O*. Alternatively, they may carry distinct *O* alleles, which have markedly differing effects in their various mutant forms. In this regard the *O* locus for anthurium may be a complex loci as seen with the antirrhinum anthocyanin regulator ROSEA (Schwinn, 1999). Again, because white lines accumulate flavones, *CHI* activity is assumed. Although the presence of the transcript does not always correlate to a functional protein, post-transcriptional modification does not appear to be a significant mechanism for regulating the expression of genes involved in phenylpropanoid biosynthesis (Martin et al., 1991). Therefore, it is likely that the transcripts being detected for *CHS, F3H, DFR* and *ANS* for 1250 do encode functional proteins. Of course a point mutation affecting protein function cannot be ruled out.

Further work is needed to make conclusive statements on the identity of the gene product coded for by the *O* locus. The cloning of the anthurium anthocyanin regulatory genes would greatly assist this objective. The primary limitation to this experimental effort was the location of stock plants in Hawaii. This meant limited RNA samples prevented experiments from being repeated or developed further.

5.4.2 Investigating the nature of the *M* locus in anthurium

To test the proposal that *M* encodes the *F3'H*, an anthurium *F3'H* clone was needed, as well as lines genotypically defined for the *M* locus. The first phase of the strategy was to isolate an anthurium *F3'H* cDNA clone and perform a comparative northern or PCR analysis with orange-coloured anthurium lines having an *O_mm* genotype and red-coloured lines of the *O_MM* or *O_Mm* genotype. If the model held true then, when probed with anthurium *F3'H*, there should be no (or at least very little) transcript in orange lines with genotype *O_mm* compared to levels in the red (*O_M_*) coloured spathe. Of course this approach to test the model assumes the predicted mutation in *F3'H* is a null mutation and therefore does not allow a full-length transcript to be produced.

As discussed in Chapter 3, no cDNA clone encoding the anthurium F3'H protein was isolated from the cDNA library or by PCR approaches. Difficulty in securing F3'H clones has been encountered for other species (Kevin Davies, personal communication).

As an alternative, experiments were performed to determine if F3'H enzyme activity existed in anthurium using tissue from Altar spathe. When compared to the petunia positive control, no definitive F3'H protein activity was detected for anthurium spathe tissue (Figure 5.2). It would, however, be useful to compare the amount of the protein in the extracts of both anthurium and petunia to determine that the extraction procedure was effective for anthurium tissue. Naringenin substrate was not detected for the anthurium assays, and was only found in the boiled controls, suggesting that the absence of naringenin is related to enzyme activity. Therefore, one possibility explaining the absence of naringenin in anthurium assay reactions is conversion to flavones by FNS or F2H as both are

microsomally located. It is possible that FNS or F2H were very active in the extract, removing all the substrate from the endogenous F3'H.

It is possible that there could have been enzyme-catalysed conversion to eriodictyol but at undetectable levels, either because of low yield of active protein, or the presence of an inhibitor within the anthurium tissue extract. If low levels of product for the anthurium assay are assumed, then testing alternative solvents for improved separation and resolution may prove worthwhile. Lengthening the time frame for incubation beyond 2 h, to allow more products to accumulate, was attempted. However, incubation times as long as 24 h were tested without success (data not shown). To determine if inhibitor compounds co-precipitated with the anthurium protein extract, assays were performed with an equal mixture of anthurium and petunia extract. This protein mixture converted naringenin to eriodictyol with the same efficiency as an assay using only petunia extract (data not shown), suggesting that there were no inhibitory compounds in the anthurium extract.

There are two alternative interpretations of the results aside from the failure of the extraction assay procedure. The first is the possibility that anthurium F3'H cannot utilise naringenin as substrate (either not at all or not efficiently). The second is the possibility that anthurium spathe tissue uses an alternative biochemical route for the synthesis of precursors for dihydroquercetin. While 4-coumaroyl-CoA is the primary physiological substrate for chalcone formation, some CHS enzymes can accept caffeoyl-CoA, forming eriodictyol chalcone (Hellar and Forkmann, 1988, 1993). In this way, the caffeoyl-CoA could provide both hydroxyl groups on the B-ring, eliminating the need for endogenous F3'H activity in the formation of dihydroquercetin. Further experimentation is necessary to investigate these possibilities and the possibility that M could encode 4-coumarate: CoA 3-hydroxylase.

Chapter 6

Cloning and characterisation of a putative anthocyanin regulator from anthurium

6.1 INTRODUCTION

The anthurium flower is a modified leaf with no such equivalent in the model species studied to date. In addition to this, anthurium is a monocot plant with a much longer floral developmental time than ornamental models such as petunia and antirrhinum. How these characteristic developmental features are reflected in the expression pattern of anthocyanin structural genes was investigated in Chapter 4. As anthocyanin structural genes are transcriptionally regulated (Paz-Ares et al., 1987; Martin et al., 1991; Avila et al., 1993; Cone et al., 1993; Grotewold et al., 1994), the expression pattern of these genes gives preliminary insight into the regulation of anthocyanin production in anthurium spathe tissue. In Chapter 5, preliminary data were presented suggesting that the *O* locus in anthurium encodes a regulatory gene. Studies with model species such as maize, antirrhinum and petunia have shown that Myb and bHLH proteins are common transcription factors involved in regulating anthocyanin biosynthesis. Consequently, it was reasoned that similar proteins might be involved in anthurium spathe tissue. The objective of this part of the thesis was to initiate investigations into the regulation of anthocyanin pigmentation in anthurium spathe tissue, towards developing a comprehensive understanding of this process in this unusual ornamental plant. The first step in such an initiative is the cloning and characterisation of key anthocyanin regulatory genes.

No *Myb* or *bHLH* anthocyanin regulatory genes have been previously isolated from an ornamental monocot species. This chapter describes the isolation of several cDNAs encoding a portion of the Myb DNA binding domain and their categorisation into distinct Myb families. A full-length cDNA clone for one group was subsequently cloned and the predicted amino acid sequence of the protein analysed to identify a potential activation domain and other features. This cDNA was further analysed and Chapter 7, describes its temporal expression pattern and the functional analysis that was performed using transient expression assays aimed at complementing an antirrhinum MYB regulatory mutant line. In addition, the promoter for the anthurium *DFR* gene was cloned and promoter-reporter experiments were performed with antirrhinum wild type and regulatory mutants.

6.2 MYB ISOLATION STRATEGY

Since transcription factors are expressed at low levels within the cells, homologous screening was considered to be the best approach to secure a full-length clone. With this in mind, RT-PCR combined with 3' RACE technology (Frohmann et al., 1988) was used to generate a partial clone, which was then used as a homologous probe in library screens. Since, the Myb domain is highly conserved in sequence, the first step was to perform PCR reactions using degenerate primers designed to this region (Romero et al., 1998). The layout of these is shown in Figure 6.1. Each primer pair was expected to generate a 180 bp product. The degenerate primers, reverse primers N1-N6 and forward primers C1-C2, are listed in Appendix I.

The primers were designed to recognise most of the known Myb groups including those involved in regulating anthocyanin biosynthesis. The PCR reactions cloned several 180 bp cDNA fragments encoding a portion of the Myb domain. The second step focused on the cloning of the 3' end of selected Myb fragments using 3' RACE, that were then used as radiolabeled probes for homologous library screening. The 3' end is usually very specific for any given gene and therefore allows greater stringency to be used in the library screen.

6.2.1 RT-PCR reactions with Myb degenerate primers

The template for the RT-PCR reactions was first-strand cDNA. Expand-reverse transcriptase was used to catalyse the synthesis of the cDNA from RNA pooled together from stages 1 to 6 of spathe development from Altar. The reaction was primed using the same 35-mer used by Frohman et al. (1988). Reactions for the first strand cDNA synthesis were performed as described in Section 2.9.3.

PCR reactions were set up with *Taq* DNA polymerase as described in Section 2.9.1 using 1-5 µL of the cDNA pool and 0. 5 µL of *Taq* DNA polymerase in at total reaction volume of 50 µL. All 18 primer combinations (N1-N6 x C1-C3) were used in the first attempt, and for all PCR reactions a negative control (no template) and a positive control (*Roseal* in

Figure 6.1. Degenerate primers for Myb RT-PCR reactions. The schematic shows the location of the degenerate primers designed to the recognition helix region, a conserved region of the Myb domain (Avila et al., 1993; Romero et al., 1998). The primers are listed below those that were most effective at amplifying anthurium Myb cDNAs (N6 x C2 and N2 x C2) are underlined.

The reverse primer sequences are:

N1: 5' CGGAATTCDKNAARAGYTGYAG 3'

N2: 5' CGGAATTCDKNAARAGYTGYCG 3'

N3: 5' CGGAATTCDKNAARTCNTGYAG 3'

N4: 5' CGGAATTCDKNAARTCNTGYCG 3'

N5: 5' CCCGGGTGYGGNAARTCNTG 3'

N6: 5' GGAATTCTGYGGNAARAGYT 3'

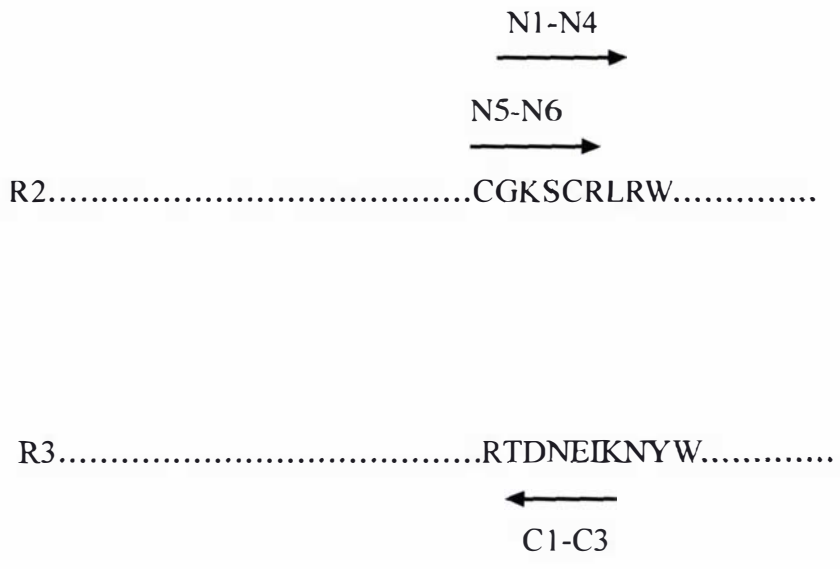
The forward primer sequences are:

C1: 5' CGGAATTCTTNAYNGCRTTRTCNGT 3'

C2: 5' CGGAATTCTTDAYYTCRTTRTCNGT 3'

C3: 5' CGGAATTCTTNAYNTKRTTRTCNGT 3'

The degenerate code used: R=A+G, Y=C+T, S=G+C, D=A+G+T and N=A+G+C+T


N1-N4
N5-N6
R2.....CGKSCRLRW.....
R3.....RTDNEIKNYW.....
C1-C3

pART7 vector) were used. The reactions were subject to PCR under the following conditions: 35 cycles of 95 °C/30 s, 50 °C/1 min, 72 °C/30 s and one cycle at 72 °C/3 min.

PCR products were electrophoresed on a 2% (w/v) TBE/EtBr gel with a 100 bp ladder and products of the desired size were gel purified as described in Section 2.10 and re-amplified over 25 cycles using the same PCR parameters as above. The gel purified, amplified PCR product was then ligated into the pGEM-T Easy vector as described in Section 2.12. Heat shock transformation was performed on ligated plasmid (Section 2.13) and 40-60 white colonies were selected and cultured for plasmid DNA extraction by alkaline lysis (Section 2.6.1). Inserts were released with a single digest using 0.5 µL of *Eco* R1(10 U μl^{-1}), 4 µl of DNA, 1 µL of buffer and 4 µL of water. Reactions were incubated at 37 °C for 1 h then separated on a 2% (w/v) TBE/EtBr gel for 1 h at 100 V. A range of colonies, with insert sizes from 180 bp-200 bp, was selected for high quality DNA preparation using Qiagen Mini Prep Kit (Section 2.6.2). These samples were commercially sequenced using the ABI automated DNA sequencer using the M13 reverse primer.

6.2.2 Cloning the 3' end of Myb cDNAs

Several *Myb* related clones (180 bp fragments) were obtained from the RT-PCR approach. Clones were grouped into six families as discussed in the results. Myb groups C and F were of most interest. Therefore, based on the sequence data, specific primers for nested PCR were designed for both groups of clones, and 3' RACE was performed using the same stock of first strand cDNA for template as used previously. The 35-mer oligonucleotide that was used initially to prime the RT reaction was also used as the C-terminal anchor for the 3' RACE PCR reactions. To perform the nested PCR, the PCR product amplified with the first primer pair was gel purified as described in Section 2.10 and an aliquot used as the template for the second round of PCR reactions using a separate primer pair. Similarly the product from the second round of PCRs, once purified from the agarose gel, was used as template for the third round of PCR reactions using a third primer pair.

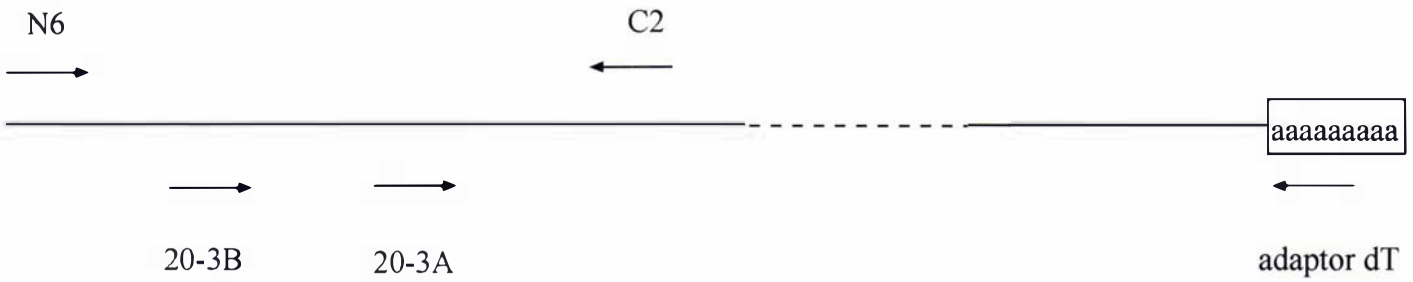
The primer pairs used in the nested PCR approach for Myb Group C clones were Myb N6 x adaptor-dT_{17} followed by Myb 20-3B x adaptor-dT_{17} and Myb 20-3A x adaptor-dT_{17} (Figure 6.2). To amplify the 3' end of Group F *Myb* clones the sequence of primers used for nested PCR were Myb N6 x adaptor-dT_{17} followed by Myb 20-5/18B x adaptor-dT_{17} and Myb 20-5/18A x adaptor-dT_{17}. The sequences of these primers are in Appendix I and a schematic showing the layout of the primers is shown in Figure 6.2.

For both Group C and Group F nested PCR reactions, a positive control reaction was performed using *Rosea1* in pART7 as template and RosSp x adaptor-dT_{17} as the primer pair. The RosSp primer (listed in Appendix I) is a gene specific primer for *Rosea1* located in the equivalent Myb domain sequence position on *Rosea1* to those of the anthurium nested PCR primers. This positive control PCR gave a band size of approximately 500 bp with the primer pair. As *Rosea1* is a *Myb* anthocyanin regulator, products of similar size were targeted for the reactions with anthurium *Myb* clones.

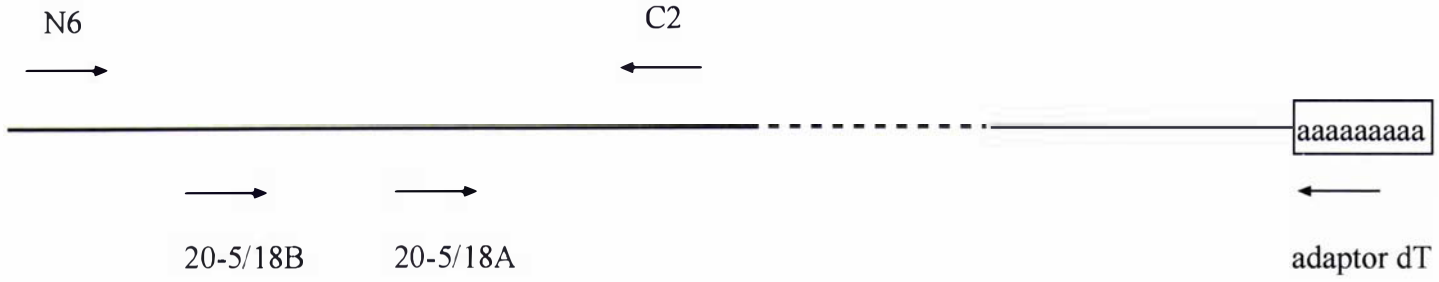
The PCR reactions were set up as described in Section 2.9.3 with 1-5 µL of cDNA from the stock and 0.5 µL of *Taq* DNA polymerase. The PCR cycling parameters were: 95 °C/30 s, 55 °C/1 min, 72 °C/30 s ending with one cycle at 72 °C/3 min. PCR products were gel purified and ligated into pGEM-T Easy vector then transformed into XL1-Blue MRF' and plated on L-Agar plates with 40 µL X-Gal (50 mg mL^{-1}), 20 µL IPTG (50 mg mL^{-1}) and 100 µg mL^{-1} ampicillin. Plasmid DNA was extracted from transformants and the plasmids yielding inserts of 500-600 bp were used as templates (1-2 µL) for further evaluation by PCR prior to sequencing. PCR reactions were set up using Myb C2 and 20-3A as the primer pair for the Group C candidates whereas Myb C2 and 20-5/18A was the primer pair used to test the Group F candidates (Figure 6.2). PCR reactions were similar to those described above except that the number of cycles and the annealing temperature were reduced to 25 cycles and 55 °C, respectively. Those plasmids giving a positive PCR product (80-100 bp) were sent for sequencing.

Figure 6.2. Primer layout for nested- 3' RACE Myb PCR reactions. The relative positions of primers for amplifying the 3' end of the Group C (Panel A) and F (Panel B) anthurium *Myb* clones are shown. The specific primers used for Group C *Myb* were 20-3B followed by 20-3A and for Group F clones 20-5/18B followed by 20-5/18A. The relative positions of the degenerate primers (N6 and C2) from Romero et al. (1998) are also shown. PCR products of the required size were evaluated prior to sequencing using 20-3A x C2 for Group C candidates and 20-5/18A x C2 for Group F candidates. The sequences of the primers are in Appendix I.

A


N6
C2
aaaaaaaaa
20-3B
20-3A
adaptor dT

B


N6
C2
aaaaaaaaa
20-5/18B
20-5/18A
adaptor dT

6.2.3 Cloning full-length cDNAs for anthurium *Myb* Groups C and F

This involved homologous screening of the amplified spathe cDNA library with the partial *Myb* clones of Group C and F. Each aliquot from the library consisted of approximately 50,000 plaques and was plated out in L-Top agarose on 20 agar plates (giving a total of 1 million plaques) then incubated overnight at 37 °C. Screening was performed as described in Section 2.5.6 but with two adjustments. For this library screen, duplicate lifts were performed on each plate. In addition, a higher stringency was used with the hybridisation and washing of the membranes. A temperature of 65 °C was used and the membranes washed with 1 x SSC (w/v) /0.1% SDS (w/v) followed by 0.1 x SSC (w/v) /0.1% SDS (w/v), both for 15 min. Positives obtained in the primary screen were taken through secondary, and on occasion tertiary screens, to obtain a single plaque pick. Subsequently, in vivo excision was performed as described in Section 2.5.7 and single colonies were cultured for plasmid DNA extraction. cDNA Inserts ranging from 800-1,500 bp were obtained and sent for sequencing with the M13 reverse primer.

6.3 RESULTS

6.3.1 Isolating cDNAs for Myb genes from anthurium spathe tissue

In the first phase of isolating a cDNA for an anthurium anthocyanin Myb transcription factor, degenerate primers for the highly conserved Myb binding domain were used in a RT-PCR approach. Of the eight primer combinations two (N6 x C2 and N2 x C2) gave distinct PCR products of the expected size, 180 bp. The remaining primer pairs either gave very faint products or none at all. Products from the N6 x C2 and N2 x C2 reactions were cloned, and a total of 23 PCR clones were sequenced. Analysis of the DNA and amino acid sequences showed that they represented several *Myb* genes. The clones shared an average of 75% identity with each other and separated out into six groups (Group A to Group F) based on deduced amino acid and DNA sequence differences (Figure 6.3). Each group scored hits for different Myb proteins on the basis of BLAST searches of the NCBI database. The equivalent regions with the respective Myb protein groups are aligned in

Figure 6.3 A comparison of the deduced amino acid sequence for Myb Groups A-F. The cDNA fragment for each Myb was obtained by RT-PCR using degenerate primers N6 x C2 and N2 x C2 designed to the recognition helix sequence of several Myb sequences (Avila et al., 1993; Romero et al., 1998) . Black shading indicates identical amino acids. Dashes indicate gaps in the sequence for alignments. The location of portions of the R2 and R3 domains is marked, along with the positions of the two primers that were successful in securing anthurium *Myb* clones. The filled area beneath the diagram identifies the point of greatest variation among the clones and corresponds to the border region of the two repeats in the Myb domain. The two candidates for Myb Group F are shown and they differ only by a single amino acid residue that is marked by an asterisk.

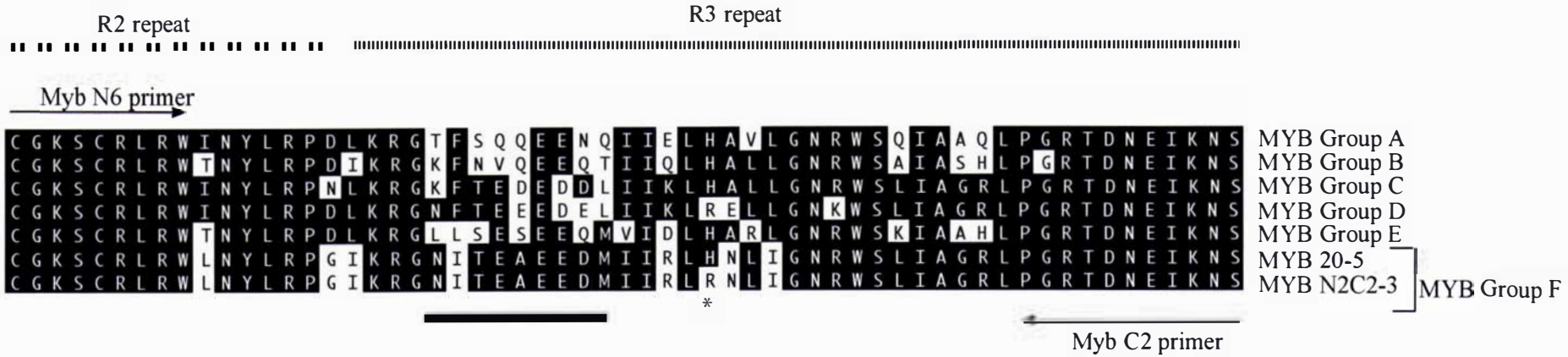

R2 repeat
R3 repeat
Myb N6 primer
CGKSCRLRWINYLRPDLKRGTFSQQEENQIIELHAVLGNRWSQIAAQLPGRTDNEIKNS
CGKSCRLRWTNYLRPDIKRGKFNVQEEQTIIQLHALLGNRWSAIASHLPGRTDNEIKNS
CGKSCRLRWINYLRPNLKRGKFTEDEDDLIIKLHALLGNRWSLIAGRLPGRTDNEIKNS
CGKSCRLRWINYLRPDLKRGNFTEEEDELIIKLRELLGNKWSLIAGRLPGRTDNEIKNS
CGKSCRLRWTNYLRPDLKRGLLSESEEQMVIDLHARLGNRWSKIAAHLPGRTDNEIKNS
CGKSCRLRWLNYLRPGIKRGNITEAEEDMIIRLHNLIGNRWSLIAGRLPGRTDNEIKNS
CGKSCRLRWLNYLRPGIKRGNITEAEEDMIIRLRNLIGNRWSLIAGRLPGRTDNEIKNS
MYB Group A
MYB Group B
MYB Group C
MYB Group D
MYB Group E
MYB 20-5
MYB N2C2-3
MYB Group F
*
Myb C2 primer

Figure 6.4 A-C and D-F, and the nucleotide sequence for each Myb Group (A-F) is in Appendix VI and VII.

Myb Group A showed highest % identity (93%) to arabidopsis MYB4 (AtMYB4), compared to 71% with maize C1. AtMYB4 was shown to be a repressor of cinnamate 4-hydroxylase activity in arabidopsis (Jin et al., 2000). The alignment of the deduced amino acid sequence of the three anthurium clones that comprised this group (Myb 20-17, 20 and N2C2-1) showed an exact match (data not shown), and the minor differences that were in the DNA sequence of the clones were not due to genomic DNA variation as they occurred in the primer binding sites (Figure 6.4A).

Myb Group B showed highest % identity to arabidopsis MIXTA-like (95%) and antirrhinum MIXTA (92%), and their alignment is shown in Figure 6.4B. In antirrhinum, MIXTA is involved in determining cell shape of petal epidermal cells (Noda et al., 1994). Interestingly, the clones from arabidopsis and antirrhinum had only 75% identity in this region. The two candidates in Group B (*Myb* 20-19 and 20-2) were identical at the amino acid level.

Myb Group C showed highest % identity (86%) to maize ZM 38, the leaf specific C1-related Myb expressed in leaves (Marocco et al., 1989) (Figure 6.4C). The two clones comprising this group (*Myb* 20-3 and 20-P) were 100% identical at the amino acid level. The variations seen at the DNA level occurred in the primer binding sites.

Myb Group D showed highest % identity to barley HV 1, which is the C1 related cDNA from barley leaves (Marocco et al., 1989). Only a single candidate (*Myb* 20-39) from the 23 clones sequenced belonged to this group. It had 93% identity to HV1, but it showed only 76 % identity to maize C1 (Figure 6.4D). It was very different from the Group C and F clones, sharing 77% and 88% identity, respectively.

Myb Group E showed highest % identity to Myb related protein2 from *P. patens* (Leech et al., 1993). The function of this Myb protein has not yet been characterised. The clones in this group (20-6, 1 and N) had identical amino acid sequences (Figure 6.4E).

Figure 6.4 A-C. Amino acid alignment Myb Groups A-C with the relevant region of the Myb proteins to which they are most similar. C1 is included in all comparisons as this is the only anthocyanin Myb protein to which the clones show reasonable similarity. Amino acid sequences from plant sources other than anthurium, in the alignment and their accession numbers are: AtMYB4 (Y099777), AtMIXTA-like (X99809), AmMYBMIXTA (X79108), ZM38 (P20025) and ZmMYBC1 (P10290).

A


1 CGKSCRLRWINYLRPDLKRGTFSQQEENQIIELHAVLGNRWSQIAAQLPGRTDNEIKN
49 CGKSCRLRWINYLRPDLKRGAFSQDEENLIIELHAVLGNRWSQIAAQLPGRTDNEIKN
49 CGKSCRLRWLNYLRPNIRRGNISYDEEDLIIRLHRLLGNRWSLIAGRLPGRTDNEIKN
MYB Group A
AtMYB4
ZmMYBC1

B


1 CGKSCRLRWTNYLRPDIKRGKFNVQEEQTIIQLHALLGNRWSAIASHLPKRTDNEIKN
49 CGKSCRLRWTNYLRPDIKRGKFNLQEEQTIIQLHALLGNRWSAIATHLPKRTDNEIKN
49 CGKSCRLRWANYLRPDIKRGPFSLQEEQTIIQLHALLGNRWSAIASHLPKRTDNEIKN
49 CGKSCRLRWLNYLRPNIRRGNISYDEEDLIIRLHRLLGNRWSLIAGRLPGRTDNEIKN
MYB Group B
AtMIXTA like
AmMYBMIXTA
ZmMYBC1

C


1 CGKSCRLRWINYLRPNLKRGKFTEDEDDLIIKLHALLGNRWSLIAGRLPGRTDNEIKN
49 CGKSCRLRWINYLRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKN
49 CGKSCRLRWLNYLRPNIRRGNISYDEEDLIIRLHRLLGNRWSLIAGRLPGRTDNEIKN
MYB Group C
ZM38
ZmMYBC1

Figure 6.4 D-F. Amino acid alignment Myb Groups D-F with the relevant region of the Myb proteins to which they are most similar. Maize C1 is included in all comparisons as this is the only anthocyanin Myb protein to which the anthurium Myb proteins show reasonable similarity. Amino acid sequences from plants sources other than anthurium, in the alignment and their accession numbers are: HV1 (P20026), PpMyb2 (P80073), HV33 (P20027) and ZmMYBC1 (P10290).

D


1 CGKSCRLRWINYLRPDLKRGNFTEEEDELIIKLRELLGNKWSLIAGRLPGRTDNEIKN
49 CGKSCRLRWINYLRPDLKRGNFSHEEDELIIKLHSLLGNKWSLIAGRLPGRTDNEIKN
49 CGKSCRLRWLNYLRPNIRRGNISYDEEDLIIRLHRLLGNRWSLIAGRLPGRTDNEIKN
MYB Group D
HV1
ZmMYBC1

E


1 CGKSCRLRWTNYLRPDLKRGLLSESEEQMVIDLHARLGNRWSKIAAHLPGRTDNEIKN
49 CGKSCRLRWTNYLRPDLKRGIFSEAEENLILDLHATLGNRWSRIAAQLPGRTDNEIKN
51 CGKSCRLRWINYLRPDLKRGCFSQQEEDHIVALHQILGNRWSQIASHLPGRTDNEIKN
49 CGKSCRLRWLNYLRPNIRRGNISYDEEDLIIRLHRLLGNRWSLIAGRLPGRTDNEIKN
MYB Group E
PpMYB2
HV33
ZmMYBC1

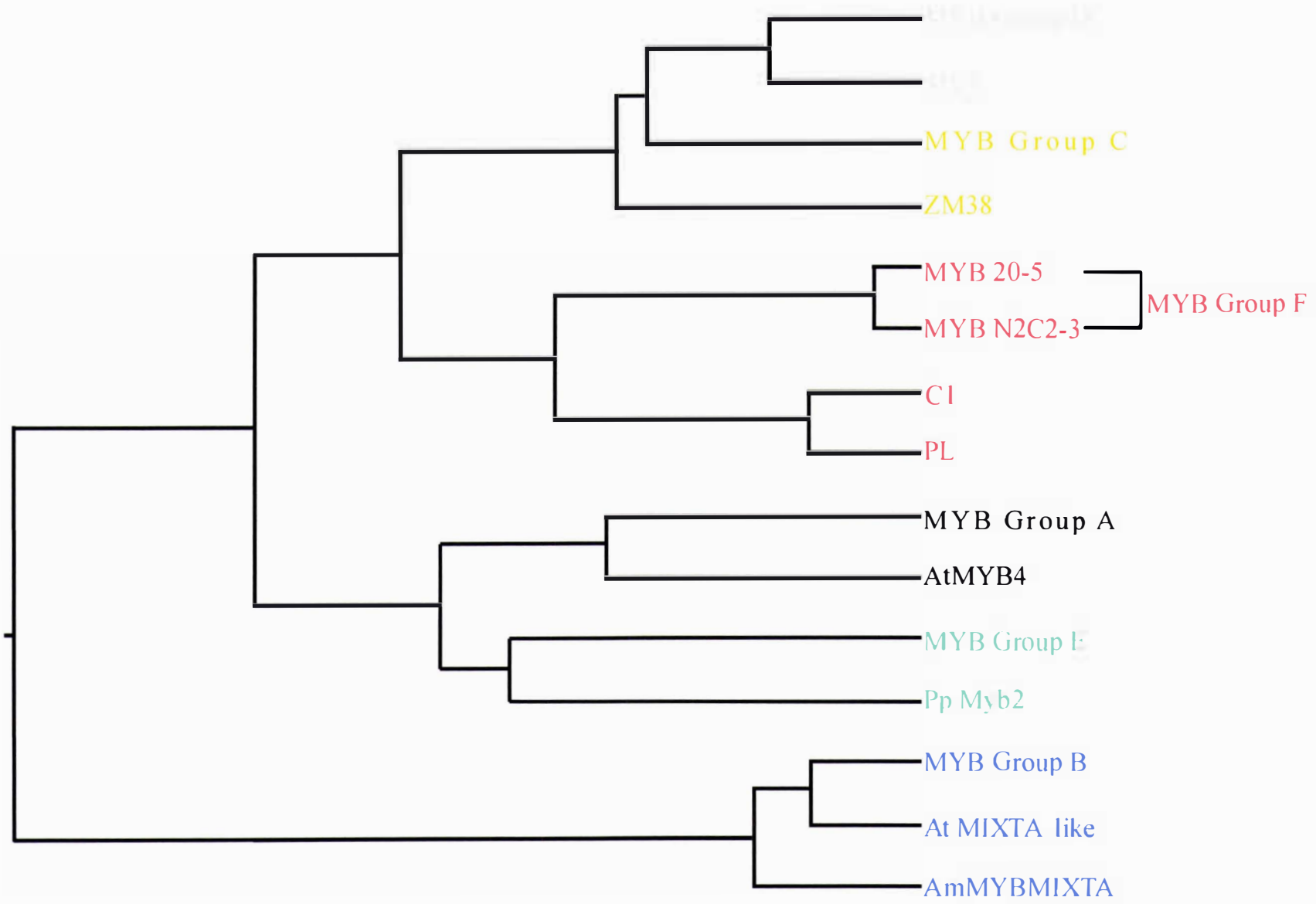
F


1 CGKSCRLRWLNYLRPGIKRGNITEAEEDMIIRLHNLIGNRWSLIAGRLPGRTDNEIKN
1 CGKSCRLRWLNYLRPGIKRGNITEAEEDMIIRLRNLIGNRWSLIAGRLPGRTDNEIKN
49 CGKSCRLRWLNYLRPNIRRGNISYDEEDLIIRLHRLLGNRWSLIAGRLPGRTDNEIKN
MYB 20-5
MYB N2C2-3
MYB Group F
ZmMYBC1

Myb Group F had the largest number of clones with 52% of the sequenced PCR products belonging to this group. It showed highest % identity to the anthocyanin Mybs maize C1 and PL. There were 12 members in this group (N2C2 - 2, 3, 4, 5, 6, 7, 11, 13, 15 and 20-5/18/O) and one possible allelic difference. The difference was seen in three N2C2 clones, N2C2- 2, 3 and 5. Whereas these three clones were identical, they differed from the other members of Group F at the 104 bp mark having a CGA instead of CAA in the coding sequence. These code for the basic amino acids arginine and histidine, respectively, and accounts for the difference at position 36 for the two Myb Group F clones in Figure 6.4F.

The distinctive features of each group and the Myb proteins to which they are most similar are represented in the phylogenetic tree shown in Figure 6.5. Of all the Myb protein fragments, Group E was the only one with highest % identity to a Myb protein, yet uncharacterised, from prokaryotes. While all the cDNA fragments and their respective deduced amino acid sequences had significant % identity to interesting Myb proteins, the scope and objective of the thesis meant that efforts needed to be focused on isolating *Myb* cDNAs that are for potential anthocyanin regulators in anthurium spathe. On this basis, Myb Group F was the prime target for further analysis, as it shared the highest similarity, in the Myb domain, to a known anthocyanin regulator (maize C1), suggesting that it may recognise similar target genes in anthurium spathe. Also selected, were the Myb Group C and D candidates, which showed highest % identity to ZM 38 and HV 1. The precise roles of ZM 38 and HV 1 have not yet been defined but they are both active in leaf tissue (Marocco et al., 1989). In addition, the Group C clones had the second highest % identity to maize C1. Specific nested primers were designed to these cDNA fragments based on the sequence data and 3' RACE PCR performed to secure the 3' end of each group. This was the chosen approach as the 180 bp cDNA fragments were two small to be radiolabeled and used as probes to screen the library. The reactions were only successful for Group C and F and these results are described in the following section.

Figure 6.5. Phylogenetic relationships of deduced anthurium Myb amino acid sequences to those of selected Myb proteins from other species. The phylogenetic tree is based on alignments of anthurium R2R3 Myb fragments, generated from RT-PCR with degenerate primers designed by Romero et al. (1998), to the equivalent region in other plant Mybs. The alignment was done using the MegAlign programme (gap penalty 10, gap length penalty 10) from the DNASTAR sequence analysis suite of software. The Myb groups A-F sequences were from this thesis and accession numbers for the sequences are: HV1 (P20026), ZM38 (P20025), ZmMYBC1 (P10290), ZmMYBPl (AAB67720), AtMYB4 (Y099777), PpMYB2 (P80073), AtMIXTA-like (X99809) and AmMYBMIXTA (X79108).


MYB Group C
ZM38
MYB 20-5
MYB N2C2-3
MYB Group F
C1
PL
MYB Group A
AtMYB4
MYB Group I
Pp Myb2
MYB Group B
At MIXTA like
AmMYBMIXTA

6.3.2 The 3' end of Group C and F

Three rounds of nested PCR with 40 cycles for each round had to be performed to get sufficient amounts of the correct PCR product. For both groups, the location of the last nested primer for the third round of PCR was such that a sufficient overlap distance with the original 180 bp fragments was present to determine if the correct clone was amplified. Three clones for each group were selected for sequencing. Both the primer sites and a polyadenylated tail were identified in all Group F clones, and clones in each of the two groups had identical sequences. The Group C amino acid and DNA sequence showed highest similarity to ZM 38 and HV1 from a BLAST Search, while the Group F sequence had maize C1 as its first hit.

6.3.2.1 Analysing the C-terminus sequence of Myb Group C

The first thing that stood out with the C-terminus of the Group C clone was its smaller size in comparison to the other clones. Counting from the end of the R3 repeat, the C-terminus of the Group C clone had 99 amino acid residues, whereas from the same point in C1, HV1 and ROSEA1 it was 158, 152 and 111 amino acid residues, respectively.

Unlike the Myb domain, the C-terminal end was quite divergent when compared to other Myb proteins. From 87% and 84% identity in the Myb domain to ZM 38 and HV 1 respectively, the C-terminus of the anthurium Group C clone showed only 30 and 32% amino acid conservation to the C-terminus of ZM 38 and HV 1, respectively. This figure declined to as low as 21% with ROSEA1. In addition, HV 1 and ZM 38 shared blocks of conserved amino acids in their C-terminal ends (Marocco et al., 1989) but no such conservation was seen with the C-terminal end of the Group C protein.

The presence of long stretches of negatively charged acidic amino acid residues and the formation of an amphipathic α-helix has been suggested as two essential elements of the transactivation domain (Giniger and Ptashne, 1987; Ptashne, 1988). Residues that are predicted to form an amphipathic α-helix in the transactivation domain have been identified for other anthocyanin regulators (Paz-Ares et al., 1990; Avila et al., 1993; Schwinn, 1999).

While some acidic hydrophobic amino acids are present in the C-terminus of anthurium Group C that can form α-helices, these were not amphipathic (data not shown). Attempts to clone a full-length cDNA for the Group C *Myb* proved unsuccessful. No hybridisation signals were obtained with a primary library screen of 1 million plaques using a radiolabeled Group C fragment.

6.3.2.2 Analysing the C-terminus sequence of Myb Group F

Unlike the C-terminus for the Group C candidate above, the C-terminus of Myb Group F clone (called *AaMYB1*) had 120 amino acids, and so was comparable in size to that of C1, HV1 and ROSEA1. However, like the C-terminus for the Group C sequence, it differed significantly from the C-terminus of the other Myb proteins, having no common conserved motifs.

The C-terminus of AaMYB1 had a region (amino acids 134-240) in which the acidic amino acids (aspartic acid and glutamic acid) are concentrated (Figure 6.6). Using PROTEASE analysis software from the DNASTAR package, two significant α-helices, from residues 171-186 and 236-244, of high probability were predicted in this domain. Helix Wheel analysis from ExPASy Proteonomics database (www.expassy.ch) predicted that both helices are amphipathic with hydrophobic residues grouping on one face of the helix and hydrophilic residues lining up on the other face (Figure 6.7 and 6.8).

6.3.3 Isolating a full-length cDNA for *AaMyb1*

Seven independent clones were obtained by screening the library with a probe made from the 3' end of *AaMyb1*. Obtaining seven independent positives from the library screen for the anthurium Group F *Myb* clone fits well with the large number of PCR (180 bp) fragments that were isolated for this group in the first stage of this cloning effort. Five of these clones were of similar size (approximately 1 kb). The other two clones (C7 and 9A) were 1.5 kb and 0.8 kb respectively. All the clones except C7 were identical in sequence. However, clone 9A lacked the ATG start codon. Clone C7 showed no similarity to existing Myb

Figure 6.6. The predicted amino acid sequence for AaMYB1. An acidic domain, potential phosphorylation sites and a putative nuclear localisation site (NLS), are shown. The two repeats of the Myb domain are underlined and the end of the first repeat is highlighted in yellow. An acidic domain (in red) that contains two potential activation domains (boxed in black) is located just outside the Myb domain (residues 134 to the end of the protein). Several acidic amino acids (marked with asterisks) are found within this domain. There are 13 serine (S) and three threonine (T) residues (marked with blue and black arrows respectively) located throughout the length of the amino acid sequence. One or more of these sites may serve as targets for kinases as part of a post-translational mechanism for the control of AaMyb1 protein activity in the cell.

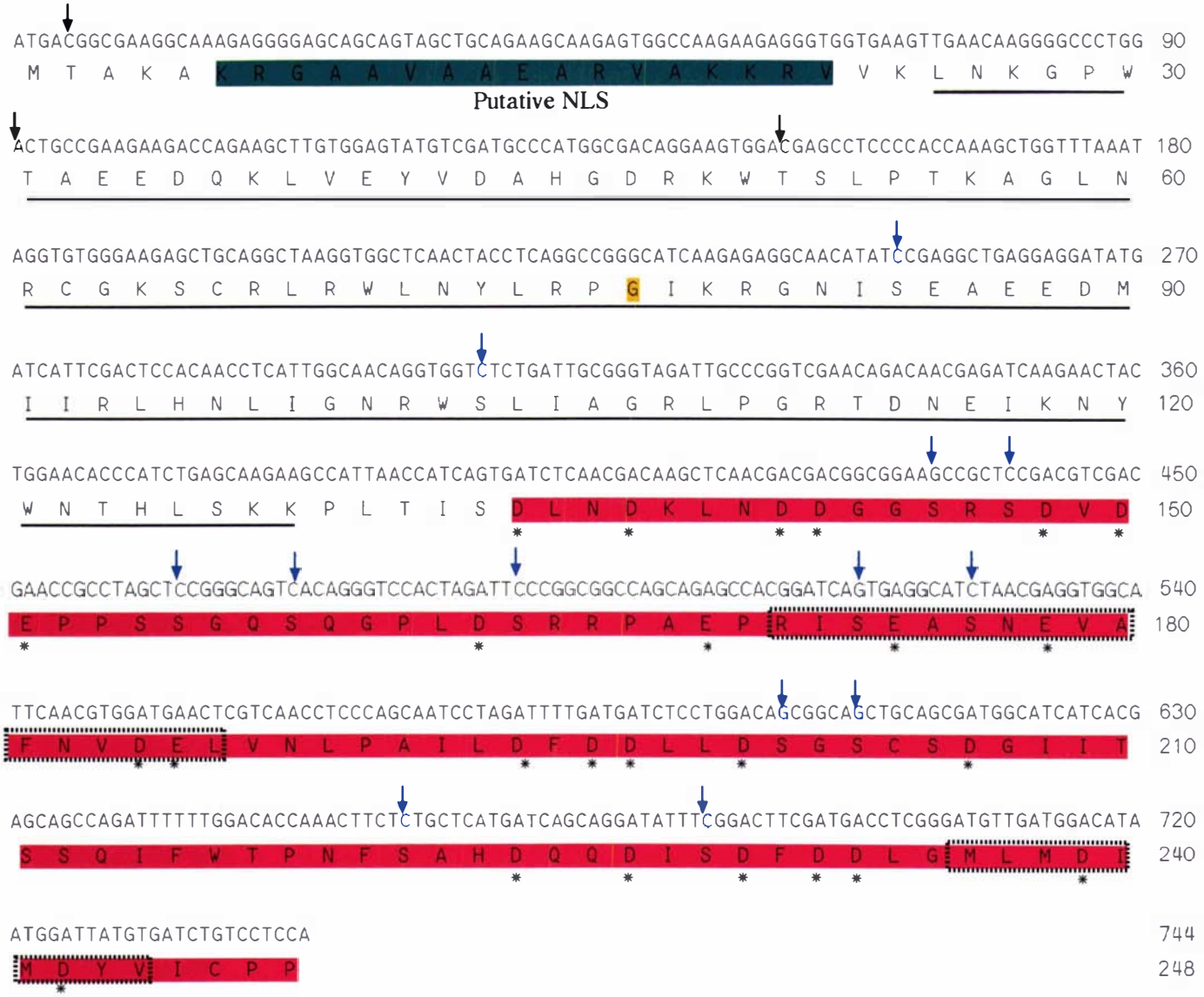

Putative NLS

Figure 6.7. α-helicity of AaMYB1 acidic domain. The helix wheel was constructed using Helix Wheel software from the ExPASy database (www.expasy.ch). Five α-helices are predicted in the acidic domain and the longest (boxed residues 171-186) is presented in this figure a second in Figure 6.8. The helical wheel has an amphipathic configuration, with polar residues (highlighted in green) lining up on one face of the helix and hydrophobic residues (highlighted in orange) lining up on the other.

171-RISEASNDVDFNVDEL-186

Polar Residues

Hydrophobic residue

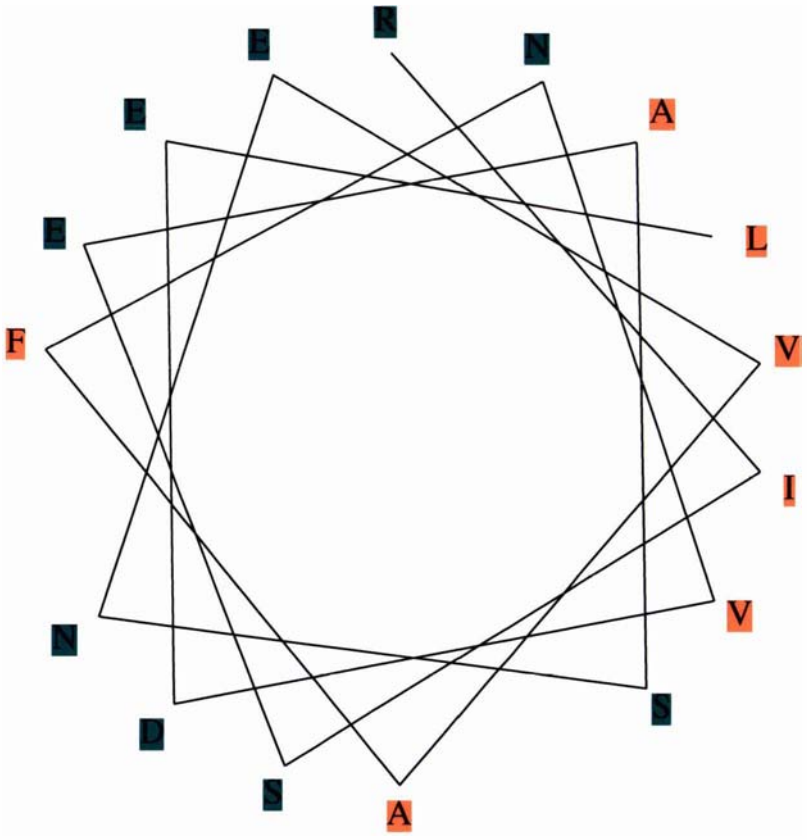
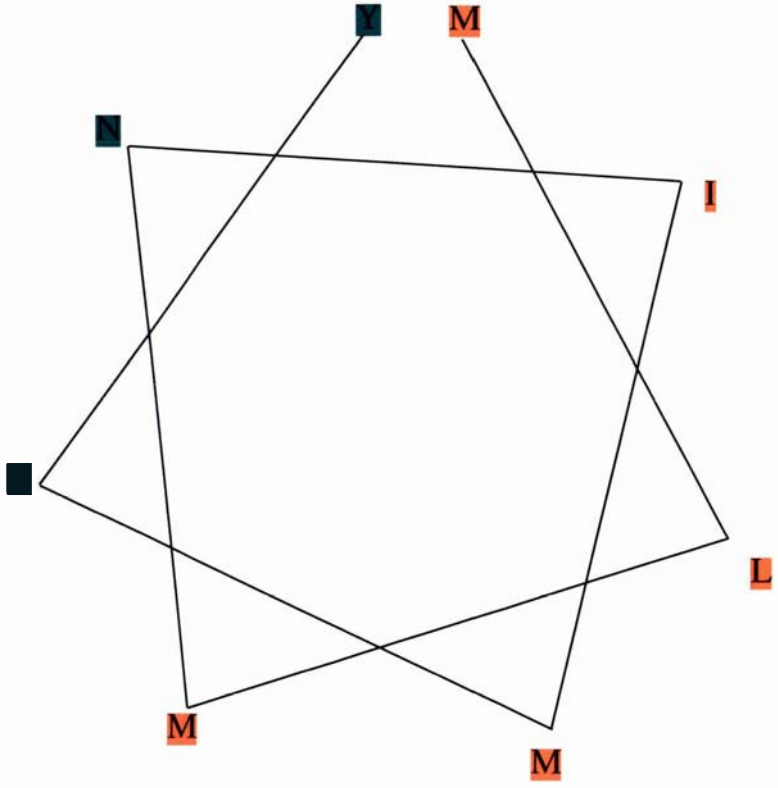


Figure 6.8. The predicted second α-helix of AaMYB1 acidic domain. The helix is formed from residues 236-244, shown in the box. The wheel was constructed using Helix Wheel software from the ExPASy database (www.expasy.ch). The analysis predicts an amphipathic wheel, with polar residues (shaded in green) on one face of the helix and hydrophobic residues (shaded in orange) on the other face.

236-MLMDIMDYV-244

Polar Residues

Hydrophobic residue



proteins in the database when both its DNA and amino acid sequence were used in BLAST searches of the NCBI database.

The nucleotide sequence for the Myb binding domain of *AaMyb1* matched exactly those of clones Myb 20-5 and 20-18 (180 bp). In addition, the 3' end of the full-length cDNA of *AaMyb1* matched exactly that of the 3' end cDNA fragment obtained by RACE PCR in Section 6.3.2.

6.3.3.1 Characteristics of the AaMyb1 cDNA

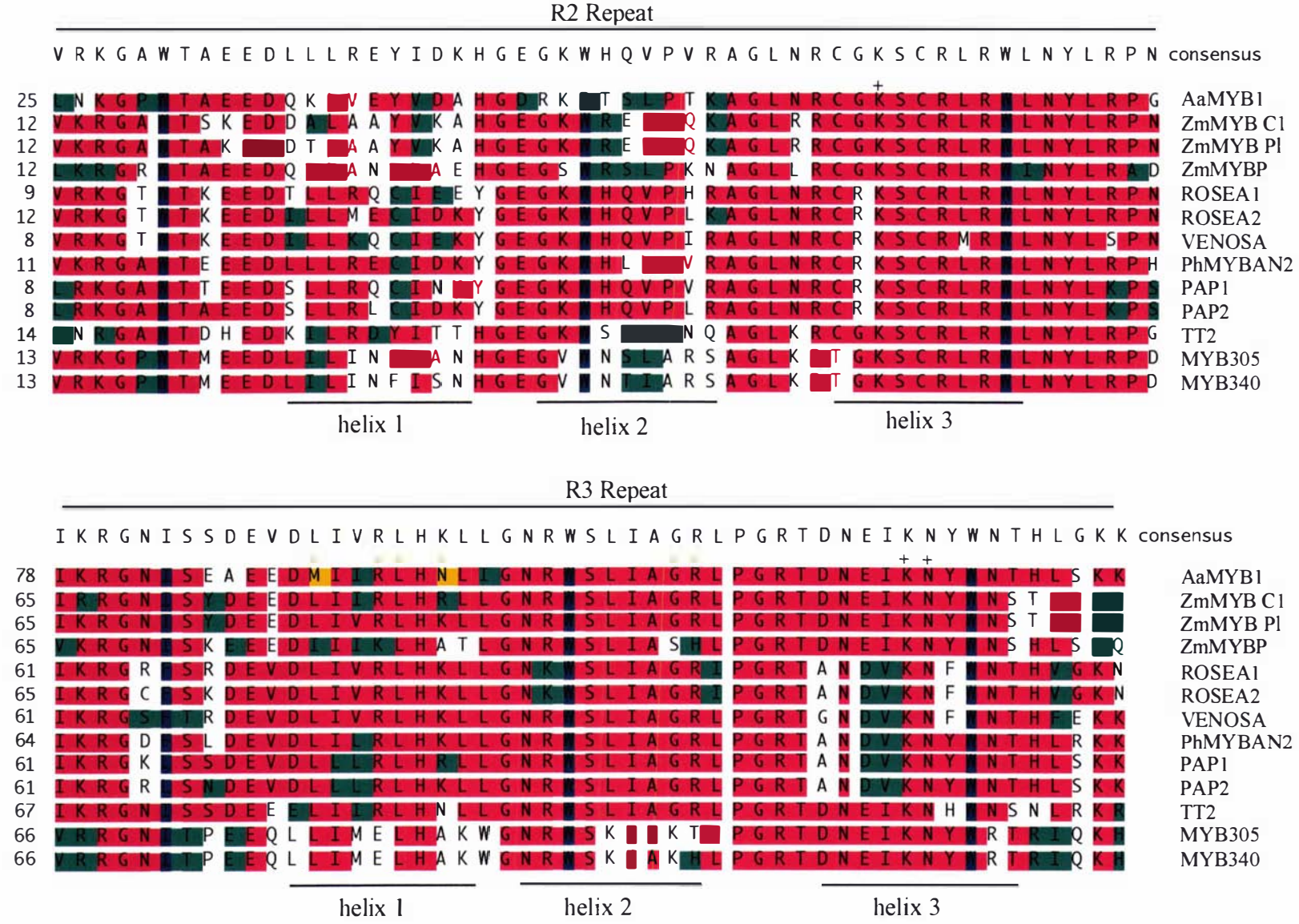
The full-length cDNA clone was 976 bp with the longest open reading frame being 747 bp. There were 63 bp of 5' untranslated sequence, 168 bp of 3' untranslated sequence and a poly $(A)^+$ extension of 16 bp. Two conserved polyadenylation signals (ATAATAT and AATAAAT, Joshi, 1987b) were located downstream of the stop codon at positions 867 and 926, respectively. The nucleotide sequence encoding AaMYB1 is presented in Appendix IX.

6.3.3.2 Characteristics of the AaMYB1 protein

The coding sequence for *AaMyb1* and the putative protein it encoded are shown in Figure 6.6. The AaMYB1 open reading frame encoded a 249 amino acid deduced polypeptide with a predicted mass of 27 kD and an estimated isoelectric point of 5.2. The N-terminus domain was composed of two repeats corresponding to the second (R2) and third (R3) repeats of the three repeat Myb domain found in animal Myb proteins, and was compared to the R2R3 repeats of related Mybs in Figure 6.9. This N-terminal R2R3 Myb domain structure predominates in the plant Myb proteins identified to date (reviewed in Martin and Paz-Ares, 1997; Jin and Martin, 1999). Each repeat consist of three helices.

The highly conserved tryptophan residues present in the binding domain of mammalian c-Myb and all R2/R3 plant Mybs (Klempnauer et al., 1982; Martin and Paz-Ares, 1997) were conserved in AaMYB1 (Figure 6.9). As in most plant Mybs the first tryptophan residue in the R3 repeat was conservatively substituted. In AaMYB1, the tryptophan residue is replaced by isoleucine (Figure 6.9). High % identities were shared between AaMYB1 and the anthocyanin related Mybs when the Myb DNA binding domains were compared. The

Figure 6.9. Sequence alignment of the R2R3 Myb domain of AaMYB1 to anthocyanin related Myb proteins. These proteins are all anthocyanin regulators except for P, TT2, MYB305 and MYB340. P was included because it is a monocot Myb regulating a branch of the flavonoid pathway (Grotewold et al., 1994). Similarly MYB305 and MYB340 were included because of their suggested roles in regulating genes acting early in the phenylpropanoid and flavonoid pathway (Sablowski et al., 1994; Moyano et al., 1996). TT2 was included because its close resemblance to AaMYB1 and because it was able to substitute for C1 in maize kernels (Nesi et al., 2001). The alignment was done using MegAlign in the DNASTAR sequence analysis suite of programmes. Conserved tryptophan residues are coloured blue, including the conserved substitution for the first tryptophan of the R3 repeat that is a common feature of plant Mybs. All other amino acids that form the consensus of each repeat are highlighted in red, while conserved substitutions are highlighted in green. The three putative α-helices in each repeat are noted below the diagram. The six residues that have been found to specify the interaction of C1 with its bHLH partner (Grotewold et al., 2000) and that are conserved in all anthocyanin related Mybs are marked with asterisks. Two of these residues, (highlighted in yellow) are not conserved in AaMYB1. Like AaMYB1, TT2 has an asparagine (N) instead of the conserved arginine (R) or lysine (K) in one of six positions. Residues that are known specific interaction points with target DNA are conserved in all the Myb proteins in the alignment and are marked with a plus (+) sign. The GenBank accession numbers (or references where accession numbers are not available) are: C1 (P10290), PL (AAB67720), P (AAC49394), ROSEA1, ROSEA2 and VENOSA (Schwinn, 1999), AN2 (AAF66727), PAP1 (AF325123), PAP2 (AF325124), TT2 (AJ299452), MYB305 and MYB340 (Jackson et al., 1991).

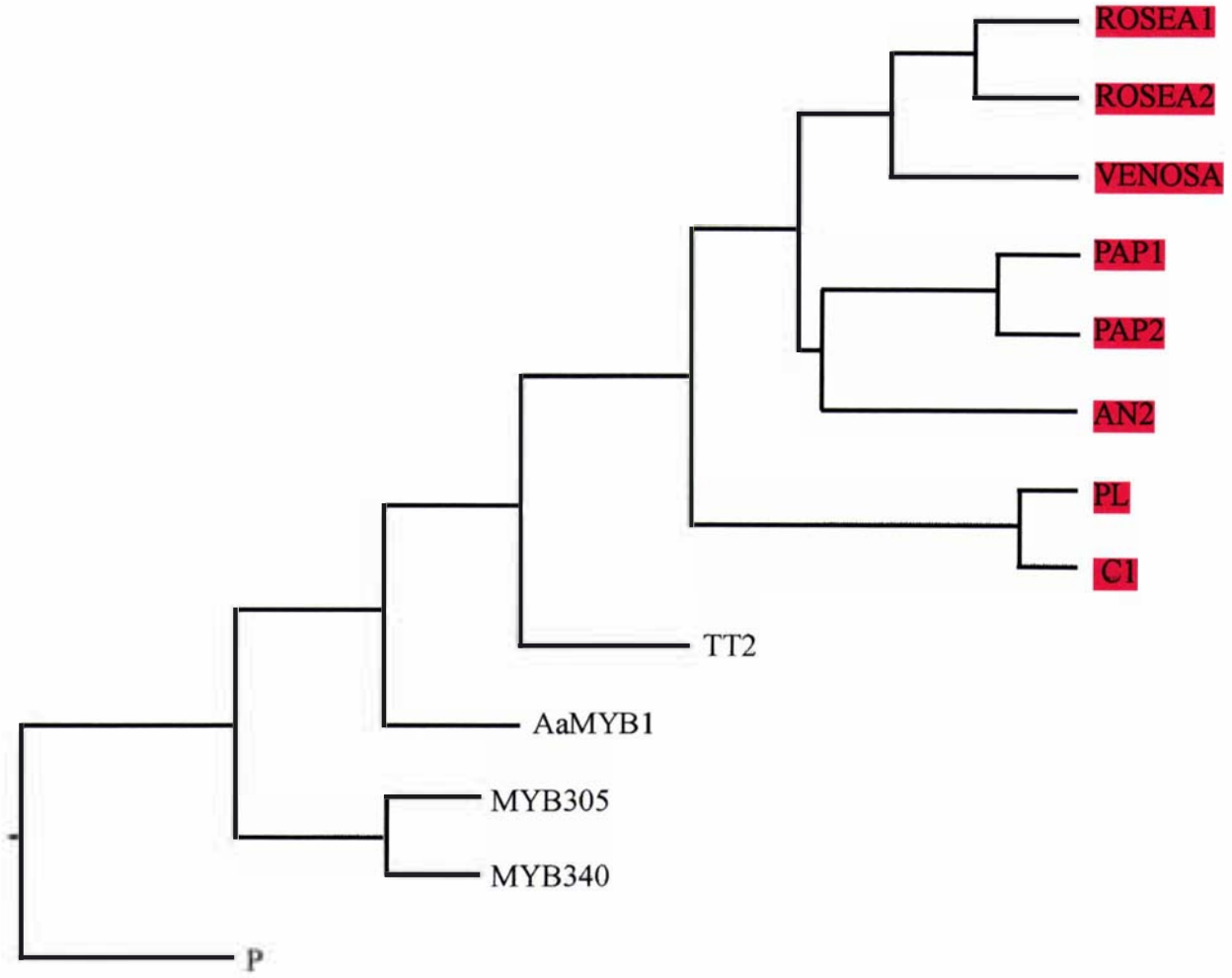

R2 Repeat
V R K G A W T A E E D L L L R E Y I D K H G E G K W H Q V P V R A G L N R C G K S C R L R W L N Y L R P N consensus
25 AaMYB1
12 ZmMYB C1
12 ZmMYB Pl
12 ZmMYBP
9 ROSEA1
12 ROSEA2
8 VENOSA
11 PhMYBAN2
8 PAP1
8 PAP2
14 TT2
13 MYB305
13 MYB340
helix 1
helix 2
helix 3
R3 Repeat
I K R G N I S S D E V D L I V R L H K L L G N R W S L I A G R L P G R T D N E I K N Y W N T H L G K K consensus
78 AaMYB1
65 ZmMYB C1
65 ZmMYB Pl
65 ZmMYBP
61 ROSEA1
65 ROSEA2
61 VENOSA
64 PhMYBAN2
61 PAP1
61 PAP2
67 TT2
66 MYB305
66 MYB340
helix 1
helix 2
helix 3

R2 repeat for AaMYB1 showed 64% and 66% identity to maize C1 and PL, respectively. PL is the C1 homologue controlling anthocyanin biosynthesis in most of the plant body (Cone et al., 1993). Even higher values are observed when the R3 repeats are compared, with a 78% identity for maize C1 and PL. However the highest % identity for R2 and R3 of AaMYB1 was with arabidopsis TT2. It shared 68% and 80% identity, respectively, to the corresponding repeat in TT2. TT2 controls proanthocyanidin accumulation in arabidopsis seeds and while it was able to substitute for C1 in maize kernels, the mutant (tt2) does not affect anthocyanin production in seedlings and vegetative tissue (Nesi et al., 2001). The AaMYB1 R3 repeat had only four of the six conserved residues (Figure 6.9) that specify interaction of maize C1 with its bHLH partner (Grotewold et al., 2000). These residues have been found to be conserved in all anthocyanin related Mybs. In contrast TT2 only has five of these residues (Nesi et al., 2001).

While there are high levels of conservation between the Myb domain of AaMYB1, TT2 and maize anthocyanin related Mybs, the % identity over the entire length of the protein, averaged at 34% (data not shown). This explains the distinct and separate evolutionary grouping of AaMYB1 in the phylogenetic tree shown in Figure 6.10.

The ExPASy Proteomics database (www.expasy.ch) was used to evaluate the amino acid sequence of AaMYB1 for other characteristics. Regarding possible post-translational modification of the AaMYB1 protein, the NetPhos program (Blom et al., 1999) identified several serine and threorine residues as potential phosphorylation sites (Figure 6.6). The PSORT programme (Nakai and Kanehisa, 1992), that predicts protein sorting signals and localization sites for eukaryotic proteins, identified the KRGAAVAAEARVAKKRV amino acid sequence in the N-terminal domain as a putative nuclear localisation signal (Figure 6.6).

Figure 6.10. Phylogenetic relationship of AaMYB1 to the anthocyanin Myb family. The rational for including P, TT2, MYB305 and MYB340 was provided in the legend of Figure 6.9. The phylogenetic tree was built using full-length amino acid sequences for the Myb proteins described in Figure 6.9 and is based on an alignment performed with the MegAlign programme (gap penalty 10, gap length penalty 10) from the DNASTAR sequence analysis suite of programmes. Amino acid sequences and GenBank accession numbers (or references where accession numbers are not available) are: ROSEA1, ROSEA2 and Venosa (Schwinn, 1999), AN2 (AAF66727), PAP1 (AF325123), PAP2 (AF325124), TT2 (AJ299452), PL (AAB67720), C1 (P10290), MYB305, and MYB340 (Jackson et al., 1991) and P (AAC49394).


ROSEA1
ROSEA2
VENOSA
PAP1
PAP2
AN2
PL
C1
TT2
AaMYB1
MYB305
MYB340
P

6.4 DISCUSSION

Using degenerate primers for the conserved region of the Myb domain was an effective strategy for isolating anthurium *Myb* cDNAs. This approach to cloning Myb cDNAs was pioneered with arabidopsis and showed that there are more than 80 R2/R3-Myb regulatory genes in its genome (Romero et al., 1998). The isolation of 20 clones from anthurium spathe RNA tissue alone, from which six distinct genes have been identified, suggests that the anthurium genome may also contain a large Myb gene family. Such large Myb families have been suggested for other plant species besides arabidopsis, such as petunia and tomato (Avila et al., 1993; Lin et al., 1996). The classification of the anthurium Myb protein fragments into six distinct groups is consistent with the phylogenetic tree in Figure 6.5, where each of the Myb fragments group with the Myb factor to which it showed the highest % sequence identity.

Although the research target was anthocyanin biosynthesis related Myb genes, the other Myb genes identified are of interest. Group A shows high % identity to arabidopsis AtMYB4 in its DNA binding domain. AtMYB4, the first plant R2R3 Myb protein identified as a repressor, regulates the accumulation of UV-protectant sinapoyl esters by repressing the activity of C4H (Jin et al., 2000). It is tempting to infer a similar mechanism in anthurium spathe tissue involving a candidate of Group A, as UV-protection is critical for this obligate shade species. Group B is another interesting clone, as it showed high % identity to antirrhinum MIXTA. MIXTA affects pigmentation, not by targeting anthocyanin structural genes but by controlling the formation of conical petal epidermal cells (Noda et al., 1994). However, despite the interest in the other genes, time did not allow for further work to be performed on these clones. Groups C and F were selected as those being most likely to encode an anthurium anthocyanin regulator.

Three rounds of nested PCR, using first the degenerate primer (N6), were necessary to successfully clone the 3' end of Group C and F *Myb* cDNA fragments. A full-length sequence for Group F was obtained from the library screen, as evidenced by the presence of a putative ATG start codon along with 5' untranslated sequence data from the *AaMyb1* cDNA clone. The putative polyadenylation signals identified by Joshi (1987b) were clearly

seen in the 3' end of the Group F clone and a clearly defined poly $(A)^+$ tail was present suggesting that the complete 3' end of the clone had also been secured. However, the Group C fragment obtained from RACE PCR reactions appeared to be truncated. This can happen if the dT_{17} adaptor-primer binds to an A-rich region in the template coding sequence upstream from the poly $(A)^+$ tail (Frohman et al., 1988). It is the ability to identify conserved polyadenylation sites in the 3' sequence data that gives the confidence that the reaction proceeded from the poly $(A)^+$ tail. However, the short poly $(A)^+$ tail, the lack of clear polyadenylation signals and the relatively short length of the Group C clone, suggest that it was generated by the dT_{17} adaptor-primer binding to an A-rich region upstream from the poly $(A)^+$ tail.

The sequence analysis of the predicted protein encoded by *AaMyb1* gave strong evidence for a role as a DNA binding protein. All the key structural features that define the Myb domain are present, particularly the high levels of conservation in the third α-helix (Figure 6.9) of each repeat, which is the predicted recognition helix (Ogata et al., 1994). The occurrence of several serine and threonine residues in the predicted amino acid sequence of AaMYB1, suggest that post-translational mechanism of phosphorylation/dephosphorylation may regulate AaMYB1 expression as may occur with other Myb proteins. In antirrhinum, dephosphorylation of MYB340 switches its DNA binding affinity from low to high (Moyano et al., 1996). Similarly, phosphorylation of a serine residue in the N-terminus region of animal c-Myb by a protein kinase, inhibits sequence specific binding of c-Myb to its target DNA (Lüscher et al., 1990), and dephosphorylation restored its DNA binding ability.

While an analysis of the N-terminal region of AaMYB1 strongly supports its role as DNA binding protein, it is the C-terminus that provides clues to its role as a transcriptional activator. AaMYB1 contains an acidic domain in its C-terminus. The acidic domains of transcription factors such as C1, animal c-Myb and AtMYB2, contain transactivation domains and are required for transcriptional activation (Sakura et al., 1989; Paz-Ares et al., 1990; Goff et al., 1991; Urao et al., 1996). This suggests that AaMYB1 also has a transactivation domain and acts as a transcriptional activator. It has been argued that

activation domains require for function an amphipathic α-helix with polar residues on the surface and hydrophobic residues within (Giniger and Ptashne, 1987; Ptashne, 1988; Franken et al., 1994). Helical wheel analysis has confirmed that such structures could form within the activation domain of AaMYB1, and similar structures are predicted in C1 (Paz-Ares et al., 1987) as well as in ROSEA1, ROSEA2 and VENOSA, which are antirrhinum anthocyanin regulators (Schwinn, 1999).

Interestingly, however, studies with maize C1, involving extensive mutational analysis of the transcriptional activation domain to introduce helix incompatible residues, showed that an amphipathic α-helix was not required for activation (Sainz et al., 1997b). Rather, activation may be dependent on specific amino acids in the activation domain of maize C1. The importance of two specific amino acid residues (aspartate 262 and especially leucine 253) were highlighted in experiments that showed significant reduction in activation when substitutions were made for these residues (Sainz et al., 1997b). As to whether these residues are essential for all anthocyanin Myb transcription activators is not certain, especially as the more important leucine 253 in maize is not conserved in the antirrhinum anthocyanin Myb transcription factor ROSEA1.

Until the putative activation domain of AaMYB1 is functionally defined, the possibility that AaMYB1 may be a repressor has to be considered. However, none of the motifs identified in the C-terminus of repressor Myb proteins (Kranz et al., 1998; Jin et al., 2000; Aharoni et al., 2001) are found in the C-terminus of AaMYB1. The absence of motifs that are conserved in repressor Mybs, suggest that AaMYB1 does not belong to this class of Myb proteins.

It is important to note that, while an activation domain is predicted in the C-terminus of AaMYB1, it can still function as a repressor in situ as was shown for b-Myb that functions both as an activator and a repressor though it possesses an activation domain (Foos et al., 1992). In this way, AaMYB1 may be able to bind to but not activate its target genes, thereby acting as competitor with the other Myb factor. A similar mechanism has been

suggested for AmMYB308 in its role as a regulator of lignin biosynthesis and phenolic acid metabolism in antirrhinum (Tamagnone et al., 1998).

None of the motifs found in the C-terminus of Myb proteins involved in phenylpropanoid biosynthesis (subgroups 5-7, Kranz et al., 1998) or in all the remaining subgroups, were found in AaMYB1. However, a similar situation was reported for TT2, yet it was shown to play a role in regulating proanthocyanidin accumulation in developing arabidopsis seeds (Nesi et al., 2001). Therefore, the absence of conserved motifs in the C-terminus of AaMYB1 compared to Mybs involved in phenylpropanoid biosynthesis does not preclude AaMYB1 as an activator in regulating some aspect of this metabolic pathway.

Whether AaMYB1 partners with a bHLH factor would be an interesting question to investigate further, given that two of the six residues important for specifying maize C1 interaction with its bHLH partner (Grotewold et al., 2000) are absent in AaMYB1. It is unlikely that the two residues in AaMYB1 are sequence errors as they were also present in both the PCR clones and the clone obtained from the library. However, not all Mybs involved in regulating flavonoid biosynthesis require a bHLH partner. Antirrhinum MYB305 and MYB340 were both able to activate the first gene of the phenylpropanoid pathway along with *CHI* and *F3H* of the flavonoid pathway in yeast and plant protoplast when supplied alone (Moyano et al., 1996). In addition, the *P* gene product of maize (that regulates phlobaphene biosynthesis) can activate a subset of the anthocyanin biosynthetic genes, independent of R, in maize cells (Grotewold et al., 1994). Interestingly, all three Mybs (MYB305, MYB340 and P) like AaMYB1, have only 4 of the six conserved residues (Figure 6.9) Therefore, AaMYB1 may not require a bHLH transcription factor to function in regulating a branch of the phenylpropanoid pathway in anthurium spathe tissue.

As an alternative, AaMYB1 may have a role in regulating flavonoid biosynthesis that requires interaction with a class of protein other than bHLH transcription factors. A bZIP protein was suggested as the binding partner of MYB305 (Sablowski et al., 1994). Therefore, if a similar situation occurs with AaMYB1 as a potential regulator of flavones, it may do so by interacting with a bZIP protein and so display a different interaction dynamic

than C1 with its bHLH partner. Further to this, it is not known if the mechanism specifying C1 interaction with its bHLH partner R applies for all other anthocyanin Mybs. The inability of C1 to complement *rosea1* in particle bombardment experiments (Dr. Kathy Schwinn, Crop & Food Research, New Zealand, personal communication) suggest that maize C1 cannot work with endogenous antirrhinum anthocyanin bHLH factors and that the mechanism of C1 interaction with R may not be conserved in all other plant species. This idea is strengthened by that fact that neither antirrhinum DELILA or petunia JAF13 were able to complement maize *r* mutants, although JAF13 could substitute for *r* in some functional assays (Mooney et al., 1995; Quattrocchio et al., 1998). In that regard, that two of the six residues are not conserved does not provide sufficient basis to exclude AaMYB1 from a role in regulating phenylpropanoid biosynthesis in anthurium spathe.

Chapter 7

Functional analysis of AaMYB1 and DFR promoter studies

7.1 INTRODUCTION

In Chapter 6, several structural features of the AaMYB1 clone were described, that suggest the encoded product acts within the nucleus as a transcriptional activator. However, further work was necessary to confirm whether it is an anthocyanin regulator in anthurium spathe tissue. Transient expression assays have been used to show that Myb proteins function in heterologous systems. By this method, AN2 was shown to substitute for maize C1 and activate the entire cascade of structural genes for pigment production in maize aleurone tissue (Quattrocchio et al., 1998). Also AN2 and JAF13, that are required for DFR expression in petunia, were shown to work effectively with Lc and C1 respectively (Quattrocchio et al., 1993, 1998). Similar experiments also confirmed a role for ROSEA1 and ROSEA2 as activators of anthocyanin biosynthetic genes in antirrhinum (Schwinn, 1999). Therefore, transient expression assays, using particle bombardment, were considered a valid approach in assessing the ability of AaMYB1 to induce anthocyanin biosynthesis in heterologous tissue. *rosea1* petal tissue was used, from the rosea^{dorsea} line, which carries a mutated *rosea1* gene resulting in pigmentation predominating in the outer epidermis of the upper lobes and tubes.

While this approach has been used with reasonable success for defining function of anthocyanin Myb transcription factors, it is difficult to make inferences from a negative result. For example, maize C1, although an established anthocyanin regulator, does not complement the *rosea1* mutation unless Lc is supplied (Kathy Schwinn, personal communication). This suggests that not all the anthocyanin regulators are interchangeable and that specific regulatory features may have evolved in particular species. Therefore, similar experiments were also conducted in an anthurium mutant (Acropolis). Results discussed in Chapter 4 suggest that Acropolis is a regulatory mutant. However, it is not known if this mutation is at a genetic locus encoding a Myb or bHLH transcription factor. Also, an anthurium *DFR* promoter fragment was cloned and attached to the GUS reporter gene for promoter-reporter studies in wild type and *rosea1* antirrhinum petal tissue.

It is important to bear in mind that the ability to recognise and bind to a given promoter does not always mean that this is the native target of the Myb or that it regulates the process

for which the structural gene is involved in (Sablowski et al., 1994). Additionally, these experiments do not distinguish whether the regulators are activating structural genes directly or are part of a more complex regulatory hierarchy in which they regulate the structural genes in an indirect manner.

7.2 MATERIAL AND METHODS

7.2.1 Northern analysis of AaMyb1 expression in anthurium spathe

The details for northern blotting are presented in Section 2.14. To examine the expression pattern for *AaMyb1*, 5 µg of poly $(A)^+$, prepared from RNA extracted from four stages of Altar spathe tissue (Stages 2-5), was electrophoresed in a denaturing gel then blotted onto Hybond N^+ membrane. The blot was prehybridised and probed with the radiolabeled C-terminal fragment of *AaMyb1* (as an *Eco* RI insert cut from the pGEM-T Easy vector) at 65 °C overnight. Membranes were washed, once with 2 x SSC/0.1 % (w/v) SDS and then with 0.1 x SSC/0.1% (w/v) SDS solution. Both washes were for 15-20 min.

7.2.2 Construction of an AaMyb1 expression vector

The anthurium *AaMyb1* cDNA was cloned in sense and antisense orientation into pART7 to form pPN122 and pPN123, respectively. pART7 is a vector derived from pGEM9Zf containing the 35S promoter of CaMV isolate Cabb B-JI, a multiple cloning site, and the transcriptional termination region of the octopine synthase gene (Gleave, 1992). Restriction sites were engineered onto the *AaMyb1* clone using Platinum *Pfx* polymerase, to directionally clone it into pART7 that was linearised by an *Eco* RI/*Cla* I digest. Primers MYBP7 and MYBP8 (see primer list in Appendix I) were used to add an *Eco* RI site and a *Cla* I site to the N-terminal and C-terminal encoding ends of the full-length clone. The *Eco* RI site was added upstream of the 5' untranslated region and the *Cla* I site was added downstream from the poly $(A)^+$ tail. PCR reactions were set up as described in Section 2.9.2 using a Techne Genius Thermocycler. The cycling parameters used were; 1 cycle at 95 °C/3 min; 35 cycles at 95 °C/30 s, 60 °C/1 min, 68 °C/1 min and 1 cycle at 68 °C/4 min. The entire PCR reaction was electrophoresed on a 1% (w/v) agarose gel and the desired

product was isolated and quantified by comparing band intensities with a DNA mass ladder as described in Sections 2.8.3. pART7 and the *AaMyb1* purified PCR products were digested with *Eco* RI/*Cla* I at 37 °C (the reaction with the PCR fragment was left overnight) and the DNA separated by gel electrophoresis. The linearised vector and the *Eco* RI/*Cla* I *AaMyb1* fragment were isolated from the agarose gel and their concentrations determined by agarose gel electrophoresis. The digested *AaMyb1* (1-3 μL) fragment was then ligated into the *Eco* RI/*Cla* I cut pART7 (1 μL) using 1 μL T4 DNA ligase and 1 μL 10 x ligase buffer. The ligation reaction was incubated overnight at 4 °C and was subsequently used to transform *E.coli* as described in Section 2.13. An antisense construct (pPN123) was made in pART7 by ligating pAa5 (*AaMyb1* in pBluescript SK+) cut with *Xba* I/ *Xho* I into pART7 cut with the same pair of restriction enzymes.

Colony hybridisation was used to screen for cell lines containing pPN122 and pPN123. Ampicillin resistant colonies were streaked (in a replicated numbered pattern) onto L-agar plates with ampicillin selection and incubated overnight at 37 °C. Colonies were then lifted onto Hybond N^+ membranes (marked to define orientation). The membrane with the colonies was then placed on 3 MM filter paper soaked with 10% (v/v) SDS to lyse the bacterial cells. After 5 min the membranes were removed and placed in succession on denaturing solution and neutralisation solution (made up as described in Section 2.5.4), each for 5 min. Membranes were then washed in 2 x SSC to remove cell debris and the DNA was then fixed to the membrane using a UV cross-linker as described in Section 2.5.4. The filters were hybridised with a radiolabeled *AaMyb1* probe at 65 °C overnight in Church and Gilbert solution. Subsequently, membranes were washed first with 3 x SSC/0.1% (w/v) SDS and then with 0.1 x SSC/0.1% (w/v) SDS, both washes at 65 °C for 15-20 min. Plasmids from colonies hybridising to *AaMyb1* were checked by restriction enzyme analysis to determine which contained the *AaMyb1* fragment. Sequence information was used to confirm the details of the plasmid construct.

7.3 PARTICLE BOMBARDMENT EXPERIMENTS

These experiments were performed to test the functionality of *AaMyb1* as well as the ability of antirrhinum anthocyanin Myb regulators to activate reporter gene expression from the promoter of an anthurium *DFR* gene.

Particle bombardment experiments were performed with a modified helium particle inflow gun based on Finer et al. (1992) but with a high-speed direct current solenoid valve that allows more accurate opening times. Anthurium spathe tissue (generally from Stage 6 flowers) was used along with lobe tissue from Stage 7 for antirrhinum (refer to Martin et al., 1991 for definition of stages). The spathe or lobe tissue was first removed from the rest of the flower, surface sterilised for 10-15 min in 10% (v/v) bleach containing 1-2 drops of TWEEN (100 mL^{-1}), and then rinsed thoroughly in sterile deionised water. Gold, 1 μm in diameter, was used as the standard microparticle DNA carrier. The gold (50 mg) was prepared by sterilisation with ethanol then resuspension in 500 μL of sterile water. This was divided into 50 μL aliquots with a final concentration of 0.1 mg μL^{-1}. The particle suspension consisted of 20 μg plasmid DNA (in 10 μL) precipitated onto the gold particles (5 mg) using 50 μL of 2.5 M $CaCl_2$ (filter sterilised) and 20 μL of 0.1 M spermidine (prepared using sterile water). The gold suspension was set to vortex at a constant speed during which the DNA was added, then $CaCl_2$ and spermidine were added simultaneously. The suspension was vortexed for a further 3 min and then briefly centrifuged. Following this, 90 μL of the supernatant was removed and the suspension vortexed again to prevent clumping.

Once the particle suspension was ready, the tissue was placed on #2 medium [1/2 x MS macro salts (Murashige and Skoog, 1962), 1 x MS micro salts, 1 x MS iron, 1 x LS vitamins (Linsmaier and Skoog, 1965), 3% (w/v) sucrose, 7.5% (w/v) agar] and placed inside the chamber of the inflow gun at 120-160 mm from the nozzle. An aliquot of 5 μL of the particle suspension was used for each bombardment and microparticles were accelerated directly in a helium stream (300 kPa in a partial vacuum of -95 kPa). The standard solenoid opening time was 30 milliseconds. Each sample was bombarded three

times and then incubated on #2 medium at 20 °C under a 16/8 h light/dark photoperiod, with 35 mmol m^{-2} sec^{-1} cool white fluorescent light. Tissue was examined the following day for GFP expression. Histochemical staining for GUS expression (using the RT99-GUS vector) required a further 24 h incubation in GUS assay buffer consisting of 50 mM phosphate buffer (pH 7.0), and 0.35 mg mL^{-1} X-Gluc substrate. Incubations were performed in the dark at 37 °C overnight.

7.4 CLONING AND CHARACTERISATION OF THE PROMOTER REGION FOR *AaDFR1*

The promoter for the anthurium *DFR* gene was cloned using the Universal Genome Walker kit (Clontech, USA). The kit is for a PCR based method that allows the cloning of genomic DNA that is adjacent to a known sequence of DNA. The first step was the isolation of high quality genomic DNA (described in Section 7.4.1). This was followed by the construction of pools of adaptor-ligated genomic DNA fragments referred to as 'libraries' (Section 7.4.2). The libraries were made from 1 µg of genomic DNA completely digested with the selected restriction enzymes that generate blunt ended fragments. Each pool of digested genomic DNA was then ligated to the Genome Walker adaptors. The libraries became the stock of template for the third stage of the protocol, consisting of PCR based genome walking that used nested general primers for the adapters and nested primers for the 5' end of the anthurium *DFR* cDNA (Section 7.4.3). The promoter fragment was then cloned into pGEM-T Easy vector producing plasmid pANA15 and sequenced. It was also ligated into the expression vector pUCBM21 (producing pPN121) as described in Section 7.5.

7.4.1 Extraction of anthurium genomic DNA

DNA was extracted from anthurium leaf tissue using the Nucleon Phytopure system (Amersham Pharmacia Biosciences, England), which was developed specifically to eliminate contaminating polysaccharides from DNA extractions without the use of phenol or CTAB. This is accomplished with the use of the Nucleon Phytopure resin, which serves two purposes. Once the cell walls are broken and the cells are lysed with Reagent 1, containing potassium SDS, the resin is added. This removes polysaccharides from the

sample as boric acid esters, formed by covalent linkages between the free boric acid groups in the resin and 1,2 dihydroxy compounds in the sample. Secondly, the resin partitions off the homogenate into an upper DNA-containing aqueous phase and a lower organic phase, thereby facilitating easy recovery of high quality DNA.

Fresh, juvenile leaf tissue (1 g) from Altar was harvested and crushed to a fine powder in liquid N_2. An aliquot of 4.6 mL of Reagent 1 and 250 µl of 2β ME was added to a 15 mL Nalgene tube. The crushed tissue was then added to the solution in the tube and vigorously shaken. Reagent 1 has potassium SDS, which is known to form complexes with proteins and carbohydrates. Clumps of tissue were fragmented with a sterile glass rod ensuring a homogenous mix. During this, 20 µL of a 10 mg mL^{-1} DNase free RNase stock was added and thoroughly mixed with the extract. This was followed by a 1 h incubation of the tube in a 37 °C water bath with intermittent shaking. Reagent 2 (1.5 mL) was then added, mixed and the tube incubated at 65 °C for 10 min, again with intermittent shaking. The sample was then placed on ice for 20 min after which 2 mL of chloroform, that had been stored at -20 °C, was added. The cold chloroform is thought to be more effective at removing complexed proteins and polysaccharides. To further assist in the removal of these contaminants, 200 µL of the Phytopure resin suspension was added to the samples and the tubes inverted several times to mix the contents. The tube was shaken on a tilt shaker for 10 min at room temperature and then centrifuged at 1300 g in a Jouan bench top centrifuge for 10 min. Once completed, the upper DNA containing aqueous phase was removed and transferred into a fresh tube without disturbing the resin suspension layer. To precipitate the DNA, an equal volume of ice-cold isopropanol was added and the tubes gently inverted until the DNA precipitated. At this point DNA was hooked out of the solution using a heat-sealed pipette and transferred directly into 500 µL of sterile water. The genomic DNA was then left overnight at 4 °C to resuspend. The isolated DNA was quantified using the Hoeschts Fluorometer as described in Section 2.8.2, except that the standard used was Calf Thymus DNA.

7.4.2 Construction of Genome Walker libraries

The construction of the libraries required very clean, high molecular weight genomic DNA as the starting material. The genomic DNA, isolated from anthurium leaf by the Nucleon Phytopure resin system (Section 7.4.1), was tested to determine if it was of sufficiently high quality. A 1 μL aliquot of the genomic DNA (0.1 μg μL^{-1}) along with 1 μL of control genomic DNA (0.1 μg μL^{-1}) was electrophoresed on a 0.5% (w/v) agarose/EtBr gel with 1 x TBE along with *Hin*d III fragments as a DNA ladder (Invitrogen, New Zealand). The DNA was found to be larger then 50 kb with minimal smearing (data not shown). This was an ideal size and so was used as the source of genomic DNA for the library. In addition, human genomic DNA supplied with the kit was used as a positive control for making the library.

Five blunt end digestions were set up, four with anthurium genomic DNA using *Eco* RV, *Stu* I, *Pvu* II and *Dra* I and one with human genomic DNA using *Pvu* II. The tubes were labeled DL1 to DL4 and positive control. Digestion contained 2.5 μg DNA, 1/10 volume 10 x buffer, 8 μL restriction enzyme, and sterile water, bringing the final volume of 100 μL.

All tubes were inverted (never vortexed as this could shear the DNA) to gently mix the contents and then incubated at 37 °C for 2 h. After this, the reactions were gently mixed again and returned to 37 °C to incubate for a further 16-18 h. To test that the digestion of each sample was complete, an aliquot of 5 μL was taken from each reaction and run on a 0.5% (w/v) agarose/EtBr gel. Once the digestion was completed, the DNA was purified.

To each of the samples, an equal volume (95 μL) of phenol:chloroform:isoamyl alcohol (24:24:1 v/v) was added and mixed by gently inverting the tubes. Samples were spun briefly to separate the aqueous and organic phases. The former were transferred to sterile 1.5 mL microcentrifuge tubes and an equal volume of chloroform was then added and mixed by inverting the tubes. Again a brief spin was performed to separate the lower organic and upper aqueous phases. The upper phases were transferred to fresh tubes and the

DNA precipitated with 2 volumes of ice-cold 95% (v/v) ethanol, 1/10 volume of 3 M NaOAc (pH 4.5) and 20 μg of glycogen (20 mg mL^{-1}).
Tubes were inverted and the DNA pelleted by centrifugation at 15,000 rpm for 10 min in a microcentrifuge. Once completed, the supernatants were removed and the pellets washed in 100 μL ice cold 80% (v/v) ethanol. Tubes were centrifuged for a further five min at 15,000 rpm, the supernatants removed and the pellets allowed to air dry. An aliquot of 20 μL of 0.1 x TE (pH 7.5) was used to resuspend the pellets. From each tube a 1 μL aliquot was removed and electrophoresed on a 0.5% (w/v) agarose/EtBr gel to determine the approximate quantity of DNA recovered from the purification.

For each of the five samples, ligation reactions were set up in a sterile 0.5 mL tube using 4 μL of digested purified DNA, 1.9 μL Genome Walker Adaptor (25 μM), 1.6 μL of 10 x ligation buffer and 0.5 μL of T4 DNA ligase (6 U μL^{-1}).

Samples were incubated at 16 °C overnight in a PCR thermal cycler, after which incubating the reaction at 70 °C for 5 min inactivated the ligase. To each tube was added 70 μL 0.1 x TE (pH 7.4) and this was the 'library stock.'

7.4.3 PCR based genome walking

Two gene specific primers were designed to the 5' end of the anthurium *DFR* cDNA clone obtained from the library screen, as discussed in Chapter 3. The two primers (DFRGSP1 and DFRGSP2 (listed in Appendix I) were designed according to the specifications in the Clontech Genome Walker protocol, and were 27 nucleotides in length with a GC content of 55% and 59%, respectively. These primers were used with AP1 and AP2 respectively for

The PCR protocol consisted of eight primary, and secondary PCR amplifications: four libraries from anthurium leaf tissue, two positive controls (positive control human Genome Walker library and one positive control library constructed from control human Genomic DNA alongside the anthurium tissue library) and two negative controls (no templates). For both positive controls the gene specific primers PCP1 and PCP2 supplied with the kit, were

used. For the primary PCR, DFRGSP1 and AP1 were used as the primer combination, while DFRGSP2 and AP2 were used for the secondary PCR reactions (AP1 and AP2 were supplied with kit and were the upstream primer (AP1) and the nested primer (AP2) to the Genome adaptor sequence refer to Table 7.1). An aliquot of 1 μL of each library was used and for the primary PCR and 1μL of a 50 x dilution of each primary PCR product was used for the secondary PCR reactions. The Clontech Advantage Genomic Polymerase Mix was used to catalyse the PCR amplification reactions. This polymerase mix contains a combination of two thermostable DNA polymerases used to increase the accuracy of the PCR. The 'Tth' polymerase is the primary polymerase that carried out most of the extension while a secondary polymerase provides the critical 3'-5' exonuclease proof reading function. The Advantage Genomic Polymerase Mix also contains the 'Tth'Start Antibody that binds to and inactivates the 'Tth' DNA polymerase at lower temperatures and thus eliminates DNA synthesis from non-specifically bound primers while reactions are bring assembled. In so doing the antibody provides a convenient and automatic form of hot-start PCR (Kellog et al., 1994). The layout for the reactions suggested in the Clontech Universal Genome Walker Kit was used and is shown in Table 7.1.

The primary PCR reaction was comprised of 37.8 μL of sterile water, 5 μL of 10 x 'Tth' PCR reaction buffer, 1 μL dNTP mix (10 mM), 2.2 μL $MgOAc_2$ (25 mM), 1 μL AP1 (10 μM), 1 μL Advantage Genomic Polymerase (50 x).

A 1 μL aliquot of primer DFRGSP1 was added to reactions 1A through 5A and 1 μL of PCP1 for reactions 6A through 8A. Subsequently, 1 μL of each DNA library (including the positive control) library was added to the respective tube, whereas 1 μL of water was added to each of the negative controls. Finally, a few drops of mineral oil were added to overlay the contents of each tube. This was followed by a brief spin in a microcentrifuge after which tubes were immediately placed in a Perkin-Elmer thermocycler with cycling parameters set for: seven cycles of 94 °C/25 s, 72 °C/3 min, followed by 32 cycles at 94 °C/25 s, 67 °C/3 min ending with one cycle at 67 °C/7 min.

An 8 μL aliquot from each reaction was analysed on a 1% (w/v) agarose/EtBr gel. For the secondary nested PCR reaction, 1 μL of each primary PCR reaction was added to a sterile 1.5 mL microcentrifuge tube and diluted with 49 μL of water. The secondary PCR reactions consisted of 37.8 μL of sterile water, 5 μL of 10 x 'Tth' PCR reaction, 1 μL dNTP mix (10 mM), 2.2 μL $MgOAc_2$ (25 mM), 1 μL AP2 (10 μM), 1 μL Advantage Genomic Polymerase (50 x).

Table 7.1 A summary of the layout of PCR reactions for cloning the *AaDFR1* promoter

DNA Library (DL)	Tube #	Primary PCR primers	Tube #	Secondary PCR primers
DL-1	1A	DFRGSP1 & AP1 [a]	1B	DFRGSP2 & AP2 [b]
DL-2	2A	„	2B	„
DL-3	3A	„	3B	„
DL-4	4A	„	4B	„
Negative control #1	5A	„	5B	„
Positive control #1 (Control library[c])	6A	PCP1 & AP1 [a]	6B	PCP2 & AP2 [b]
Negative control #2	7A	„	7B	„
Positive control #2 (Pre-constructed [d]control library)	8A	„	8B	„

[a]AP1 is the upstream primer to the Genome adaptor sequence

[b]AP2 is the nested primer to the Genome adaptor sequence.

[c]Control library that was constructed (simultaneously with the anthurium leaf library) from the control human genomic DNA supplied with the kit.

[d]This was the pre-constructed library that came with the kit and serves as a positive control for the PCR.

Primer DFRGSP2 (1 μL) was added to tubes 1B through to 5B; primer PCP2 (1 μL) was added to tubes 6B through to 8B. A 1 μL aliquot was taken from each diluted primary PCR product and added to its respective tube. The contents of each tube was overlaid with mineral oil and briefly spun in a microcentrifuge as before and PCR cycling commenced

with the following parameters: Five cycles at 94 °C/25 s, 72 °C/3 min followed by 20 cycles at 94°C/25 s, 67 °C/3 min ending with one cycle at 67 °C/7 min.

Once completed, an aliquot of 5 µL from each secondary PCR reaction was analysed on a 1% (w/v) agarose/EtBr gel. Only the *Dra* I and *Eco* RV libraries gave a final product and in both cases it was a 1 kb product. This was purified from the agarose gel as described in Section 2.10 and cloned pGEM-T Easy vector according to the ligation protocol described in Section 2.12. XL1-Blue MRF' competent *E. coli* cells were transformed with ligated DNA products (Section 2.13) and plated onto LB-plates containing IPTG/X-Gal (for blue/white selection of transformants) in addition to the appropriate antibiotic.

Plasmid DNA extraction was performed on overnight cultures of selected transformants and each tested for the presence of an insert with single enzyme (*Eco* RI) digest. Those yielding the expected insert size were sent for sequencing using the M13 reverse and forward primers (Section 2.5.8).

7.5 CONSTRUCTION OF *AaDFR1* PROMOTER EXPRESSION VECTOR

The pANA15 (anthurium *DFR* promoter fragment, AaDFRP1, sense orientation in pGEM-T Easy vector) fragment was cloned into a modified version of the pUCBM21 (Roche Molecular Diagnostics, USA) vector to make pPN121. pUCBM21 is a derivative of the pUC 18/19 plasmids that has additional sites for restriction enzymes in the multiple cloning site. The modified version of pUCBM21 had the coding sequence for GUS reporter gene engineered in frame into the *Nco* I site of the vector (Dr. Simon Coupe, Crop & Food Research, New Zealand, personal communication). *Hin*d III and *Nco* I restriction sites were added by PCR to the 5' and 3' ends respectively, of *AaDFRP1* using primers DFRP27 and DFRP28 (Appendix I), to facilitate directional cloning into the modified pUCBM21 vector.

The reaction conditions were as described in Section 7.2.2. The PCR products were separated on a 1% (w/v) agarose gel and purified and quantified. An overnight digestion reaction was performed on both the vector and the pANA15 PCR product, first, with *Hin*d III and then with *Nco* I. The isolation of the promoter fragment and the linearised vector

from the gel, the subsequent ligation of the promoter fragment into the vector, the transformation of *E. coli* and the selection of recombinants by colony hybridisation, was as described in Section 7.3. The orientation of inserts was confirmed by sequencing using the sequence specific primer PSeqA (Appendix I). This primer allowed sequence data to be obtained for the junction of the promoter 3' terminus and the start of the GUS fragment. Particle bombardment experiments were performed with this expression vector as described in Section 7.3.

7.6 RESULTS

7.6.1 Temporal expression of *AaMyb1*

The expression of *AaMyb1* was examined in Altar tissue from Stages 2 to 5 of spathe development and compared with the pattern of flavonoid and anthocyanin biosynthetic gene expression and flavonoid production profile as determined in this thesis. Stages 1 and 6 were excluded because the expression pattern of structural genes in these stages is equivalent to that in Stages 2 and 5, respectively, for Altar. In addition, the low yields of RNA from Stage 6 tissue made poly $(A)^+$ extraction to difficult.

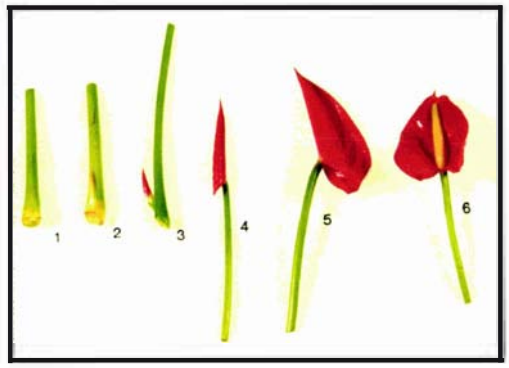
The temporal pattern of *AaMyb1* transcripts did not correlate strongly or consistently with expression pattern of anthocyanin biosynthetic genes (Figure 7.1). *AaMyb1* transcript peaks early in spathe development but declined throughout the later stages examined, suggesting it may have a role in flavonoid production, early in spathe development. This transcript pattern is inconsistent with AaMYB1 regulating transcription of *CHS, F3H* and *ANS*, as transcripts for these genes become more abundant during Stages 4 and 6. However, the expression profile for *AaMyb1* appears consistent with a role in regulating flavone accumulation, which peaks early in spathe development and declines as the spathe matures.

7.6.2 Transient expression studies

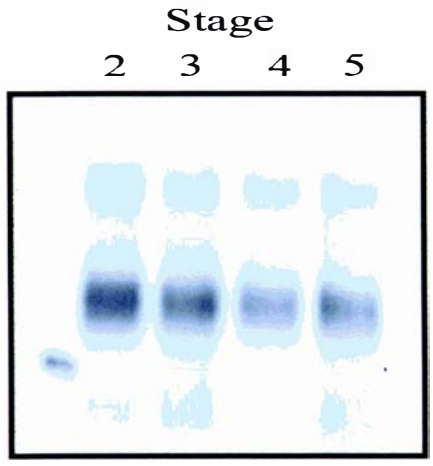
Several particle bombardment experiments were performed to determine if the *AaMyb1* clone could complement the antirrhinum *rosea1* mutant. The first attempt was transient

Figure 7.1. *AaMyb1* gene expression pattern. Panel A was taken from Figure 4.1 and reprinted here for convenience. It shows the size and pigmentation of the spathe throughout the six Stages of development. Transcript abundance for *AaMyb1* was measured for Stages 2 to 5 of spathe development for Altar (Panel B). 5 μg poly $(A)^+$ RNA was loaded per lane and resolved on a 1% agarose denaturing gel. The gel was photographed under UV-light to and shows equal loading for each lane (panel C). The poly $(A)^+$ RNA was blotted unto nylon membrane and hybridised with a radiolabeled 3' end fragment of *AaMyb1* (as an *Eco* Rl fragment from pGEM-T Easy vector).

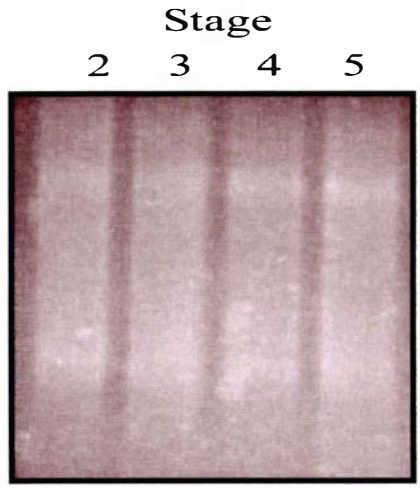
A



B



C



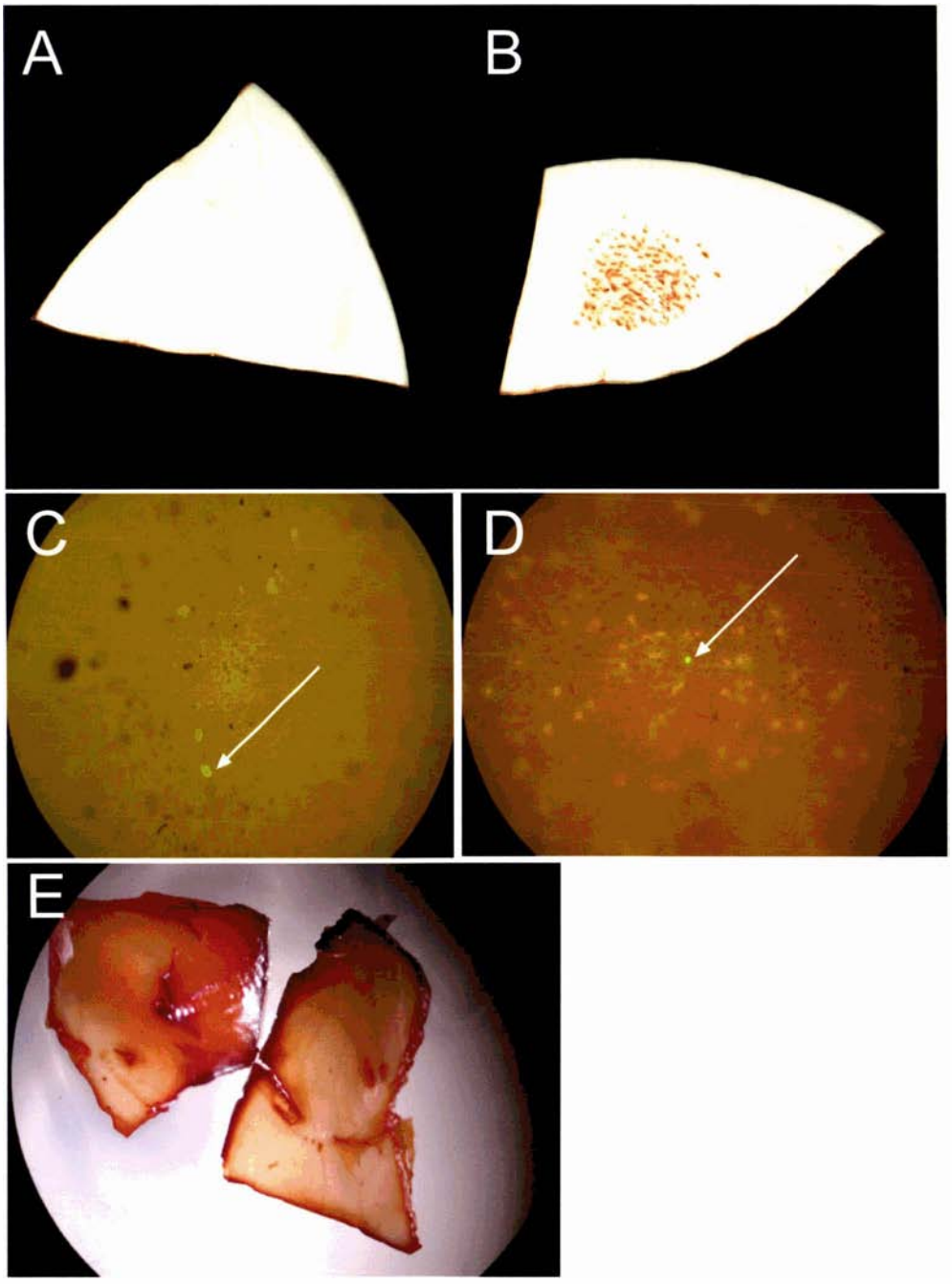
expression of *AaMyb1* driven by the 35S promoter in pART7 (pPN122) with pRT99-GFP included in the particle suspension as an internal control. This plasmid is a pUC18 derivative containing a chimaeric GFP gene controlled by the 35S promoter (Simon Coupe, personal communication). However, although GFP expression was observed in *rosea1* petal tissue, the wild type phenotype was not restored by the introduction of pPN122 into the *rosea1* mutant. Another positive control was done alongside these experiments where antirrhinum *rosea1* tissue was bombarded with pLN81 (*Rosea1* cDNA in pART7) and as expected, the wild type phenotype was restored, producing single pigmented epidermal cells (data not shown).

The possibility that *AaMyb1* was unable to complement the *rosea1* mutant phenotype because it lacked a suitable bHLH partner in antirrhinum tissue was tested by co-bombardment of *rosea1* mutant tissue with the maize *Lc* gene (also under 35S in pART7). However, this did not restore pigmentation (data not shown). Again, pRT99-GFP was included in the suspension as a positive control and GFP expression was observed in the tissue upon examination.

As functionality of AaMYB1 could not be confirmed in a heterologous system, the possibility that the anthurium Myb required its own endogenous bHLH partner was tested using particle bombardment experiments with white spathe tissue from Acropolis. Acropolis is likely to be a regulatory mutant based on gene expression studies (Chapter 4) and this may be because of a *Myb* mutation, making it a likely candidate for complementation with AaMYB1. Initially, little headway was made with these experiments, as the high phenolic levels in anthurium tissue resulted in heavy browning of the spathe immediately upon bombardment with the particle suspension (Figure 7.2). In these tissues, no GFP expression was observed. Several parameters for the experiment were adjusted in attempting to alleviate this problem. A range of helium pressures (200 kPa-450 kPa) was used in combination with adjusting the distance of the tissue from the nozzle.

However, none of these factors changed the rate at which browning occurred or the extent to which the browning developed. As an alternative, liquid media with rafts, as opposed to

Figure 7.2. Transient expression assays in anthurium spathe tissue. Panel A shows normal Acropolis spathe tissue that has not been wounded by bombardment compared to panel B showing the same tissue following particle bombardment. Note the wound response in panel B. The browning started within 2 min of wounding and intensified during the overnight incubation on agar. In panel C, the arrow identifies a GFP signal in Montana tissue following bombardment with pRT99-GFP. Panel D shows the most conspicuous and convincing GFP expression (identified by the arrow). This was obtained when Montana was bombarded with Tungsten-20. Panel E shows the extensive browning that occurred when the GUS histochemical assay was attempted with anthurium tissue.


A
B
C
D
E

bombardment on agar was used. Enough culture was used just to cover the surface of the rafts and tissue was bombarded while on the raft. The distance from the nozzle was 13 cm from the tissue and 1 μM gold was used with a helium pressure of 300 kPa in a partial vacuum of -14 kPa and 30 s solenoid opening time. A helium pressure of 300 kPa was chosen, because above this, the wounding of the tissue was too severe as evidenced by more rapid and extensive browning. No significant improvement was obtained with this procedure as neither GFP expression nor complementation, as evidenced by anthocyanin pigmentation, was observed.

As it was thought to be oxidation of the phenolics that was resulting in browning, ascorbic acid and PVP (concentrations ranging from 0.05 M-0.5 M) were tried individually or in combination. Two approaches were used. Firstly, tissue was soaked in solutions of the two reagents before and after bombardment. In some cases the tissue was incubated overnight in the solutions as opposed to incubation on #2 Media. As another approach, the tissue was vacuum infiltrated with ascorbic acid and/or PVP and then bombarded with the particle suspension. Of these two approaches, only the second was able to prevent the browning. However, vacuum infiltration appeared to have a detrimental effect on the tissue, as no GFP expression was observed in antirrhinum tissue subjected to the same treatment.

Although no GFP expression was obtained with Acropolis, there was some success with cultivars with coloured spathes, with Montana providing the best background (Figure 7.2). However, GFP expression was still faint by comparison to that obtained with antirrhinum. The optimum GFP expression was obtained when Tungsten-20 was used as the DNA microparticle carrier (Figure 7.2). All other parameters remained unchanged.

7.6.3 In vitro promoter studies

Given the lack of success at complementing the *roseal* mutant phenotype with pPN122, alternative experiments were designed to compare the regulatory mechanism in anthurium with that of antirrhinum. A promoter fragment for anthurium *DFR* was cloned using the

Genome Walker protocol, and comparative studies were performed to test the activation of promoters in the two plant species.

7.6.3.1 Cloning an anthurium DFR promoter fragment

A 905 bp promoter fragment was cloned from the Genome Walker library digested with *Eco* RV and *Dra* I. Both fragments were identical when their sequence data were analysed and only the *Eco* RV fragment was used in further experiments. The nucleotide sequence is shown in Figure 7.3 and several TATA and CAAT box sites are marked out. The position of the two *DFR* gene specific primers is shown. The -GAGA-, -CACA- and -TATA- stretches that were present in the 5' untranslated region of the *AaDFR1* cDNA clone were also present in the promoter fragment. The DFRGSP2 primer was clearly identifiable from the M13 forward sequencing reaction.

An examination of the promoter for *cis*-elements using the PLACE database (www.dna.affrc.go.jp; Higo et al., 1999) and the TRANSFAC database on GENOMENET (www.genome.ad.jp), identified the Myb DNA binding consensus site MBSIIG (Jin and Martin, 1999) at position -327 relative to the translation initiation ATG. The computer-aided searches also identified two maize P consensus sequences (Grotewold et al., 1994) at -327 and -848, (Figure 7.3). Several other *cis*-elements were identified and the complete listing is available in Appendix X.

7.6.3.2 Heterologous promoter activation experiments

Experiments were performed to determine if antirrhinum anthocyanin regulatory proteins can recognise and bind to promoters for structural genes involved in anthocyanin biosynthesis in anthurium, and reciprocal experiments were planned for antirrhinum structural gene promoter-constructs introduced into anthurium tissue. Antirrhinum wild type tissue was bombarded with the pPN121 (chimaeric anthurium *DFR* promoter fragment attached to the GUS reporter gene in pUCBM21 vector). Four control experiments were also performed. One used gold alone as a negative control for the wounding produced by bombardment. An antirrhinum *DFR::GUS* construct (pJAM1311, supplied by Dr. Cathie Martin of the John Innes Centre) was introduced into wild type and *roseal* petal tissue as a

Figure 7.3. Nucleotide sequence and conserved motifs in the anthurium DFR promoter. The portion of the promoter (905 bp) was analysed for conserved *cis*-regulatory elements using the PLACE database and TRANSFAC databases. The Myb motif, MBSIIG (marked in red), is the binding site preferentially recognised by several anthocyanin Myb regulators (Jin and Martin, 1999). Two core consensus sequences for the maize P protein (Grotewold et al., 1994) (marked in blue) were also identified. The start of the promoter sequence is numbered from -1 from the proposed translation initiation ATG. Nucleotides downstream of the promoter are marked in green and include the two DFR gene specific primers (DFRGSP1 and DFRGSP2) used with the Genome Walker kit (highlighted in yellow). DFRGSP2 includes the first 3 bases upstream of the first in frame ATG. Several potential TATA and CAAT box sites have been underlined.

maize P core consensus

-905 ATCCAATATAGTTACTCAACTTGTTGAGAGGTCTCTTAGACCAGAAGTGTAGCGAAA GGAAGGTGATCCAGATGGGAATCGAGG

-815 TCAGACTATCCTTCTCCAATCGAAGTTTTGGGATCCATTTTTGCAAAACAAAGGGATGCCCCTTGATGCCCCAGAACCCTTCCAATGCCG

-725 GGACCATATCTTCCTCAAATACTGTAATCTAGTAATAATATAGGTTTTGTGATTTTGTATTGAGAGGAGGCCGATGCTTTTGAAGAGAAA

-635 GGGAGCCCGACCTCCCACCTTCTCTAGGAAGGTCCTCCATCCGAGTCAAACTCGCGGTTCGGAGAGGTGCAACGAGAGGATCCGGCCTAG

maize P core consensus

-545 AGGATCCAAATTCCAGAGTCCGCAATCCGAAGGTTCGGAAGAAGCCG AGCTACCAG TCTTCGCGCGTCCGTCTCCTTGTACCAACGGTAA

-455 ACAGAACTCTTCCCGCCACGGGGACTACCAAACGACGAACCACTACACGACCCATATGCATATCCGTACAAAATGACGAATTACTCTCGC

MBSIIG binding site

-365 ACATCGCCTAATTGTCCTCTCCTCATGCAGCGAAGAGTGTTAGGTGGTTAGTTTGAACACATGACTACAGCAGGAGGTAGATGACGGCCG

-275 GGGCCCTGAGACTGAACGCTCAAGCTGTCATATGTGACATGTCCCAGCAAATTTATGAGCTACCCACGTCTATATAAATAGAGAGAGAAG

-185 AGCGTAGGGTAGACACAGTTGCCTCCGATTTGCTGCTCCATCGATCTCTCCTTCCCACACACACACACACACACACACACACACAC

-95 ACATATATATATAGAGAGAGAGAGAGAGAGAGGGAGAGAGCTTTTATCAGTTTCTGCCTCTATAAACAACAGCTGGTTATATATAATCTG

-1

-5 CGAGG ATGATGCACAAGGGCACCGTGTGCGT GACGGGCGCTGCCGGGTTC TTTGGCTCATGGCTGATCATGAGGCTC

DFRGSP2 DFRGSP1

positive and negative control, respectively. As a positive control for the GUS staining process, a regal pelargonium (*Pelargonium X domesticum)* cultivar Dubonnet stably transformed with a *35S::GUS* construct (Dr. Murray Boase, Crop & Food Research, New Zealand, personal communication) was used.

Positive GUS foci were observed for antirrhinum wild type tissue bombarded with pPN121. GUS foci were also evident in the *roseal* mutant tissue bombarded with the same construct (Figure 7.4). In contrast, the positive control using pJAM1311 gave GUS foci in antirrhinum wild type but not *roseal* petal tissue. The gold-only negative control had no GUS foci indicating the results observed with the promoters were not a wound response.

The reverse experiments (to determine if anthurium anthocyanin regulatory proteins can recognise and activate promoters of antirrhinum structural genes for anthocyanin biosynthetic enzymes) were attempted. However, heavy browning of the tissue occurred during incubation in the GUS assay buffer. To circumvent this problem *promoter::GFP* constructs were made. However, time did not allow for the transgenic experiments to be performed prior to submitting the thesis.

7.7 DISCUSSION

While highly conserved Myb protein structural motifs, including some of those found in anthocyanin related Myb proteins, were observed in the AaMYB1 predicted protein, it is well known that Myb proteins with very similar DNA binding domains can have distinct physiological functions. For example, antirrhinum MIXTA and maize C1 have been classified into the same Myb subgroup based on the similarity of their DNA binding domains, yet MIXTA is involved in specifying the formation of conical cells in petal epidermis while C1 is involved in regulating anthocyanin biosynthesis (reviewed in Jin and Martin, 1999). Therefore, for Myb regulatory proteins it is not possible to accurately assign function simply based on the amino acid sequence. Consequently, further evidence that *AaMyb1* encodes a protein that regulates anthocyanin or other flavonoid biosynthesis was required.

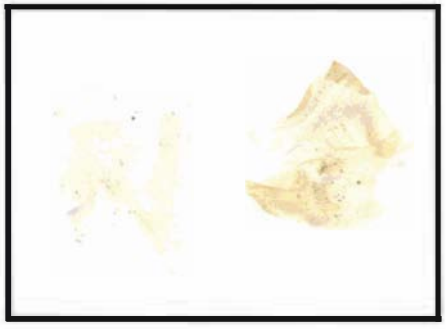
Figure 7.4. Transient expression experiments with antirrhinum and anthurium *DFR promoter::GUS* constructs. The constructs are shown beneath the respective panels as pJAM1311 and pPN121 respectively. The anthurium DFR promoter was activated in both wild type and *rosea1* mutant petal tissue (A1 and A2 respectively) whereas the antirrhinum promoter was activated in wild type tissue (B1) but not in *rosea1* tissue (B2). Anthocyanin pigment in wildtype flowers degraded during incubation thus allowing GUS foci to be observed.

Antirrhinum *rosea1* mutant

Antirrhinum wild type

A1

A2

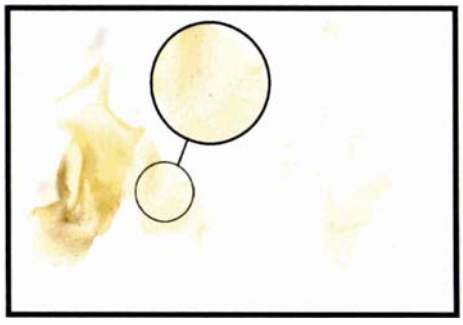


Antirrhinum wild type

Antirrhinum *rosea1* mutant

B1

B2



Anthurium *DFR::GUS* construct: pPN121

Antirrhinum *DFR::GUS* construct: pJAM1311

One would expect transcript levels of a transcription factor to peak just prior to that of its target gene(s). The high transcript levels at Stage 2 for *AaMyb1* might be expected, given the expression profile of *CHS, F3H, DFR* and *ANS* in early spathe development. However, the decline in *AaMyb1* transcript levels as the spathe matures, does not correspond to the high transcript levels detected for *CHS, F3H* and *ANS* in the equivalent stages of the spathe. Therefore, the temporal distribution of *AaMyb1* transcripts is inconsistent with a role for the protein in regulating anthocyanin biosynthesis in anthurium spathe.

With the expression pattern of anthurium *DFR* (Chapter 4), it is possible that a separate regulator exists in anthurium spathe tissue that specifically targets *DFR*. Should this be AaMYB1 then it is possible that its expression pattern oscillates with the same diurnal rhythm as its target gene, making estimation of AaMYB1 development pattern difficult. This pattern, where transcription levels of regulatory genes cycle in the same rhythm as their target genes, has been suggested for several regulators (Harmer et al., 2000).

Another possibility is that AaMYB1 regulates flavone biosynthesis in the spathe. The biochemical profile of non-anthocyanin flavonoids (that for anthurium would primarily be flavones) showed decreasing levels from Stage 2 (as discussed in Chapter 4). This fits well with the transcript profile for *AaMyb1* and is consistent with the earlier suggestion that *AaMyb1* may encode a protein that has a role in flavonoid biosynthesis starting at the early stages of spathe development. The abundance of flavones in anthurium tissue may also explain the predominance of this clone among the *Myb* candidates obtained from the PCR reactions (Chapter 6). If anthurium clones for *CHI* and *FNS* were available, experiments could be performed to test this hypothesis. Separate regulators for flavone and anthocyanin biosynthesis in anthurium would then suggest that *CHS* and *CHI* are subject to different regulatory control. However, this is not supported by northern data discussed in Chapter 4 that showed *CHS* and *F3H* have a similar temporal expression pattern to *ANS*. One possibility that harmonises the two is that anthurium *CHS* exists as a multigene family and different members of the family respond differently to the specific regulators. Differential regulation of gene family members by Myb transcription factors was shown for *CHS* family in petunia (Quattrocchio et al., 1993).

It is well known that for some plants, different regulators control the activity of EBGs and branches in the flavonoid pathway, and that they may not activate the anthocyanin specific genes (Jackson et al., 1992; Quattrocchio et al., 1993, 1998). Further to this, even though C1 activates all the genes for flavonoid biosynthesis in maize kernels, it could not activate the EBGs in petunia tissue. On the contrary, AN2 was able to substitute for PL and activate the entire suite of genes in maize, even though it does not activate the EBGs in petunia. This implies that regulatory proteins are evolutionarily conserved between species and that species-specific patterns of regulatory gene activity are probably due to changes in the *cis*-elements of target genes (Quattrocchio et al., 1998).

One draw back of these assays is that they cannot tell if the Myb factor being tested was at least able to bind to the target DNA. Should it be that AaMYB1 binds but cannot activate the anthocyanin structural genes in antirrhinum *rosea1* petals, then it may be acting as a competitor in a role similar to that suggested for antirrhinum MYB305, which has been implicated in regulating flavonol production. MYB305 can bind to *DFR* and other LBGs but is unable to activate them (Sablowski et al., 1994).

In suggesting that AaMYB1 regulates EBGs and flavone biosynthesis, the possibility of alternative Myb interaction partners is introduced. bZIP proteins that recognise the ACGT core sequence in the target DNA, have been implicated as possible partners for MYB305 in activating early genes such as *PAL* in antirrhinum (Sablowski et al., 1994).

The existence of a separate pathway for the regulation of flavone and flavonol biosynthesis has been suggested to have evolutionary significance in flowers, given the roles of these pigments in screening out damaging UV-B rays and as insect attractants (Martin et al., 1991). The fact that there is no reduction in flavone levels in white anthurium spathes is further evidence for the existence of a separate regulatory pathway for these compounds. However, further work is necessary to confirm that AaMYB1 is involved in flavone production.

Undoubtedly, a much stronger indicator that AaMYB1 plays a role in anthocyanin biosynthesis would have been its ability to complement the *roseal* mutant phenotype and induce the formation of anthocyanins. Transient expression assay in heterologous systems, by particle bombardment, was chosen because it has been used successfully to investigate the functionality of other anthocyanin regulators (de Majnik et al., 1998; Quattrocchio et al., 1998; Schwinn, 1999; Nesi et al., 2001), including one from the gymnosperm black spruce (Charest et al., 1994). Antirrhinum mutants were chosen because of the ready availability of the *roseal* mutant line and the necessary antirrhinum plasmid DNA constructs, along with their ease for use in particle bombardment-mediated transient gene expression.

The inability of anthurium AaMYB1 to complement the *roseal* mutant phenotype suggests it does not encode a Myb involved in regulating anthocyanin biosynthesis. However, aspects of the antirrhinum regulatory system for anthocyanin biosynthesis may make heterologous complementation with Myb proteins difficult. In previous work, it was found that C1 could not complement *roseal* on its own, but required its own bHLH partner (Lc) in the bombardment experiments (Kathy Schwinn, personal communication). This may be the case for AaMYB1 also. However, this specificity for the bHLH partner is variable, as shown in the ability of arabidopsis TT2 MYB to substitute for C1 function in maize kernels when used in combination with a member of the maize bHLH family (B-Peru) (Nesi et al., 2001). Also both AN2 and C1 were able to interact with JAF13 or Lc (Quattrocchio et al., 1998).

Further evidence that suggests an alternative role for AaMYB1 other than the regulation of anthocyanin biosynthesis came from sequence analysis. Six residues within the Myb DNA binding domain that specify interaction between Myb proteins and bHLH transcription factors were identified by in vitro mutagenesis studies in C1 (Grotewold et al., 2000). These residues are conserved in all Myb proteins to date that have been shown to be involved in regulating anthocyanin biosynthesis. Two non-conservative substitutions are found in AaMYB1 as discussed in the previous chapter (Figure 6.9). However, it is not known if the mode of action observed in maize Myb/bHLH interaction is the same in other

plant species, and the possibility of a different mode of interaction is being investigated in antirrhinum (Cathie Martin, personal communication).

Given the results from transient expression in *rosea1*, performing transient expression in an anthurium background would allow AaMYB1 to interact with the endogenous bHLH (or some other transcription partner). Acropolis was chosen because experiments in Chapter 4 suggested that it is a mutant for a regulatory factor. However, in the absence of direct genetic proof, it was assumed that this mutation was in the Myb regulatory loci and not the bHLH factor. The data gained from these experiments was inconclusive because of the extensive and rapid browning that was characteristic of the tissue upon wounding. It is likely that the phenolics released by the tissues severely affected the viability of the cells within the range of its spread. Furthermore, in some of the experiments using coloured tissue, GFP expression (though marginal) was only observed outside the wound site. In these zones of minimal GFP expression it is possible that fewer phenolics were released and consequently more cells were able to survive and express the DNA construct.

The use of antioxidants such as ascorbic acid was able to alleviate the browning in white spathe tissue only when tissue was vacuum infiltrated and left overnight in the ascorbic acid solution. The effect of this treatment on cell viability was estimated by subjecting antirrhinum lobe tissue to the same treatment, the result of which was the complete elimination of GFP expression. Therefore, while ascorbic acid and PVP can eliminate browning in transient expression experiments involving anthurium spathe tissue, both individually or together, have a deleterious effect on cell viability (as indicated by the loss of GFP expression) that makes their use in such experiments counterproductive. In the final analysis, the heavy browning in the white anthurium spathe prevented meaningful functional analysis of AaMYB1 in the homologous background.

Using rafts and liquid culture in highly vented vessels has reported to be successful for transient expression in tissues with high phenolic levels, such as *Aconitum nepalus* (Gerald Tanny, Osmotek Ltd, personal communication). Solid media is a disadvantage because of the slower rate of diffusion compared to liquid media. Therefore, on solid media gradients

are established and high molecular weight compounds such as phenolics accumulate in the immediate vicinity around the wound, resulting in cell death. Liquid medium should allow the phenolics to diffuse away quickly, minimizing cell death. It is possible that with further adjustments this protocol could be successful for transient expression in anthurium.

A protocol for transient expression in anthurium spathe tissue was developed for some cultivars of anthurium. Particle bombardment experiments for GFP expression were only successful when coloured anthurium lines were used. It is possible that the anthocyanin glycosides provided an endogenous supply of a non-deleterious antioxidant to counter the phenolic oxidation. However, initially, the expression level of GFP was marginal (Figure 7.2). The characteristic strong luminous fluorescent glow definitive of robust GFP expression was only realised when Tungsten-20 was used instead of gold as the DNA carrier (Figure 7.2). Unlike gold with its smooth, round edges, tungsten is irregularly shaped with sharp edges on its surface. This translates to more wounding but also deeper penetration within the tissue (Dr. Simon Deroles, Crop & Food Research, New Zealand, personal communication). This could be significant for anthurium given the thickness of the mature spathe tissue. Also, as anthocyanin production is confined to sub-epidermal tissues this may affect the use of spathe tissue for anthocyanin transient analysis. The procedure would need to be optimised for consistency and greater levels of target gene expression.

Securing a *DFR* promoter fragment allowed further research into the regulation of anthocyanin biosynthesis in anthurium. It is likely that the promoter fragment obtained by PCR is that for the gene that corresponds to *AaDFR1*. Although it is not known if anthurium has a multigene family for DFR, as Southern analysis was inconclusive, the position of the *DFR* gene specific primer (DFRGSP2) at the junction of *AaDFR1* open reading frame provided the 5' untranslated sequence region as a check for the specific DFR clone corresponding to *AaDFR1*. The 5' region of genes is very divergent and the *AaDFR1* cDNA clone had in excess of 200 bp of 5' untranslated region (Chapter 3). Analysis of the promoter sequence and the cDNA clone confirmed that this region is identical in the two DNA fragments.

The positive GUS expression obtained when antirrhinum wild type lobe tissue was bombarded with anthurium *DFR::GUS* (Figure 7.4) suggests that the fragment cloned from the Genome Walker library contains a functional promoter region for anthurium *DFR*. Also, the absence of GUS expression when petal tissue from the *rosea1* mutant line (Figure 7.4) was bombarded with antirrhinum *DFR::GUS* confirms that ROSEA is one of the anthocyanin regulatory proteins required for antirrhinum *DFR* promoter activity. That the antirrhinum anthocyanin regulatory proteins (in wild type tissue) can recognise and activate the anthurium *DFR* promoter, suggests that the petal tissue of the dicot antirrhinum shares common regulatory mechanisms with spathe tissue of the monocotyledon anthurium. However, GUS activity in the *rosea1* mutant bombarded with pPN121 construct indicates that ROSEA was not required for the activation of this monocot anthocyanin biosynthetic promoter (Figure 7.4). If it is that ROSEA did not activate anthurium *DFR* in wild type tissue, then it suggests that either ROSEA is restricted in the range of recognition sequences it can bind to, or that, in antirrhinum the activation of the GUS gene through the anthurium *DFR* promoter does not require a Myb factor or uses a different Myb to ROSEA1.

ROSEA1 is only one of a family of Myb transcription factors that are involved in regulating flavonoid and anthocyanin biosynthesis in antirrhinum petal tissue (Jackson et al., 1991; Martin et al., 1991; Schwinn, 1999). It is possible that one of the other Mybs is responsible for the activation of the anthurium *DFR* promoter, although the antirrhinum *DFR* promoter construct gave no GUS foci in the *rosea1* line, suggesting lack of other Myb proteins that are active on that promoter. Furthermore, the lack of pigment in the epidermal tissue of the rosea1 mutant suggests that none of the other anthocyanin Mybs are active.

Activation of the GUS gene through the anthurium *DFR* promoter in *rosea1* petal tissue suggests the involvement of other transcription factors, a suggestion supported by the presence of a large number of other putative *cis*-regulatory elements in the anthurium *DFR* promoter. The activation of plant genes by more than one transcription factor may allow the plant to respond to a variety of environmental stimuli such as wounding, light, temperature for example. Anthurium *DFR* and possibly the other genes in the flavonoid pathway may

have *cis*-regulatory elements that are wound responsive. Possibly a wound responsive factor in the bombarded antirrhinum tissue was induced and was able to activate the *DFR* promoter. This possibility is supported by the rapid wound response characteristic of anthurium spathe tissue. This may be a feature of anthurium flavonoid biosynthetic gene promoters not shared by the antirrhinum biosynthetic gene promoters.
Another possibility for the activation achieved through the anthurium *DFR* promoter in *rosea1* petal tissue comes from recent studies on the promoter region of the grape *DFR* gene (Gollop et al., 2002). High external calcium concentration (up to 50 mM) induced *DFR* expression under white light in grape cell cultures. Anthocyanin gene expression, was also shown to be positively correlated with calcium concentration, under UV light in pharmological studies of *CHS* gene expression in arabidopsis and soybean cell suspension cultures, and parsley protoplasts culture (Christie and Jenkins, 1996; Frohnmeyer et al., 1997, 1998). Fifty μL of 2.5 M $CaCl_2$ was used in preparing the gold suspension for particle bombardment experiments in this thesis and it is thus possible that the anthurium DFR promoter is activated by calcium via a ROSEA1-independent route.

The results suggest that, multiple signals, and possibly multiple Myb/bHLH regulatory factors (or other regulatory proteins), act on the anthurium *DFR* promoter. These signals may be related to development, the diurnal rhythm, and possibly a wound-or calcium-mediated response. This complex network, that possibly involves one or more signal transduction pathways, could compensate for the absence of ROSEA1 in the promoter/reporter experiments described in this chapter.

While an anthocyanin specific Myb was not obtained, it was necessary and logical to pursue and investigate the role of the Myb family of genes as possible anthurium anthocyanin regulators. The case for conserved regulatory systems in antirrhinum and anthurium was partially supported but would be made stronger by performing transient expression experiments with antirrhinum anthocyanin biosynthetic gene *promoter::GUS* or *promoter::GFP* constructs in anthurium tissue. This would also allow confirmation that white spathe lines correspond to regulatory mutants. This was the original intention, but the difficulties of reporter gene expression in anthurium hindered progress in this regard.

The promoter experiments infer a role for transcription factors other than those of the Myb type, that, activate the promoters of anthocyanin specific genes in anthurium spathe. The results do not eliminate the role of other Myb transcription factors but clearly show that ROSEA1 is not required for the activation of the anthurium *DFR* promoter. These experiments also strengthen one of the key observations from the studies to date, that of *DFR* being a key regulatory target with unique expression features.

The evidence in this chapter is probably against *AaMyb1* encoding an anthocyanin regulator. Possible explanations for the lack of success in isolating a cDNA for the appropriate Myb are that the Myb may have very low expression levels, (although the strong pigmentation in the spathe contradicts this idea), or it may be specific for *DFR* and consequently exhibits unusual transcript abundance patterns with diurnal variations peaking at one or a few times of the day. Consequently, this point could have been missed in library construction and the PCR reactions with the Myb degenerate primers. One must also bear in mind that of the 80 plus Myb fragments generated in arabidopsis (Romero et al., 1998), only two of these were anthocyanin related MYBs (MYB75/PAP1 and MYB90/PAP2), suggesting that numerical representation is critical in obtaining anthocyanin Myb clones in plants given the large size of the Myb family.

Identifying the anthocyanin Myb and bHLH factors that regulate anthocyanin biosynthesis in the spathe would provide useful additions to the existing body of knowledge in this field, as there are no equivalent organ structures in the traditional models studied, and limited information is available for ornamental monocots. Heterologous screening, of the cDNA library with *Rosea1* and *C1* or other anthocyanin *Myb* cDNAs may prove a successful alternate route to isolating such factors. It may also be interesting to over-express AaMYB1 in arabidopsis or maize lines lacking expression of specific Mybs, for example the *tt2* or *C1* mutant.

Chapter 8

General discussion and conclusions

Full-length clones for *CHS, F3H, DFR* and *ANS (AaCHS1, AaF3H1 AaDFR1* and *AaANS1*, respectively) were isolated from spathe tissue of the anthurium cultivar Altar. The predicted amino acid sequences encoded by the cDNA clones have structural features that are highly conserved in equivalent functionally defined proteins. Phylogenetic analyses for the different clones strengthened the case for their putative identities.

Although anthurium is a monocot, only two of the four biosynthetic clones grouped with similar proteins from other monocot species. This first molecular study in anthurium suggests that while its botany defines it as a monocot, the molecular characteristics for some genes involved in anthocyanin biosynthesis converge towards those of dicots. AaCHS1, in particular, showed a weak association with CHS proteins from other monocots. In fact a non-parametric bootstrapping analysis showed that AaCHS1 was locally unstable in the phylogeny with little support for grouping with the monocot and even less for non-related CHS proteins (Figure 3.2). The weak grouping of AaCHS1 with the dicots and more distant phylogenetic distance from the monocots suggests that AaCHS1 is genetically distant from other homologous sequences.

Undoubtedly the ultimate proof of identity is functional characterisation by biochemical assay with the anthurium protein expressed from suitable expression vectors or transgenic complementation of mutants in a compatible heterologous or homologous host by transient assay or stable transformants. However, barring functional characterisation, it can be said with reasonable confidence that the anthurium cDNA clones obtained were derived from the respective anthocyanin biosynthetic enzyme mRNAs in anthurium tissue.

Some of the flavonoid biosynthetic genes are known to exist as gene families, such as *CHS* in petunia (Koes et al., 1989) and *DFR* in gerbera (Helariutta et al., 1993). It is not known whether such gene families are present in the anthurium genome as Southern analysis for all the genes proved inconclusive because of technical difficulties. Therefore, it is not known whether the cDNA clones obtained represent the only copies of the genes in the genome. However, in all the isolation efforts, several fragments were obtained for each clone and all the same restriction fragment patterns. Nevertheless, it would be worthwhile

repeating the Southern analyses, as the complete and accurate restriction maps for each clone are available.

Anthurium presented a good model for investigating the regulation of anthocyanin biosynthesis in an ornamental monocot. In addition, there is no equivalent structure to the spathe of anthurium in the ornamental dicots studied. Added to this, anthurium flowers have an extended flower development period of eight-nine weeks. In some ways, the divergence seen in the four-anthurium cDNA clones away from monocots, was also reflected in the transcript abundance profiles of the genes. The pattern of coordinate transcriptional regulation of the genes in the flavonoid pathway as seen in maize, does not apply to anthurium. Furthermore, there are dicot species where all the flavonoid genes appear to be regulated together (Boss et al., 1996; Gong et al., 1997). The findings of this work support the proposal that the classification of regulatory models based on monocots and dicots is no longer scientifically accurate or relevant.

Clearly *DFR* is selected out and regulated differently in the spathe to *CHS, F3H and ANS* Not only is its temporal expression pattern different to that of the other three genes (Figure 4.6), there is, in addition, a diurnal rhythm to its expression (Figure 4.9). It is not known how widespread this pattern may be, although rhythmic expression of genes in the phenylpropanoid pathway has been reported for arabidopsis (Davis and Millar, 2001). Therefore, it is important to factor in rhythmic expression, whether circadian or diurnal, of flavonoid biosynthetic genes when investigating temporal transcript abundance in plant tissue.

The evidence suggests that, at least for the spathe, anthocyanin production is concomitant with *DFR* expression (Figure 4.5 and 4.7). However, the diurnal flux in *DFR* transcript levels, suggests that *DFR* expression is linked to other processes within the plant apart from anthocyanin production. This is supported by the transient expression studies with the *DFR* promoter. In this same vein, the diurnal rhythm of *DFR* expression implies the involvement of light responsive signal transduction pathways affecting *DFR* expression. *DFR*

expression, therefore, may serve as a point of regulation for anthocyanin biosynthesis as well as that of other flavonoids in the spathe.

This, however, does not eliminate the possibility of an equally significant role for genes such as *UFGT* (as seen with grape berries, Boss et al., 1996) or *GST* as regulatory targets for anthocyanin biosynthesis in the spathe, along with *DFR*. That flux could be invested in other points within the pathway besides *DFR* is strengthened by studies in the spadix, which clearly showed that for Altar, high *DFR* levels are not necessarily correlated with anthocyanin production. It further implies that the reduced anthocyanin level in the Altar spadix is probably due to a mutation in a structural gene such as *UF3GT*. Kamemoto et al. (1988) suggested that *C* and *R* control spadix colour and that they are absent in yellow spadix lines. These are distinct from the *M* and *O* locus that control pigmentation in the spathe. Possibly, *C* and *R* are anthocyanin specific genes acting after *ANS* in the pathway and these are mutated in the yellow spadix of Altar.

It is possible that in the spathe there is a separate regulatory apparatus that specifically targets *DFR* expression and another that regulates *CHS, F3H and ANS*. This could explain the uniqueness of *DFR* transcript profiles in this tissue (Figure 4.6 and 4.7). Furthermore, northern analysis of Acropolis does strengthen the idea that mutations at regulatory loci account for acyanic anthurium lines. However, the fact that white lines continue to accumulate flavones at levels comparable to wild type lines, points to separate regulatory machinery controlling flavone production in the spathe. On this basis, genes acting early in the pathway, such as *CHS* and *CHI*, may be targets for such a regulator and their expression should remain unchanged to support flavone production in white lines. However, a reduction in transcript levels of all four genes including *CHS* was detected in the white spathe of Acropolis (Figure 4.7). This apparent contradiction can be reconciled if there is more than one copy of *CHS* in the anthurium genome and that each is targeted separately by different regulators as is seen with petunia (Quattrocchio et al., 1993).

The regulatory model suggested for anthurium spathe tissue (Figure 4.10) is a hybrid between coordinate regulation (Ludwig and Wessler, 1990; Boss et al., 1996; Gong et al.,

1997) and regulation of the genes in distinct subsets (Martin et al., 1991; Davies et al., 1993; Quattrocchio et al., 1993; Shirley et al., 1995; Oren-Shamir et al., 1999). For anthurium, the regulation of the pathway is divided into two subsets, one comprised of *CHS, F3H and ANS* and the other subset includes *DFR*. However, within this pattern, an aspect of coordinate regulation is seen, with *CHS* being regulated with anthocyanin-specific genes such as *ANS*.

The results add new perspectives to the genetics of anthurium flower colour. The two genetic factors controlling spathe colour are *M* (determining whether pelargonidin 3-rutinoside or cyanidin 3-rutinoside are produced) and *O* (controlling whether any anthocyanins are produced). Hypothetically, *O* could have been any of the structural genes *CHS, CHI, F3H, DFR, ANS, UFGT* or *GST*, as its homozygous recessive state prevents anthocyanin production irrespective of the nature of the *M* locus. However, the reduction in transcript levels for *CHS, F3H, DFR* and *ANS* in genotypically defined cultivars recessive at the *O* locus (1349 and 1244) along with Acropolis, suggests that *O* is a regulatory gene that targets these four genes. However, it is not yet clear how this fits in with the suggestion that there exists a separate regulator for *DFR* and for flavone production. Furthermore, the relationship between *O* and *M* requires further investigation. While the most likely candidate for *M* is *F3'H*, no cDNA clone for this gene was obtained from the library and neither could biochemical assays confirm its functional presence. While technical factors cannot be excluded as a possible explanation for the assay results, it is also possible that anthurium uses an alternative pathway to accumulate DHQ, possibly through its CHS utilising caffeoyl-CoA rather than the usual 4-coumaroyl-CoA as a substrate (Heller and Forkmann, 1988, 1993). With the availability of the putative anthurium *CHS* cDNA clone, biochemical assays can be set up to compare the effectiveness of both caffeoyl-CoA and 4-coumaroyl-CoA as substrates.

Clearly more detailed experiments are required to elucidate the apparently complex regulatory network involved in flavonoid production in anthurium spathe tissue. A detailed molecular analysis of more cultivars with defined genotypes is required. These experiments would be more conclusive with the cloning of anthurium *CHI, F3'H, UFGT, GST* and Myb/

bHLH regulatory genes. In addition, northern analysis experiments should be performed with the gene specific 3' end of the cDNA as opposed to the full-length clone. Although high stringency washes were used, probes made from the 3' end of the clones would have ensured that the hybridising signals obtained were to the specific gene probe and not another member of the gene family, assuming that one exists.

That combinatorial action between Myb and bHLH proteins is the mechanism for regulating anthocyanin biosynthetic genes is well known. Consequently, *Myb* genes were the first targets in an effort to identify the regulatory proteins that control anthocyanin pigment production in anthurium spathe tissue. A full-length clone (*AaMyb1*) was obtained and analysis of its N-terminal end confirmed its identity as an R2R3 Myb with high % identity in its DNA binding domain to anthocyanin Mybs from both monocot and dicot species. Further to this, the presence of an acidic domain in the C-terminus suggests that AaMYB1 may be involved in transcriptional activation.

Although AaMYB1 showed highest % identity to TT2, C1 and PL in its Myb domain, this does not imply commonality of function. This is because a number of other factors, apart from the sequence of the Myb domain, such as protein-protein interactions, may influence the ability of the protein to recognise target sites (Jin and Martin, 1999). Consequently, Myb proteins with very similar DNA binding domains may have distinct target genes and therefore regulate distinct physiological plant processes.

Two of the six conserved residues that specify C1 interaction with its bHLH proteins (Grotewold et al., 2000) are absent in AaMYB1. It is unlikely that these two particular residues in AaMYB1 are sequence errors because they were present in independent clones obtained by PCR and from the cDNA library. However, it suggests that if AaMYB1 was an anthurium anthocyanin regulator, and required interaction with bHLH protein for activation, then it may have developed a separate mechanism of interaction that does not depend on the six residues identified in maize. In fact, while these residues are important for C1 they do not appear to define a mechanism that applies to all plants species. TT2 only has five of these residues, but in transient assays was able to complement C1 in maize

kernels when used in combination with B-Peru (Nesi et al., 2001). Performing similar tests with AaMYB1 in maize kernels could be very informative.

It may also be that AaMYB1 is involved in regulating flavones, a suggestion supported by its temporal transcript levels that peak early in spathe development (Figure 7.1), coinciding with the production profile for non-anthocyanin flavonoids in anthurium spathe tissue (Figure 4.5). It was suggested that in antirrhinum, MYB305, which is believed to play a role in regulating EBGs and the first gene of the phenylpropanoid pathway (Moyano et al., 1996), interacts with bZIP proteins and not bHLH proteins (Sablowski et al., 1994). In suggesting a similar role and molecular partner for AaMYB1, one would expect it to have a different interaction dynamic than C1 with bHLH. Consequently, the absence of two of the six conserved residues may not preclude functionality of the AaMYB1 or exclude it from involvement in the regulation of the phenylpropanoid pathway in anthurium spathe tissue.

Although AaMYB1, under the 35S promoter, was unable to complement *roseal* in particle bombardment experiments with antirrhinum petal tissue, this alone does not exclude it as an anthocyanin regulator, as maize C1 was only able to complement *roseal* when Lc was included in the particle suspension. Therefore it was suggested that, like C1, AaMYB1 could not interact with antirrhinum bHLH (DELILA) and required its own endogenous bHLH partner. However, the requirement for the endogenous bHLH protein seems to vary from plant to plant and cannot be stated as a prerequisite in transient expression assays.

The rapid browning of anthurium tissue upon wounding hampered the assay of AaMYB1 activity in endogenous tissue. However, a protocol for transient expression in coloured anthurium lines was developed providing another tool for investigations of anthurium flower colour.

While only inferences on the regulation of anthocyanin biosynthesis in anthurium spathe tissue could be gleaned from the functional characterisation studies of AaMYB1, initial analysis of the anthurium DFR promoter fragment suggested that it is recognised by anthocyanin Myb proteins. This is because conserved *cis*-elements of the MBSIIG type,

along with binding sites for maize P Myb, were found in the promoter fragment (Figure 7.3). This hinted to some commonality in the regulation of anthocyanin biosynthesis in anthurium spathe with what is observed in flowers of other species such as petunia and antirrhinum.

However, the activation of the *DFR* promoter in both wild type and *rosea1* petal tissue during particle bombardment experiments showed that ROSEA1 is not necessary for *DFR* promoter activity in the monocot anthurium. The experiments did, however, successfully replicate the results of Schwinn (1999) showing that ROSEA1 activates the promoters of antirrhinum anthocyanin genes, as the antirrhinum *DFR* promoter was only activated in wild type tissue. The results suggest that other Myb proteins besides those involved in anthocyanin biosynthesis target the conserved Myb protein recognition *cis*-elements in the anthurium *DFR* promoter.

The recent report that calcium activates the grape *DFR* promoter in cell cultures grown under white light (Gollop et al., 2002) supports the idea that other regulatory cues could be involved in anthurium *DFR* gene expression. It may be significant that calcium is used in the laboratory for preparing the particle suspension for bombardment. Whether this explains *DFR* promoter activation in *rosea1* petal tissue requires further investigation.

However, as was reported for the grape *DFR*, it appears that more than one mechanism is involved in the control of anthurium *DFR* transcript levels. The first being an anthocyanin related mechanism, mediated through Myb (and possible bHLH) interaction, as suggested by the presence of conserved *cis*-motifs in the promoter. From the evidence it does not appear that AaMYB1 is this protein. However, a role for AaMYB1 in the regulation of flavone biosynthesis cannot be ruled out. The second mechanism is a light mediated signal transduction pathway that is reflected in the diurnal rhythm of *DFR* gene expression. The third possible mechanism is a calcium signal pathway. Therefore, at any one point in time, the *DFR* expression being observed could be due to factors apart from activation by anthocyanin related Mybs. However, from the biochemical data available on anthurium

there are no significant branches of the flavonoid pathway in which *DFR* activity is implicated or linked making *DFR* seem an unusual regulatory target.

An analysis of the promoters of all the structural genes for anthocyanin biosynthesis in anthurium and comparisons with that of *DFR* may be useful in defining specific changes that may be of functional significance. Also, promoter deletion experiments to define the motif that conveys diurnal rhythms to *DFR* are possible frontiers for future work. The diurnal rhythm to *DFR* gene expression raises interesting questions about the turnover or degradation of its transcripts. mRNA destabilising elements, DST, located in the 3' untranslated region of some unstable soybean transcripts (Newman et al., 1993; McClure et al., 1989) have been found associated with specific clock controlled genes in arabidopsis (Gutierrez et al., 2002). The role of DSTs in the rhythmic patterns of *DFR* gene expression could be another research target for future work.

With a suggested role for AaMYB1 in regulating flavone and EBG levels, a comparison of its Myb binding domain or activation domains with that of Myb proteins with suggested similar functions, such as MYB308/330 or MYB305/340, may allow similarities to be identified and investigated.

Further functional analysis of AaMYB1 is necessary to determine its precise role in anthurium spathe tissue. Performing the reverse promoter/reporter experiments in anthurium tissue will be a beneficial series of experiments that will enable the regulatory systems in both plants to be compared in parallel. GFP would have to be used as the reporter gene given the difficulty with GUS assays in anthurium spathe tissue. During this study, an anthurium *DFR promoter::GFP* construct had already been designed for these experiments (unpublished work). A similar construct will be made for antirrhinum.

Transformation experiments in anthurium are a possibility, but the long regeneration time is a discouragement to such an approach. More ready information may be obtained with targeted experiments with arabidopsis carrying mutations for *PAP1, PAP2* or *TT2*. Additionally, it may be useful to use quantitative PCR and compare transcripts levels for

AaMyb1 in white and red (preferably genotypically defined) cultivars. It is also possible that AaMYB1 is involved in regulating anthocyanin biosynthesis but as part of a hierarchy of proteins, and is therefore not the direct activator. This would explain its inability to complement *rosea1*. Specific promoter binding experiments are needed to provide answers to this question. Also, promoter-binding studies in yeast or plant protoplast can be used to determine if AaMYB1 targets genes in the phenylpropanoid or flavonoid pathway or both and whether it can activate these genes independent of bHLH transcription factors.

While the scope of this work did not allow the pursuit of bHLH and possible WD repeat proteins in anthurium spathe tissue, these are also involved in regulating anthocyanin biosynthesis in other plants (de Vetten et al., 1997; Walker et al., 1999) and as such should be pursued, as this would add to our understanding of how the regulation of anthocyanin is accomplished in the modified leaf of anthurium.

In terms of industry applications from this work there are several potential benefits. One of the challenges facing the industry is to generate desirable colours in a suitable genetic background, particularly one offering resistance to disease and showing high productivity. White anthurium, are particularly in demand but while several white lines are grown, none of the cultivars show appreciable levels of resistance. With the suite of anthurium flavonoid biosynthetic cDNAs available, gene manipulation experiments, to silence CHS, F3H, DFR or ANS endogenous gene activity in a suitable cultivar, could be performed generating white anthurium in a desired genetic background.

Further to this, sense or antisense *CHS* experiments may be particularly interesting because of the possibility of generating patterned spathe phenotypes, as was observed with petunia and lisianthus (van der Krol et al., 1988, 1990; Deroles et al., 1998). Some anthurium cultivars have a mottled spathe (Kamemoto and Kuehnle, 1996) suggesting that patterned spathe phenotypes could occur in anthurium tissues. That patterned spathe pigmentation already exists in anthurium may be significant, since antisense *CHS* experiments did not generate patterned pigmentation in *Dendranthema* X *grandiflorum* (Courtney-Gutterson et

al., 1994), or gerbera (Elomaa et al., 1993), both of which traditionally lack patterned pigmented varieties.

As mentioned, an anthurium with a blue spathe is highly desired by the industry. Generating blue colours requires a suitable anthocyanin (usually delphinidins), slightly acidic to alkaline pH values and appropriate copigmentation. The anthurium spathe starts off with an advantage having flavone *C*-glycosides, which are considered the best copigments for blue colour (Asen et al., 1986) accumulating in the spathe tissue. In addition, crude measurements of cell sap from anthurium spathe tissue scored pH values of 5.5-6.2 for all anthurium lines used in this research (Appendix X1). Such values are quite sufficient for supporting blue colours (Goto and Kondo, 1991). In this regard, completing experiments to test the ability of anthurium DFR to catalyse DHM is crucial and such information will be invaluable. The cloning of an anthurium *F3'H* gene may also be useful in gene manipulation approaches towards generating blues. Therefore, prospects for developing blues, or at least the purple hues, appears promising.

There is probably an equal demand for a yellow anthurium flower as there is for blue. Chalcones are one source or yellow pigmentation in flowers (Brouillard and Dangles, 1993) Conventional approaches aimed to eliminate *CHI* activity by sense or antisense suppression in an effort to accumulate chalcones. However, this proved unsuccessful (van Blokland et al., 1993) with one possible reason being the noted spontaneous isomerisation of certain chalcones. Redirecting flavonoid biosynthesis is an alternative approach established by Davies et al. (1998). A chalcone reductase (CHR) from *Medicago sativa* was introduced into petunia resulting in the accumulation of novel 6'-deoxychalcones (as opposed to the common 6'-hydroxychalcones), and the development of pale yellow flowers. This approach presents new possibilities for flower manipulation in anthurium towards developing yellow lines. Furthermore, because 6'-deoxychalcones exhibit less spontaneous isomerisation than 6'-hydroxychalcones (Miles and Main, 1985), this approach overcomes the difficulties associated with inhibiting *CHI* activity.

In conclusion, this project has significantly advanced our understanding of the development of flower colour in anthurium, and the prospects of expanding the range of flower colours available to the local growers of anthurium in the West Indies.

References

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

Afchar, D., Cave, A., Guinaudeau, H. and Vaquette, J. (1984) Étude des réglisses d' iran. III. Flavonoides des racines de *Glycyrrhizza echinata*. L. Plantes Medicinales Phytotherapie. **18**, 170-174.

Aharoni, A., De Vos, C.H., Wein, M., Sun, Z., Greco, R., Kroon, A., Mol, J.N. and O'Connell, A.P. (2001) The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. Plant J. **28**, 319-332.

Akashi, T., Aokin, T. and Ayabe, S. (1998) Identification of a cytochrome P450 cDNA encoding (2S)-flavanone 2-hydroxylase of licorice (*Glycyrrhiza echinata* L. Fabaceae) which represents licodione synthase and flavone synthase II. FEBS Lett. **431**, 287-290.

Akashi, T., Fukuchi-Mizutani, M., Aoki, T., Ueyama, Y., Yonekura-Sakakibara, K., Tanaka, Y., Kusumi, T. and Ayabe, S. (1999) Molecular cloning and biochemical characterisation of a novel cytochrome P450, flavone synthase II, that catalyses direct conversion of flavanones to flavones. Plant Cell Physiol. **40**, 1182-1186.

Alfenito, M.R., Souer, E., Goodman, C.D., Buell, R., Mol, J., Koes, R. and Walbot, V. (1998) Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione S-transferases. Plant Cell **10**. 1135-1149.

Almeida, J., Carpenter, R., Robbins, T.P., Martin, C. and Coen, E.S. (1989) Genetic interactions underlying flower colour patterns in *Antirrhinum majus*. Genes Dev. **3,** 1758-1767.

Amaravadi, L. and King, M.W. (1994) Characterisation and expression of Xenopus c-Myb homolog. Oncogene **9**, 971-974.

Anton, I.A. and Frampton, J. (1988) Tryptophans in Myb proteins. Nature. **336**,719.

Asen, S., Stewart, R.N., Norris, K.H. and Massie, D.R. (1970) A stable blue non-metallic co-pigment complex of delphinidin and C-glycosylflavones in Prof. Blaauw iris. Phytochemistry. **9**, 619-627.

Asen, S., Grisebach, R.J., Norris, K.H. and Leonhardt, B.A. (1986) Flavonoids from *Eustoma grandiflorum* flower petals. Phytochemistry. **25**, 2509-2513.

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

Bairoch, A. (1991) PROSITE: a dictionary of sites and patterns in proteins. Nucl. Acids Res. **19**, 2241-2245.

Baker, M.E. and Blasco, R. (1992) Expansion of the mammalian 3β-hydroxysteroid dehydrogenase/plant dihydroflavonol reductase superfamily to include a bacterial cholesterol dehydrogenase, a bacterial UDP-galactose-4-epimerase, and open reading frames in vaccinia virus and fish lymphocystis disease virus. FEBS Lett. **301**, 89-93.

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

Beld, M., Martin, C., Huits, H., Stuitje, A.R. and Gerats, A.G.M. (1989) Flavonoid synthesis in *Petunia hybrida*: partial characterisation of dihydroflavonol 4-reductase genes. Plant Mol. Biol. **13**, 491-502.

Bianchi, A., Smith, S., Chong, L., Elias, P. and de Lange, T. (1997) TRF1 is a dimer and bends telomeric DNA. EMBO J. **16**, 1785-1794.

Biedenkapp, H., Borgenmeyer, U., Sippel, A.E. and Klempnauer, K.-H. (1988) Viral myb oncogene encodes a sequence-specific DNA-binding activity. Nature. **335**, 835-837.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline lysis procedure for screening recombinant plasmid DNA. Nucl. Acids Res. **7**, 1515-1522.

Blom, N., Gammeltoft, S. and Brunak, S. (1999) Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. **29**, 1351-1362.

Bloor, S. J. (1996) Blue flower colour derived from flavonol-anthocyanin copigmentation in *Ceanothus Papillosus*. Phytochemistry. **45**, 1399-1405.

Bohm, B.A. (1993) The minor flavonoids. In: Harbourne, J.B. (ed.), *The Flavonoids Advances in Research since 1986.* Chapman & Hall, London, pp. 387-440.

Borevitz, J.O., Xia, Y., Blount, J., Dixon, R.A. and Lamb, C. (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell. **12**, 2383-2393.

Boss, P.K., Davies, C. and Robinson, S.P. (1996) Expression of anthocyanin biosynthesis pathway genes in red and white grapes. Plant Mol. Biol. **32**, 565-569.

Bradley, J.M., Davies, K.M., Deroles, S.C., Bloor, S.J. and Lewis, D.H. (1998) The maize *Lc* regulatory gene up-regulates the flavonoid biosynthetic pathway of *Petunia*. Plant **J**. 13, 381-392.

Bradley, J.M., Deroles, S.C., Boase, M.R., Bloor, S., Swinny, E. and Davies, K.M. (1999) Variation in the ability of the maize *Lc* regulatory gene to upregulate flavonoid biosynthesis in heterologous systems. Plant Sci. 140, 31-39.

Braun, E.L. and Grotewold, E. (1999) Newly discovered plant c-myb-like genes rewrite the evolution of the plant myb gene family. Plant Physiol. **121**, 21-24.

Britsch, L. (1990) Purification and characterisation of flavone synthase I, a 2-oxoglutarate-dependent desaturase. Arch. Biochem. Biophy. **276**, 348-354.

Britsch, L., Dedio, J., Saedler, H. and Forkmann, G. (1993) Molecular characterisation of flavanone 3β-hydroxylases-consensus sequence, comparison with related enzymes and the role of conserved histidine residues. Euro. J. Biochem. **217**,745-754.

Brouillard, R. (1981) Origin of the exceptional colour stability of the *Zerbrina* anthocyanin. Phytochem. **20**, 143-145.

Brouillard, R. and Dangles, O. (1993) Flavonoids and flower colour. In: Harborne, J.B. (ed.), *The Flavonoids: Advances in Research Since 1986.* Chapman & Hall, London, pp. 565-587.

Brugliera, F., Rewell-B, G., Holton, T.A. and Mason, J.G. (1999) Isolation and characterisation of a flavonoid 3' hydroxylase cDNA corresponding to the *Ht1* locus of *Petunia hybrida*. Plant J. **19**, 441-451.

Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. Cell. **56**, 549-561.

Cadwell, M.M., Robberecht, R. and Flint, S.D. (1983) Internal filters: Prospects for UV-acclimation in higher plants. Physiol. Plant. **58**, 445-450.

Campbell, D.H. (1900) Studies on the Araceae I. Ann. Bot. **14,** 2-25.

Carre, I.A. and Kim, J.Y. (2002) MYB transcription factors in the *Arabidopsis* circadian clock. J. Exp. Bot. **53,** 1551-1557.

Charest, P.J., Xue, B., Devantier, Y. and Rutledge, R. (1994) Activation of the corn *Bronze2* promoter by the corn Myb-related C1 gene or the C1 black spruce homologue (MBF1) in larch and spruce tissues via microparticle bombardment. Abstract 62, 4th Annual Queenstown Molecular Biology Meeting, August 14-19, Queenstown, New Zealand.

Chen, S.M. and Coe, E.H. Jr. (1977) Control of anthocyanin synthesis by the *C-* locus in maize. Biochem. Genet. **15,** 333-346.

Chen, F. and Kuehnle, A.R. (1996) Obtaining transgenic anthurium through Agrobacterium-mediated transformation of etiolated internodes. J. Amer. Soc. Hort. Sci. **121,** 47-51.

Christensen, O.V. (1971) Morphological studies on the growth and formation of *Anthurium scherzerianum* Schott and *Anthurium andraeanum* Lind. Tiddsskrift for Planteavl. **75,**793-798.

Christie, J.M. and Jenkins, G.I. (1996) Distinct UV-B and UV-A/blue light signal tranduction pathways induce chalcone synthase gene expression in Arabidopsis cells. Plant Cell. **8,** 1555-1567.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. Proc. Natl. Acad. Sci. USA. **81,** 1991-1995.

Clement, J.S., Mabry, T.J., Wyler, H. and Dreiding, A.S. (1994) Chemical review and evolutionary significance of the betalains. In: Behnke, H.-D. and Mabry, T.J. (eds.), *Caryophyllales; Evolution and Systematics*, Springer-Verlag, Berlin, pp. 247-261.

Cline, J., Braman, J.C. and Hogrefe, H.H. (1996) PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. Nucl Acids Res. 24, 3546-3551.

Cocciolone, S.M. and Cone, K.C. (1993) Pl-Bh, an anthocyanin regulatory gene of maize that leads to variegated pigmentation. Genetics. **135**, 575-588.

Coe, E.H., Mc Cormick, S.M. and Modena, S.A. (1981) White pollen in maize. J. Heredity. **72**, 318-320.

Coe, E.H., Jr. (1985) Phenotypes in corn: control of pathways by alleles, time and place. In: Freeling, M. (ed.), *Plant Genetics*. Alan R. Liss, Inc., New York, pp. 509-521.

Cone, K.C., Burr, F.A. and Burr, B. (1986) Molecular analysis of the maize regulatory locus *C1*. Proc. Natl. Acad. Sci. USA. **83**, 9631-9635.

Cone, K.C., Cocciolone, S.M., Burr, F.A. and Burr, B. (1993) *Maize* anthocyanin regulatory gene *pl* is a duplicate of *c1* that functions in the plant. Plant Cell. **5**, 1795-1805.

Cosma, M.P., Tanaka, T. and Nasmyth, K. (1999) Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle-and developmentally regulated promoter. Cell. **97**, 299-311.

Courtney-Gutterson, N., Napoli, C., Lemieux, C., Morgan, A., Firoozabady, E. and Robinson, K.E.P. (1994) Modification of flower color in florist's *Chrysanthemum*: production of a white-flowering variety through molecular genetics. Bio/Technology. **12**, 268-271.

Davies, K.M., Bradley, J.M., Schwinn, K.E., Markham, K.R. and Podivinsky, E. (1993) Flavonoid biosynthesis in flower petals of five lines of lisianthus (*Eustoma grandiflorum* Grise). Plant Sci. **95**, 67-77.

Davies, K.M. and Schwinn, K.E. (1997) Flower colour. In: Geneve, R.L., Preece, J.E. and Merkle, S.A. (eds.), *Biotechnology of Ornamental Plants.* Biotechnology in Agriculture Series. No. 16, pp 259-294.

Davies, K.M., Bloor, S.J., Spiller, G.B. and Deroles, S.C. (1998) Production of yellow colour in flowers: redirection of flavonoid biosynthesis in *Petunia.* Plant J. **13**, 259-266.

Davies, J., Badiani, P. and Weston, K. (1999) Cooperation of Myb and Myc proteins in T-cell lymphomagenesis. Oncogene. **18**, 3643-3647.

Davis, S.J. and Millar, A.J. (2001) Watching the hands of *Arabidopsis* biological clock. Genome Biol. **2**, 1008.1-1008.4.

Dawson, R.M.C., Elliot, D.C., Elliot, W.H. and Jones. K.M. (1986) Data for biochemical research. Oxford Science publication. pp. 380-381.

Debeaujon, I., Peeters, A.J.M., Kloosterziel-Leon, M.K. and Koornneef, M. (2001) The TRANSPARENT TESTA12 gene of arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. Plant Cell. **13**, 853-871.

de Majnik, J., Tanner, G.J., Joseph, R.G., Larkin, P.J., Weinman, J.J., Djordjevic, M.A. and Rolfe, B.G. (1998) Transient expression of maize anthocyanin regulatory genes influences anthocyanin production in white clover and peas. Aust. J. Plant Physiol. **25**, 335-343.

Deboo, G.B., Albertsen, M.C. and Taylor, L.P. (1995) Flavonone 3-hyroxylase transcripts and flavonol accumulation are temporally coordinated in maize anthers. Plant J. **7**, 703-713.

Deroles, S.C., Bradley, J.M., Schwinn, K.E., Markham, K.R., Bloor, S., Manson, D.G. and Davies, K.M. (1998) An antisense chalcone synthase cDNA leads to novel colour patterns in lisianthus (*Eustoma grandiflorum*) flowers. Mol. Breeding. **4**, 59-66.

de Vetten, N., Quattrocchio, F., Mol, J. and Koes, R. (1997) The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev. **11**, 1422-1434.

de Vlaming, P., Cornu, A., Farcy, E., Gerats, A.G.M., Maizonnier, D., Wiering, H. and Wijsman, H.J.W. (1984) *Petunia hybrida*: A short description of the action of 91 genes, their origin and their map location. Plant Mol. Biol. Rep. **2**, 21-42.

Diesperger, H., Muller, C.R. and Sandermann, H. Jr. (1974) Rapid isolation of plant microsomal fraction by Mg^{2+} precipitation. FEBS. lett. **43**, 155-158.

Dixon, R.A. and Lamb, C.J. (1990) Molecular communications in interactions between plants and microbial pathogens. Ann. Rev. Plant Physiol. Plant Mol. Biol. **41**, 339-367.

Dooner, H.K. (1983) Co-ordinate genetic regulation of flavonoid biosynthetic enzymes in maize. Mol. Gen. Genet. **189**, 136-141.

Dooner, H.K. and Nelson, O.E. (1977) Genetic control of UDP-glucose: flavonol-O-glucosyltransferase in the endosperm of maize. Biochem. Genet. **15**, 509-515.

Eckermann, S., Schroder, G., Schmidt, J., Strack, D., Edrada, R.A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpelainen, I., Proksch, P., Teeri, T.H. and Schroder, J. (1998) New pathway to polyketides in plants. Nature. **396**, 387-390.

Ellenberger, T. (1994) Getting a grip on DNA recognition: structures of the basic region leucine zipper, and the basic region helix-loop-helix DNA-binding domains. Curr. Opin. Struct. Biol. **4**, 12-21.

Elomaa, P., Honkanen, J., Puska, R., Seppänen, P., Helariutta, Y., Mehto, M., Kotilainen, M., Nevalainen, L. and Teeri, T.H. (1993) Agrobacterium-mediated transfer of antisense chalcone synthase cDNA to *Gerbera hybrida* inhibits flower pigmentation. Bio/Technology. **11**, 508-511.

Elomaa, P., Helariutta, Y., Griesbach, R.J., Kotilainen, M., Seppänen, P. and Teeri, T.H. (1995) Transgene inactivation in *Petunia hybrida* influenced by the properties of the foreign gene. Mol. Gen. Genet. **248**, 649-656.

Fedoroff, N., Furtek, D. and Nelson, O.E. (1984) Cloning of the *bronze* locus in maize by a simple and generalizable procedure using the transposable element Activator (Ac). Proc. Nat. Acad. Sci. USA. **81**, 3825-3829.

Feldbrügge, M., Sprenger, M., Hahlbrock, K. and Weisshaar, B. (1997) PcMYB1, a novel plant protein containing a DNA-binding domain with one MYB repeat, interacts in vivo with a light regulatory promoter unit. Plant J. **11**, 1079-1093.

Ferré-D' Amaré, A.R., Prendergast, G.C., Ziff, E.B. and Burley, S.K. (1993) Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. Nature. **363**, 38-45.

Ferrer, J.L., Jez, J.M., Bowman, M.E., Dixon, R.A. and Noel, J.P. (1999) Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat. Struct. Biol. **6**, 775-784.

Finer, J, J., Vain, P., Jones, M.W. and McMullen, M.D. (1992) Development of the particle inflow gun for DNA delivery to plant cells. Plant Cell Rep. **11**, 323-328.

Fliegmann, J., Schröder, G., Schanz, S., Britsch, L. and Schröder, J. (1992). Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. Plant Mol. Biol. **18**, 489-503.

Foos, G., Grimm, S., Klempnauer, K.H. (1992) Functional antagonism between members of the Myb family: b-MYB inhibits v-MYB-induced gene activation. EMBO J. **11**, 4619-4629.

Forkmann, G. (1991) Flavonoids as flower pigments: The formation of the natural spectrum and its extension by genetic engineering. Plant Breeding. **106**, 1-26.

Forkmann, G. and Dangelmayr, B. (1980) Genetic control of chalcone isomerase activity in flowers of Dianthus caryophyllus. Biochem. Genet. **18**, 519-527.

Forkmann, G. and Stotz, G. (1981) Genetic control of flavanone 3-hydroxylase activity and flavonoid 3´-hydroxylase activity in *Antirrhinum majus* (Snapdragon). Z. Naturforsch. **35c**, 411-416.

Forkmann, G. and Ruhnau, B. (1987) Distinct substrate specificty of dihydroflavonol 4-reductase from flowers of *Petunia hybrida*. Z. Naturforsch. **42c**, 1146-1148.

Framptom, J., Leutz, A., Gobson, T.J. and Graf, T. (1989) DNA binding domain ancestry. Nature. **342**, 134.

Franken, P., Schrell, S., Peterson, P.A., Saedler, H. and Wienand, U. (1994) Molecular analysis of protein domain function encoded by the myb-homologous maize genes *C1*, *Zm* 1 and *Zm* 38. Plant J. **6**, 21-30.

Fritsch, H. and Griesbach, H. (1975) Biosynthesis of cyanidin in cell cultures of *Happlopappus gracillis.* Phytochemisty. **14**, 2437

Froemel, S., deVlaming, P., Stotz, G., Weiring, H., Forkmann, G. and Schram, A.W. (1985) Genetic and biochemical studies on the conversion of flavanones to dihydroflavonols in flowers of *Petunia hybrida*. Theor. Appl. Genet. **70,** 561-568.

Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA. **85,** 8998-9002.

Frohnmeyer, H., Bowler, C. and Schafer, E. (1997) Evidence for some signal transduction elements involved in UV-light-dependent responses in parsley protoplasts. J. Exper. Bot. **48,** 739-750.

Frohnmeyer, H., Bowler, C., Zhu, J.K., Yamagata, H., Schafer, E. and Chua, N.H. (1998) Different roles for calcium and calmodulin in phytochrome and UV-regulated expression of chalcone synthase. Plant J. **13,** 763-772.

Galego, L and Almeida, J. (2002) Role of DIVARICATA in the control of dorsoventral asymmetry in Antirrhinum flowers. Genes Dev. **1,** 880-891.

Gerats, A.G.M., Vlaming, P., Doodeman, M., Al, B. and Schram, A.W. (1982) Genetic control of the conversion of dihydroflavonols into flavonols and anthocyanins in flowers of *Petunia hybrida.* Planta. **155,** 364 368.

Gill, G. and Ptashne, M. (1987) Mutants of GAL4 protein altered in an activation function. Cell. **51,** 121-126.

Gill, G., Sadowski, I. and Ptashne, M. (1990) Mutations that increase the activity of a transcriptional activator in yeast and mammalian cells. Proc. Natl. Acad. Sci. USA. **87,** 2127-2131.

Giniger, E. and Ptashne, M. (1987) Transcription in yeast activated by a putative amphipathic α-helix linked to a DNA binding unit. Nature. **330**, 670-672.

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

Goff, S.A., Cone, K.C. and Fromm, M.E. (1991) Functional analysis of the maize C1 transcriptional activator: Comparison of wild-type and dominant inhibitor proteins. Genes Dev. **5**, 298-309.

Goff, S.A., Cone, K.C. and Chandler, V.L. (1992) Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. Genes Dev. **6**, 864-875.

Gollop, R., Even, S., Tsolova-C. V. and Perl, A. (2002) Expression of the grape dihydroflavonol reductase gene and analysis of its promoter region. J. Experi. Bot. **53**, 1397-1409.

Gong, Z.Z., Yamazaki, M., Sugiyama, M., Tanaka, Y. and Saito, K. (1997) Cloning and molecular analysis of structural genes involved in anthocyanin biosynthesis and expressed in a forma-specific manner in *Perilla frutescens*. Plant Mol. Biol. **35**, 915-927.

Gong Z, Z., Yamazaki, M. and Saito, K. (1999a) A light-inducible Myb-like gene that is specifically expressed in red *Perilla frutescens* and presumably acts as a determining factor of the anthocyanin forma. Mol. Gen. Genet. **262**, 65-72.

Gong Z, Z., Yamagishi, E., Yamazaki, M. and Saito, K. (1999b) A constitutively expressed Myc-like gene involved in anthocyanin biosynthesis from *Perilla frutescens*: molecular characterisation, heterologous expression in transgenic plants and transactivation in yeast cells. Plant Mol. Biol. **41**, 33-44.

Goodrich, J., Carpenter, R. and Coen, E. (1992) A common gene regulates pigmentation pattern in diverse plant species. Cell. **68,** 955-964.

Goto, T. and Kondo, T. (1991) Structure and molecular stacking of anthocyanins-flower color variation. Angew. Chem. Int. Ed. Engl. **30,** 17-33.

Grew, N. (1682) The anatomy of plants, with an idea of a philosophical history of plants, and several other lectures, read before the Royal Society, London. In: Bohm, B. (ed.), *Introduction to flavonoi*ds: *Chemistry and biochemistry of Organic Natural products.* Harwood Academic Publishers, The Netherlands. pp. 1-4

Grotewold, E., Drummond, B.J., Bowen, B. and Peterson, T. (1994) The myb-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell. **76,** 543-553.

Grotewold, E. and Peterson, T. (1994) Isolation and characterisation of a maize gene encoding chalcone flavanone isomerase. Mol. Gen. Genet. **242,** 1-8.

Grotewold, E., Chamberlin, M., Snook, M., Siame, B., Butler, L., Swenson, J., Maddock, S., St. Clair, G. and Bowen, B. (1998) Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. Plant Cell. **10,** 721-740.

Grotewold, E., Sainz, M.B., Tagliani, L., Hernandez, J.M., Bowen, B. and Chandler, V.L., (2000) Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R. Proc.Natl. Acad. Sci. USA. **97,** 13579-13584.

Gubler, F., Kalla, R., Roberts, J.K. and Jacobsen, J.V. (1995) Gibberellin-regulated expression of a myb gene in barley aleurone cells: Evidence for Myb transactivation of a high-pl α-amylase gene promoter. Plant Cell. **7,** 1879-1891.

Guo, K., Anjard, C., Harwood, A., Kim, H-J., Newell, P.C. and Gross, J.D. (1999) A myb-related protein required for culmination in Dictyostelium. Development. **126**, 2813-2822.

Gutierrez, R.A., Ewing, R.M., Cherry, J.M. and Green, P.J. (2002) Identification of unstable transcripts in arabidopsis by cDNA microarray analysis: Rapid decay is associated with a group of touch-and specific clock-controlled genes. Proc.Natl. Acad. Sci. USA. **99**, 11513-11518.

Hahlbrock, K. and Griesbach, H. (1975) Biosynthesis of the flavonoids. In: Harborne, J.B., Mabry, T.J. and Mabry, H. (eds.), *The Flavonoids.* Chapman and Hall, London, pp 866-915.

Hanson, K,R. and Havir, E.A.(1972) The enzymatic elimination of ammonia. In: Boyer P,D. (ed.), *The enzymes.* Academic Press, New York, pp 75-166.

Harborne, J.B. (1976) Functions of flavonoids in plants. In: Goodwin, T.W. (ed.), *Chemistry and Biochemistry of Plant Pigments*, Vol. 1. Academic Press, New York, pp. 736-778.

Harborne, J.B. (1988) The flavonoids: recent advances. In: Goodwin, T.W. (ed.), *Plant Pigments*, Academic Press, London, pp. 299-343.

Harmer, S.L., Hogenesch, L.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A. and Kays, S.A. (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science. **290**, 2110-2113.

Harrison, B.J. and Stickland, R.G. (1978) Precursors and genetic control of pigmentation. 4. Hydroxylation and methoxylation stages in anthocyanidin biosynthesis. Heredity. **40**, 127-132.

Hashim, M.F., Hakamatsuka, T., Ebizuka, Y. and Sankawa, U. (1990) Reaction mechanism for oxidative rearragement of flavanone in isoflavone biosynthesis. FEBS Lett. **271**, 219-222.

Helariutta, Y., Elomaa, P., Kotilainen, M., Seppanen, P. and Teeri, T.H. (1993) Cloning of cDNA coding for dihydroflavonol-4-reductase (DFR) and characterisation of DFR expression in the corollas of *Gerbera hybrida* var. Regina (Compositae). Plant Mol. Biol. **22**, 183 193.

Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R.J., Schroder, J., and Teeri., T.H. (1995) Chalcone synthase-like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (asteraceae). Plant Mol Biol. **28**, 47-60.

Heller, W. and Hahlbrock, K. (1980) Highly purified "flavanone synthase" from parsley catalyses the formation of naringenin chalcone. Arch. Biochem. Biophys. **200**, 617-619.

Heller, W., Forkmann, G., Britsch, L. and Grisebach, H. (1985) Enzymatic reduction of (+) -dihydroflavonols to flavan-3,4-cis-diols with flower extracts from *Matthiola incana* and its role in anthocyanin biosynthesis. Planta. **165**, 284 287.

Heller, W. and Forkmann, G. (1988) Biosynthesis of flavonoids. In: Harborne, J.B. (ed.), *The Flavonoids: Advances in Research Since 1980*. Chapman & Hall, London, pp. 399-425.

Heller, W. and Forkmann, G. (1993) Biosynthesis of flavonoids. In: Harborne, J.B. (ed.), *The Flavonoids: Advances in Research Since 1986*. Chapman & Hall, London, pp. 499-536.

Higaki, T., Rasmusan H, P. and Carpenter, W. J. (1984) A study of some morphological and anatomical aspects of *Anthurium andraeanum* Lind. HITAHR Research Series 030, University of Hawaii. pp 1-12.

Higgins, D. G. and Sharp, P.M. (1989) Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS. **5**, 151-153.

Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database. Nucl. Acids Res. **27**, 297-300.

Holton, T.A., Brugliera, F. and Tanaka, Y. (1993) Cloning and expression of flavonol synthase from *Petunia hybrida*. Plant J. **4**, 1003-1010.

Holton, T.A. and Tanaka, Y. (1994) Blue roses-a pigment of our imagination? TIBTECH. **12**, 40-42.

Howe, K.M., Reakes, C.F. and Watson, R.J. (1990) Characterisation of the sequence-specific interaction of mouse c-myb protein with DNA. EMBO J. **9**, 161-169.

Hrazdina, H. (1982) Anthocyanins. In: Harborne, J.B. and Marbray, T.J. (eds.), *The Flavonoids: Advances in Research since 1980*. Chapman and Hall, London. pp166-189.

Hughes, J. and Hughes, M.A. (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava. DNA Seq. **5**, 41-49.

Huits, S.M., Gerats, A.G.M., Kreike, M.M., Mol, J.M.N. and Koes, R.E. (1994) Genetic control of dihydroflavonol 4-reductase gene expression in *Petunia hybrida*. Plant J. **6**, 295-310.

Inostroza, J.A., Mermelstein, F.H., Ha, I., Lane, W.S. and Reinberg, D. (1992) Drl, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. Cell. **70,** 477-489.

Inoue, H., Nojima, H. and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. Gene. **96,** 23-28.

Ito, M., Araki, S., Matsunaga, S., Itoh, T., Nishihama, R., Machida, Y., Doonan, J.H. and Watanabe, A. (2001) G2/M-phase-specific transcription during the plant cell cycle is mediated by c-Myb-like transcription factors. Plant Cell. **13,**1891-905.

Itturriaga, G., Leyns, L., Villegas, A., Gharaibeh, R., Salamini, F. and Bartels, D. (1996) A family of novel myb-related genes from the resurrection plant *Craterostigma plantaginuem* are specifically expressed in callus and roots in response to ABA or desiccation. Plant Mol. Biol. **32,** 707-716.

Iwata, R,Y., Tang, C.S. and Kamemoto, H. (1979) Anthocyanins of *Anthurium andraeanum* Lind. J. Amer. Soc. Hort. Sci. **104,** 464-466.

Iwata, R.Y., Tang, C.S. and Kamemoto, H. (1985) Concentration of anthocyanins affecting spathe colour in Anthuriums. J. Amer. Soc. Hort. Sci. **110,** 383-385.

Jackson, D., Culianez-Macia, F., Prescott, A.G., Roberts, K. and Martin, C. (1991) Expression patterns of Myb genes from Antirrhinum flowers. Plant Cell. **3,** 115-125.

Jackson, D., Roberts, K. and Martin, C. (1992) Temporal and spatial control of expression of anthocyanin biosynthetic genes in developing flowers of *Antirrhinum majus*. Plant J. **2,** 425-434.

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

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B. and Martin, C. (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. EMBO J. **15**, 6150-6161.

Johnson, E.T., Yi, H., Shin, B., Oh, B.J., Cheong, H. and Choi, G. (1999) *Cymbidium hybrida* dihydroflavonol 4-reductase does not efficiently reduce dihydrokaempferol to produce orange pelargonidin-type anthocyanins. Plant J. 19, 81-85.

Johnson, E.T., Ryu, S., Yi, H., Shin, B., Cheong, H. and Choi, G. (2001) Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. Plant J. **25**, 325-33.

Jonsson, L.M.V., Aarsman, M.E.G., Schram, A.W. and Bennink, G.S.H. (1982) Methylation of anthocyanins by cell-free extracts of flower buds of *Petunia hybrida*. Phytochem. **21**, 2457-2459.

Jonsson, L.M.V., Aarsman, M.E.G., Poulton, J.E. and Schram, A.W. (1984a) Properties and genetic control of four methyltransferases involved in methylation of anthocyanins in flowers of *Petunia hybrida*. Planta **160**, 174-179.

Jonsson, L.M.V., Aarsman, M.E.G., de Vlaming, P. and Schram, A.W. (1984b) On the origin of anthocyanin methyltransferase isozymes of *Petunia hybrida* and their role in regulation of anthocyanin methylation. Theor. Appl. Genet. **68**, 459-463.

Joshi, C.P. (1987b) Putative polyadenylation signals in nuclear genes of higher plants; a compilation and analysis. Nucl. Acid Res. **15**, 9627-9640.

Joshi, C.P. (1987a) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucl. Acid Res. **15**, 6643-6653.

Kamemoto, H. and Nakasone, H.Y. (1963) Evaluation and improvemnet of Anthurium clones. Hawaii Agr. Expt. Sta., Univ. of Hawaii. Tech. Bul. 58.

Kamemoto, H., Iwata, R.Y. and Marutani, M. (1988) Genetics of the major spathe colours in Anthuriums. HITAHR Research Series 056, College of Tropical Agriculture and Human Resources. University of Hawaii, pp 1-11.

Kamemoto, H. and Kuehnle, A.R. (1996) Breeding anthuriums in Hawaii. University of Hawaii Press. pp 37-75.

Kay, Q.O.N., Daoud, H.S. and Stirton, C.H. (1981) Pigment distribution, light reflection and cell structure in petals. Botanical Journal of the Linnaean Society **83**, 57-84.

Kellogg, D.E., Rybalkin, I., Chen, S., Mukhamedova, N., Vlasik, T., Siebert, P. and Chenchik, A. (1994) *Taq*Start Antibody: Hotstart PCR facilitated by neutralising monoclonal antibody directed against *Taq* DNA polymerase. Bio Tequniques. **16,** 1134-1137.

Kerscher, F. and Franz, G. (1987) Biosynthesis of vitexin and isovitexin: enzymatic synthesis of the C-glucosylflavones vitexin and isovitexin with an enzyme preparation from *Fagopyrum esculentum* M. seedlings. Z. Naturforsch. **42c,** 519-524.

Kimura, H., Tao, Y., Roeder, R.G. and Cook, P.R. (1999) Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: Little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Mol. Cell. Biol. **19,** 5383-5392.

King, G. and Davies, K. (1992) Identification, cDNA cloning, and analysis of mRNAs having altered expression in tips of harvetsed asparagus spears. Plant Physiol. **100,** 1161-1169.

Kirik, V. and Bäumlein, H. (1996) A novel leaf-specific Myb-related protein with a single binding repeat. Gene. **183,** 109-113.

Klempnauer, K-H., Gonda, T.J. and Bishop, J.M. (1982) Nucleotide sequence of the retroviral leukemia gene v-Myb and its cellular progenitor c-Myb: the architecture of a tranduced oncogene. Cell. **31**, 453-463.

Kochs, G. and Griesbach, H. (1986) Enzymic synthesis of isoflavones. Eur. J. Biochem. **155**, 311-318.

Koes, R.E., Spelt, C.E., Mol, J.N.M. and Gerats, A.G.M. (1987) The chalcone synthase multigene family of *Petunia hybrida* (V30): Sequence homology, chromosomal location and evolutionary aspects. Plant Mol. Biol. **10**, 375-385.

Koes, R.E., Spelt, C.E. and Mol, J.N.M. (1989) The chalcone synthase multigene family of *Petunia hybrida* (V30): differential, light-regulated expression during flower development and UV light induction. Plant Mol. Biol. **12**, 213-225.

Koleske, A.J. and Young, R.A. (1994) An RNA Polymerase II holoenzyme responsive to activators. Nature. **368**, 466-469.

Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R.C., Petroni, K., Urzainqui, A., Bevan, M., Martin, C., Smeekens, S., Tonelli, C., Paz-Ares, J. and Weisshaar, B. (1998) Towards functional characterisation of the members of the R2R3-Myb gene family from *Arabidopsis thaliana*. Plant J. **16**, 263-276.

Kranz, H., Scholz, K. and Weisshar, B. (2000) c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. Plant J. **21**, 231-235.

Krebs, J.E., Kuo, M.H., Allis, C.D. and Peterson, C.L. (1999) Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. Genes and Dev. **13**, 1412-1421.

Kreps, J.A. and Kay, S.A. (1997) Coordination of plant metabolism and development by the circadian clock. Plant Cell. **9**, 1235-1244.

Kreuzaler, F. and Hahlbrock, K. (1975) Enzymic synthesis of an aromatic ring from acetate units. Eur. J. Biochem. **56**, 205-213.

Kristiansen, K.N. and Rohde, W. (1991) Structure of the *Hordeum vulgare* gene encoding dihydroflavonol4-reductase and molecular analysis of ant18 mutants blocked in flavonoid synthesis. Mol. Gen. Genet. **230**, 49-59.

Kuehnle, A.R., Kamemoto, H., Amore, T.D., Kunisaki, J.T., Lichty, J.S. and Uchida J.Y. (2000) White Lady1 Anthurium. Univ. of Hawaii CTAHR Coop. Ext. Ser. Series NPH-A-4.

Kuehnle, A.R., Amore, T.D., Kamemoto, H., Kunisaki, J.T., Lichty, J.S.and J.Y. Uchida. (2002) 'Hilo Moon' and 'Hokuloa' Anthurium. Univ. of Hawaii CTAHR Coop. Ext. Ser. Series NPH-A-5.

Kuhn, B., Forkmann, G. and Seyffert, W. (1978) Genetic control of chalcone-flavanone isomerase activity in *Callistephus chinensis*. Planta. **138**, 199-203.

Landry, L.G., Chapple, C.C., and Last, R. (1995) Arabidopsis mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol. **109**, 1159-1166.

Leech, M.J., Kammerer, W., Cove, D.J., Martin, C. and Wang, T.L. (1993) Expression of Myb-related genes in the moss, *Physcomitrella patens*. Plant J. **3**, 51-61.

Lemon, B and Tjian, R. (2000) Orchestrated response: a symphony of transcription factors for gene control. Genes Dev. **14**, 2551-2569.

Liew, C.-F., Loh, C.-S., Goh, C.-J. and Lim, S.-H. (1998) The isolation, molecular characterisation and expression of dihydroflavonol 4-reductase cDNA in the orchid, *Bromheadia finlaysoniana*. Plant Sci. **135**, 161-169.

Lin, Y. and Green, M.R. (1991) Mechanism of action of an acidic transcriptional activator in vitro. Cell. **64**, 971-981.

Lin, Q., Hamilton, W.D.O. and Merryweather, A. (1996) Cloning and initial characterisation of 14 Myb-related cDNAs from tomato (*Lycopersicon esculentum* cv. Ailsa Craig). Plant Mol. Biol. **30**, 1009-1020.

Linsmaier, E.M. and Skoog, F. (1965) Organic growth factor requirements of tobacco tissue cultures. Physiologia Plantarum. **18** 100-127.

Lipsick, J.S. (1996) One billion years of Myb. Oncogene. **13**, 223-235.

Lloyd, A.M., Walbot, V. and Davis, R.W. (1992) Arabidopsis and Nicotiana anthocyanin production activated by maize regulators R and C1. Science. **258**, 1773-1775.

Loomis, W.D. (1974) Overcoming problems of phenolics and quinones in the isolation of plant enzymes and organelles. Meth. Enzymol. **31**, 528-545.

Lu, C-A., Ho, D.T-h., Ho, S-H. and Yu, M-S. (2002) Three novel Myb proteins with one DNA binding repeat mediate sugar and hormone regulation of α-amylase gene expression. Plant Cell. **14**, 1963-1980.

Ludwig, S.R., Habera, L.F., Dellaporta, S.L. and Wessler, S.R. (1989) Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region. Proc. Natl. Acad. Sci. USA. **86**, 7092-7096.

Ludwig, S.R. and Wessler, S.R. (1990) Maize R gene family: Tissue-specific helix-loop-helix proteins. Cell. **62**, 849-851.

Lukacin, R. and Britsch, L. (1997) Identification of strictly conserved histidine and arginine residues as part of the active site in *Petunia hybrida* flavanone 3β-hydroxylase. Euro. J. Biochem. **249**,748-757.

Lukacin, R., Groning, I., Pieper, U. and Matern, U. (2000) Site-directed mutagenesis of the active site serine 290 in flavanone 3 beta-hydroxylase from *Petunia hybrida*. Euro. J. Biochem. **267**, 853-860.

Lüscher, B., Christenson, E., Litchfield, D.W., Krebs, E.G. and Eisenman, R.M. (1990) MYB DNA binding inhibited by phosphorylation at a site during oncogenic activation. Nature. **344**, 517-522.

Ma, J. and Ptashne, M. (1987) Deletion analysis of GAL4 defines two transcriptional activating segments. Cell. **48**, 847-853.

Markham, K.R., Tanner G.J., Caasi-Lit, M., Whitecross, M.I., Nayudu, M. and Mitchell, K.A. (1998) Possible protective role for 3',4'-dihydroxyflavones induced by enhanced UV-B in a UV-tolerant rice cultivar. Phytochemistry. **49**, 1913-1919.

Marocco, A., Wissenbach, M., Becker, D., Paz-Ares, J., Saedler, H., Salamini, F. and Rohde, W. (1989) Multiple genes are transcribed in Hordeum vulgare and Zea mays that carry the DNA-binding domain of the Myb oncoproteins. Mol. Gen. Genet. **216**, 183-187.

Marrs, K.A., Alfenito, M.R., Lloyd, A.M. and Walbot, V. (1995) A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. Nature. **375**, 397-400.

Martens, S and Forkmann, G. (1999) Cloning and expression of flavone synthase II from *Gerbera hybrida*. Plant J. **20**, 611-618.

Martin, C., Prescott, A., Mackay, S., Bartlett, J. and Vrijlandt, E. (1991) Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. Plant J. **1**, 37-49.

Martin, C. and Gerats, T. (1993) Control of pigment biosynthesis genes during petal development. Plant Cell. **5**, 1253-1264.
Martin, C. and Paz-Ares, J. (1997) Myb transcription factors in plants. TIG. **13**, 67-73.

Marutani, M., Tang, C.S., Paul, R. and Kamemoto, H. (1987) Anthocyanins in the lavendar anthurium. Hort. Sci. **22**, 620-622.

McClung, C.R. (2000) Circadian rhythms in plants: a millennial view. Physiologia. Plantarum. **109**, 359-371.

McClung, C.R. (2001) Circadian rhythms in plants. Ann. Rev.Plant Physiol. and Plant Mol. Biol. **52**,139- 162.

McClure, B.A., Hagen, G., Brown, C.S., Gee, M.A. and Guilfoyle, T.J. (1989) Transcription organisation and sequences of an auxin related gene cluster in soybean. Plant Cell. **1**, 229-239.

Menting, J.G.T., Scopes, R.K. and Stevenson, T.W. (1994) Characterisation of flavonoid 3'5' -hydroxylase in microsomal membarne fraction of *Petunia hybrida* flowers. Plant Physiol. **106**, 633-642.

Meshi, T. and Iwabuchi, M. (1995) Plant transcription factors. Plant Cell Physiol. **36**, 1405-1420.

Meyer, P., Heidmann, I., Forkmann, G. and Saedler, H. (1987) A new petunia flower colour generated by transformation of a mutant with a maize gene. Nature. **330**, 677-678.

Michael, T.P. and McClung C.R. (2002). Phase-specific circadian clock regulatory elements in *Arabidopsis thaliana.* Plant Physiol. **130**, 627-638.

Middleton, E. Jr. and Kandaswami, C. (1992) Effects of flavonoids on immune and inflammatory cell functions. Biochem. Pharm. **43**,1167-1179.

Middleton, E. Jr., Kandaswami, C. and Theoharides, T.C. (2000) The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease and cancer. Pharmacol. Rev. **52**, 673-751.

Miles, C.O. and Main, L. (1985) Kinetics and Mechanism of the cyclisation of 2',6'-dihydroxy-4,4'-dimethoxychalcone; influence of the 6'-hydroxyl group on the rate of cyclisation under neutral conditions. J. Chem. Soc. Perkin Trans. II. 1639-1642.

Millar, A.J. and Kay, S.A. (1996) Integration of circadian phototransduction pathways in the network controlling CAB gene transcription in Arabidopsis. Proc. Natl. Acad. Sci. USA. **93**,15491-15496.

Millar, A.J. and Kay, S.A. (1997) The genetics of phototransduction and circadian rhythms in Arabidopsis. Bioessays. **19**, 209-214.

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H. R., Carre I, A. and Coupland, G. (2002) LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis. Dev. Cell. **2**, 629-641.

Mo, Y., Nagel, C. and Taylor, L.P. (1992) Biochemical complementation of chalcone synthase mutants defines a role for flavonols in functional pollen. Proc. Natl. Acad. Sci. USA. **89**, 7213-7217.

Mol, J.N.M., Robbins, M.P., Dixon, R.A. and Veltkamp, E. (1985) Spontaneous and enzymic rearrangement of naringenin chalcone to flavanone. Phytochemistry. **24**, 2267-2269.

Mol, J., Grotewold, E., and Koes, R. (1998) How genes paint flowers and seeds. Trends Plant Sci. **3**, 212-217.

Mooney, M., Desnos, T., Harrison, K., Jones, J., Carpenter, R. and Coen, E. (1995) Altered regulation of tomato and tobacco pigmentation genes caused by the *delila* gene of *Antirrhinum*. Plant J. **7**, 333-339.

Mori, A., Nishino, C., Enoki, N. and Tawata, S. (1987) Antibacterial activity and mode of action of plant flavonoids against *Proeteus vulgaris* and *Staphyloccus aureus*. Phytochemistry. **26**, 2213-2234.

Moustafa, E. and Wong, E. (1967) Purification and properties of chlacone-flavanone isomerase from soybean seed. Phytochemistry. **6**, 625-632.

Moyano, E., Martinez-Garcia, J.F. and Martin, C. (1996) Apparent redundancy in Myb gene function provides gearing for the control of flavonoid biosynthesis in Antirrhinum flowers. Plant Cell. **8**, 1519-1532.

Mur, L. (1995) Characterisation of members of the Myb gene family of transcription factors from *Petunia hybrida*. PhD Thesis, Vrije Univeristeit.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. **15**, 743-497.

Murre, C., McCaw, P.S. and Baltimore, D. (1989a) A new DNA binding and dimerisation motif in immunoglobulin enhancer binding, *daughterless*, MyoD and *myc* proteins. Cell. 56, 777-783.

Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B., Weintraub, H. and Baltimore, D. (1989b) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell. **58**, 536-544.

Nair, P.M. and Vining, L.C. (1995) Cinnamic acid hydroxylase in Spinach. Phytochemistry. **4**, 161-168.

Nakai, K. and Kanehisa, M. (1992) A knowledge base for predicting protein localisation sites in eukaryotic cells. Genomics. **14**, 897-911.

Nakajima, J., Tanaka, Y., Yamazaki, M. and Saito, K. (2001) Reaction mechanism from leucoanthocyanidin to anthocyanidin 3-glucoside, a key reaction for coloring in anthocyanin biosynthesis. J. Biol. Chem. **13**, 25797-25803.

Nakajima, T., Uchida, C., Anderson, S.F., Lee, C.G., Hurwitz, J., Parvin, J.D. and Montminy, M. (1997) RNA helicase A mediates association of CBP with RNA polymerase II. Cell. **90**, 1107-1112.

Nesi, N., Jond, C., Debeaujon, I., Caboche, M. and Lepiniec, L. (2001) The arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. Plant Cell. **13**, 2099-2114.

Newman T.C., Ohme-Takagi, M., Taylor, C.B. and Green, P.J. (1993) DST sequences, highly conserved among plant SAUR Genes, target reporter transcripts for rapid decay in tobacco. Plant Cell. **5**, 701-714.

Nikolov, D. B. and Burley, S. K. (1997) RNA polymerase II transcription initiation: A structural view. Proc. Natl. Acad. Sci. USA. **94**, 15-22.

Nishikawa, T., Okamura, H., Nagadoi, A., Konig, P., Rhodes, D. and Nishimura Y. (2001) Solution structure of a telomeric DNA complex of human TRF1. Structure. **9**, 1237-1251.

Noda, K.-L., Glover, B.B., Linstead, P. and Martin, C. (1994) Flower colour intensity depends on specialised cell shape controlled by a Myb-related transcription factor. Nature. **369**, 661-664.

Nomura, N., Takahashi, M., Matsui, M., Ishii, S., Date, T., Sasamoto, S., and Ishizaki, R. (1988) Isolation of human cDNA clones of Myb-related genes, a-Myb and b-Myb. Nucl. Acids Res. **16**, 11075-11089.

Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S. and Nishimura, Y. (1992) Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. Pro. Natl. Acad. Sci. USA. **89**, 6428-6432.

Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S. and Nishimura, Y. (1994) Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. Cell. **79**, 639-648.

Ohi, R., McCollum, D., Hirani, B., Haese, G.J.D., Zhang, X., Burke, J.D., Turner, K. and Gould, K.L. (1994) The *Schizosaccharomyces pombe* cdc5+ gene encodes an essential protein with homology to c-Myb. EMBO J. **13**, 471-483.

Oppenheimer, D.G., Herman, P.L., Sivakumaran, S., Esch, J. and Marks, M.D. (1991) A Myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. Cell. **67**, 483-493.

Oren-Shamir, M., Shaked-Sachray, L., Nissim-Levi, A. and Ecker R. (1999) Anthocyanin pigmentation of lisianthus flower petals. Plant Sci. **140**, 81-86.

Parvin, J.D. and Young, R.A. (1998) Regulatory targets in the RNA polymerase II holoenzyme. Curr. Opin. Genet. Dev. **8**, 565-570.

Payne, C.T., Zhang, F. and Lloyd A.M. (2000) GL3 encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with GL1 and TTG1. Genetics. **156**, 1349-1362.

Paz-Ares, J., Wienand, U., Peterson, P.A. and Saedler, H. (1986) Molecular cloning of the C-locus of *Zea mays*: a locus regulating the anthocyanin pathway. EMBO J. **5**, 829-833.

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A. and Saedler, H. (1987) The regulatory *C1* locus of *Zea mays* encodes a protein with homology to Myb proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. **6**, 3553-3558.

Paz-Ares, J., Peterson, P.A. and Saedler, H. (1990) Molecular analysis of the C1-I allele from *Zea mays*: a dominant mutant of the regulatory *C1* locus. EMBO J. **9**, 315-321.

Pelletier, M.K. and Shirley, B.W. (1996) Analysis of flavanone 3-hydroxylase in arabidopsis seedlings. Coordinate regulation with chalcone synthase and chalcone isomerase. Plant Physiol. **111**, 339-345.

Pelletier, M.K., Burbulis, I.E. and Winkel-Shirley, B. (1999) Disruption of specific flavonoid genes enhances the accumulation of flavonoid enzymes and end-products in arabidopsis seedlings. Plant. Mol. Biol. **40**, 45-54.

Peters, C.W.B., Sippel, A.E., Vingron, M. and Klempnauer, K.H. (1987) Drosophila and vertebrate Myb proteins share two conserved regions, one of which functions as a DNA-binding domain. EMBO J. **6**, 3085-3090.

Peters, N.K., Frost, J.W. and Long, S.R. (1986) A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. Science. **233**, 977-980.

Phillips, D.A. and Tsai, S.M. (1992) Flavonoids as plant signals to rhizosphere microbes. Mycorrhiza. **1**, 55-58.

Pinson, B., Kongsrud, T. L., Ording, E., Johansen, L., Fornier-D, B. and Gabrielsen, O.S. (2000) Signaling through regulated transcription factor interaction: mapping of a regulatory interaction domain in the Myb-related Bas1p. Nucl. Acids. Res. **28**, 4665-4673.

Powell, R.G., Tepaske, M.R., Plattner, R.D., White, J.F. and Clement, S.L. (1994) Isolation of resveratol from *Festuca versuta* and evidence or the widespread occurencen of this stilbene in the Poaceae. Phytochemistry. **35**, 335-338.

Prescott, A. and Martin, C. (1987) Rapid method for the quantitative assessment of levels of specific mRNAs in plants. Plant Mol. Biol. Rep. **4**, 219-224.

Ptashne, M. (1988) How eukaryotic transcriptional activators work. Nature. **335**, 683-689.

Pueppke, S.G. (1996) The genetic and biochemical basis for nodulation of legumes by rhizobia. Crit. Rev. Biotec. **16**, 1-51.

Quattrocchio, F. (1994) Regulatory genes controlling flower pigmentation in *Petunia hybrida*. PhD Thesis, Vrije Univeristeit, Amsterdam.

Quattrocchio, F., Wing, J.F., Leppen, H.T.C., Mol, J.N.M. and Koes, R.E. (1993) Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. Plant Cell. **5**, 1497-1512.

Quattrocchio, F., Wing, J.F., Van der Woude, K., Mol, J.N.M. and Koes, R. (1998) Analysis of bHLH and Myb-domain proteins: Species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J. **13**, 475-488.

Rabinowicz, P.D., Braun, E.L., Wolfe, A.D., Bowen, B. and Grotewold, E. (1999) Maize R2R3 Myb genes: sequence analysis reveals amplification in the higher plants. Genetics. **153**, 427-444.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.-Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M. Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K. and Yu, G.-L. (2000) Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. Science. **290,** 2105-2109.

Romero, I., Fuertes, A., Benito, M.J., Malpica, J.M., Leyva, A. and Paz-Ares, J. (1998) More than 80 R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana.* Plant J. **14,** 273-284.

Rosinski, J.A. and Atchley, W.R. (1998) Molecular evolution of the Myb family of the transcription factors: Evidence for polyphyletic origin. J Mol. Evol. 46, 74-83.

Sablowski, R.W.M., Moyano, E., Culianez-Macia, F.A., Schuch, W., Martin, C. and Bevan, M. (1994) A flower specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. EMBO J. **13,** 128-137.

Saikumar, P., Gabriel, J.L. and Reddy, E.P. (1994) The Myb oncogene product induces DNA-bending. Oncogene. **9,** 1279-1287.

Sainz, M.B., Grotewold, E. and Chandler, V.L. (1997a) Evidence for direct activation of an anthocyanin promoter by the maize C1 protein and comparison of DNA binding by related Myb domain proteins. Plant Cell. **9,** 611-625.

Sainz, M.B., Goff, S.A. and Chandler, V.L. (1997b) Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation in vivo. Mol. Cell Biol. **17,** 115-122.

Saito, K., Kobayashi, M., Gong, Z., Tanaka, Y. and Yamazaki, M. (1999) Direct evidence for anthocyanin synthase as a 2-oxoglutarate-dependent oxygenase: molecular cloning and functional expression of cDNA from a red forma of *Perilla frutescens.* Plant J. 17, 181-189.

Sakura, H., Kanei-Ishii, C., Nagase, T., Nakagoshi, H., Gonda, T.J. and Ishii, S. (1989) Delineation of three functional domains of the transcriptional activator encoded by the c-Myb protooncogene. Proc. Natl. Acad. Sci. USA. **86**, 5758-5762.

Salzman, R.A., Fujita, T., Zhu-Salzman, K., Hasegawa, P.M. and Bressan, R.A. (1999) An impoved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. Plant Mol. Biol. Rep. **17**, 11-17.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning: A laboratory manual*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

Sawa , S. (2002) Overexpression of the AtmybL2 gene represses trichome development in arabidopsis. DNA Res. **30**, 31-34.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carré, I.A. and Coupland, G. (1998) The late elongated hypocotyl mutation of arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. Cell. **93**, 1219-1229.

Schindler, U., Beckmann, H. and Cashmore, A.R. (1992a) TGA1 and G-box binding factors: Two distinct classes of arabidopsis leucine zipper proteins compete for the G-box-like element TGACGTGG. Plant Cell. **4**, 1309-1319.

Schindler, U., Menkens, A.E., Beckmann, H., Ecker, J.R. and Cashmore, A.R. (1992b) hetrodimerisation between light regulated and ubiquitously expressed arabidopsis GDF bZIP proteins. EMBO J. **11**, 1261-1273.

Schmitz, G., Tillmann, E., Carriero, F., Fiore, C., Cellini, F. and Theres, K. (2002) The tomato *Blind* gene encodes a MYB transcription factor that controls the formation of lateral meristems. Proc. Natl. Acad. Sci. USA. **99**, 1064-1069.

Schröder, J. (1997) A family of plant-specific polyketide synthases: facts and predictions. Trends Plant Sci. **2,** 373-378.

Schröder, J. (1999) The chalcone/stilbene synthase-type family of condensing enzymes. In: Sankawa, U. (ed.), *Comprehensive Natural Products Chemistry, Vol. 1: Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives*. Elsevier Science, Oxford, pp. 749-771.

Schröder, J. (2000) The family of chalcone synthase-related proteins: functional diversity and evolution. Recent Adv. Phytochemistry. **34,** 55-89.

Schulze-Lefert, P., Dangl, J.L., Becker-Andre, M., Hahlbrock, K. and Schulz, W. (1989) Inducible in-vivo DNA footprints define sequences necessary for UV light activation of the parsley chalcone ynthase gene. EMBO J. **8,** 651-656.

Schwechheimer, C., Zourelidou, M. and Bevan, M.W. (1998) Plant transcription factor studies. Annu. Rev. Plant Physiol. Plant Mol. Biol. **49,** 127-150.

Schwinn, K.E. (1994) Novel flavonoid 3'-hydroxylating activity in petunia. Polyphenols Actualities. **11,** 58.

Schwinn, K. E. (1999) Thesis: The regulation of anthocyanin biosynthesis in *Antirrhinum majus*. University of East Anglia.

Sheffer, R. D. and Kamemoto, H. (1977) Interspecific hybridisation involving *Anthurium andraeanum* Lind. and related species. Proc. Trop. Reg. Am. Soc. Hort. Sci. **19,** 275-283.

Shirley, B.W., Kubasek, W.L., Storz, G., Bruggenmann, E., Koornneef, M., Ausubel, F.M. and Goodman, H.M. (1995) Analysis of arabidopsis mutants deficient in flavonoid biosynthesis. Plant J. **8,** 659-671.

Smolen, G., Pawlowski, L., Wilensky, S. E., and Bender, J. (2002) Dominant alleles of the basic-helic-loop-helix transcription factor ATR2 activate stress-response genes in Arabidopsis. Genet. **161**, 1235-1246.

Solano, R., Nieto, C., Avila, J., Cañas, L., Diaz, I. and Paz-Ares, J. (1995a) Dual DNA binding specificity of a petal epidermis specific MYB transcription factor (MYBPh3) from *Petunia hybrida*. EMBO J. **14**, 1773-1784.

Solano, R., Nieto, C. and Paz-Ares, J. (1995b) MYBPh3 transcription factor from *Petunia hybrida* induces similar DNA-bending/distortions on its two types of binding site. Plant J. **8**, 673-682.

Sompornpailin, K., Makita, Y., Yamazaki, M. and Saito, K. (2002) A WD-repeat-containing putative regulatory protein in anthocyanin biosynthesis in *Perilla frutescens*. Plant Mol. Biol. **50**, 485-495.

Spelt, C., Quattrocchio, F., Mol, J.N.M. and Koes, R. (2000) *Anthocyanin1* of petunia encodes a bHLH protein that directly activates transcription of structural anthocyanin genes. Plant Cell. **12**, 1619-1631.

Stafford, H.A. (1991) Flavonoid evolution: An enzymic approach. Plant Physiol. **96**, 680-685.

Stafford, H.A. and Lester, H.H. (1982) Enzymatic and non-enzymatic reduction of (+) dihydroquercetin to its 3,4-diol. Plant Physiol.**70**, 695-698.

Stich, K., Eidenberger, T., Wurst, F. and Forkmann, G. (1992) Enzymatic conversion of dihydroflavonols to flavan-3,4-diols using flower extracts of *Dianthus caryophyllus* L. (carnation). Planta. **187**, 103-108.

Stober-Grasser, U., Brydolf, B., Bin, X., Grasser, F.A., Firtel, R.A. and Lipsick, J.S. (1992) The Myb DNA-binding domain I is highly conserved in *Dictyostelium discoideum.* Oncogene 7, 589-596.

Stringer, K. F., Ingles, C. J. and Greenblatt, J. (1990) Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TRIID. Nature. **345,** 783-786.

Styles, E.D., Ceska, O. and Seah, K.-T. (1973) Developmental differences in action of R and B alleles in maize. Can. J. Genet. Cytol. **15,** 59-72.

Suzuki, A., Suzuki, T., Tanabe, F., Toki, S., Washida, H., Wu, C.Y. and Takaiwa, F. (1997) Cloning and expression of five Myb-related genes from rice seed. Gene. **198,** 393-398.

Takeda, K., Kubota, R. and Yagioka, C. (1985) Copigments in the the blueing of sepal colour of *Hydrangea macrophylla.* Phytochemistry. **24,** 1207-1209.

Tamagnone, L., Merida, A., Parr, A., Mackay, S., Culianez-Macia, F.A., Roberts, K. and Martin, C. (1998) The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. Plant Cell. **10,** 135-154.

Tanikawa, J., Yasukawa, T., Enari, M., Ogata, K., Nishimura, Y., Ishii, S. and Sarai, A. (1993) Recognition of specific DNA sequences by the c-Myb proto-oncogene product role of three repeat units in the DNA-binding domain. Proc. Natl. Acad. Sci. USA. **90,** 9320-9324.

Taylor, L.P. and Briggs, W. (1990) Genetic regulation and photocontrol of anthocyanin accumulation in maize seedlings. Plant Cell. **2,** 115-128.

Thain, S.C., Murtas, G., Lynn, J R., McGrath, R. B. and Millar A J. (2002) The circadian clock that controls gene expression in Arabidopsis is tissue specific. Plant Physiol. **130**, 102-110.

Tice-Baldwin, K., Fink, G.R. and Arndt, K. (1989) BAS1 has a Myb motif and activates HIS4 transcription only in combination with BAS2. Science. **246**, 931-935.

Urao, T., Noji, M.-A., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1996) A transcriptional activation domain of ATMYB2, a drought-inducible arabidopsis Myb-related protein. Plant J. **10**, 1145-1148.

van Blokland, R., de Lange, P., Mol, J.N.M. and Kooter, J.M. (1993) Modulation of gene expression in plants by antisense genes. In: Crooke, S.T. and Lebleu, B. (eds.), *Antisense Research and Applications*. CRC Press, London, pp 125-148.

van der Krol, A.R., Lenting, P.E., Veenstra, J., van der Meer, I.M., Koes, R.E., Gerats, A.G.M., Mol, J.N.M. and Stuitje, A.R. (1988) An antisense chalcone synthase gene in transgenic plants inhibits flower pigmentation. Nature. **333**, 866-869.

van der Krol, A.R., De Lange, P., Gerats, A.G.M., Mol, J.N.M. and Stuitje, A.R. (1990) Antisense chalcone synthase genes in *Petunia*: Visualisation of variable transgene expression. Mol. Gen. Genet. **220**, 204-212.

van Tunen, A. J., Koes, R. E., Spelt, C.E., van der Krol, A. R., Stuitje, A. R. and Mol J.N.M. (1988) Cloning of two chalcone flavanone isomerase genes from *Petunia hybrida*: coordinate, light regulated and differential expression of flavonoid genes. EMBO J. **7**, 1257-1263.

van Tunen, A. J. and Mol, J. N. M. (1991) Control of flavonoid synthesis and manipulation of flower colour. In: Grierson, D. (ed.), *Developmental regulation of plant gene expression.* Blackie and Son Ltd, Glasgow, pp. 94-130.

Venkatachalam, P., Thanseem, I. and Thulaseedharan, A. (1999) A rapid and efficient method for isolation of RNA from bark tissues of *Hevea brasiliensis*. Curr. Sci. **77**, 635-637.

Vogt, T. and Jones, P. (2000) glycosyltransferases in plant natural product synthesis: characterisation of a supergene family. Trends Plant Sci. **5**, 380-386.

Wada, T., Tachibana, T., Shimura, Y. and Okada, K. (1997) Epidermal cell differentiation in arabidopsis determined by a Myb homolog, CPC. Science. **277**, 1113-1116.

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 dorsoventrality of lateral organs in Antirrhinum. Cell. **93**, 779-789.

Walker, A.R., Davison, P.A., Bolognesi-Winfield, A.C., James, C.M., Srinivasan, N., Blundell, T.L., Esch, J.J., Marks, M.D. and Gray, J.C. (1999) The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. Plant Cell. **11**, 1337-1349.

Wang, Z.-Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, M.S. and Tobin, E.M. (1997) A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. Plant Cell. **9**, 491-507.

Wang Z.-Y. and Tobin, E.M. (1998) Constitutive expression of the gene disrupts circadian rhythms and suppresses its own expression. Cell. **93**, 1207-1217.

Wannakrairoj, S. and Kamemoto, H. (1990a) Histological distribution of anthocyanins in Anthurium spathes. Hort. Sci. **25**, 809.

Wannakrairoj, S. and Kamemoto, H. (1990b) Inheritance of purple spathe in Anthurium. J. Amer. Soc. Hort. Sci. **115**, 169-171.

Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T.K., Turner, D., Rupp, R., Hollenberg, S., Zhuang, Y. and Lassar, A. (1991) The myoD gene family: Nodal point during specification of the muscle cell lineage. Science. **251**, 761-766.

Weiss, D., van der Huit, A.H., Kroon, J.T.M., Mol, J.N.M. and Kooter, J.M. (1993) The petunia homologue of the antirrhinum candi and *Zea mays* A2 flavonoid genes; homology to flavanone 3-hydroxylase and ethylene-forming enzyme. Plant Mol. Biol. **22**, 893-897.

Weisshaar, B., Armstrong, G.A., Block, A., da Costa e Silva, O. and Hahlbrock, K. (1991) Light-inducible and constitutively expressed DNA-binding proteins recognising a plant promoter element with functional relevance in light responsiveness. EMBO J. **10**, 1777-1786.

Weston, K. (1998) Myb proteins in life, death and differentiation. Curr. Opin. Gen. Dev. **8**, 76-81.

Whedale, M.W. (1914) Our present knowledgeof the chemistry of the Mendelian factors for flower colour. J. Genetics. **4**, 109-120.

Whedale, M.W. (1915) Our present knowledge of the chemistry of the Mendelian factors for flower colour Part 2. J. Genetics. **4**, 369-376

Wieser, J. and Adams, T.H. (1995) FlbD encodes a Myb-like DNA-binding protein that coordinates initiation of *Aspergillus nidulans* conidiophore development. Genes Dev. **9**, 491-502.

Wilkins, T.A. and Smart, L.B. (1996) Isolation of RNA from plant tissue. In: Krieg P.A. (ed.), *A Laboratory Guide to RNA:Isolation, Analysis, and Synthesis*. Wiley-Liss, Inc, New York, pp. 21-41.

Williams, C.A., Harborne, J.B., and Mayo, S.J. (1981) Anthocyanin pigments and leaf flavonoids in the family Araceae. Phytochemistry. **20,** 217-234.
Williams, C.E. and Grotewold, E. (1997) Differences between plant and animal Myb domains are fundamental for DNA binding activity, and chimeric Myb domains have novel DNA binding specificities. J. Biol. Chem. **272,** 563-571.

Wilmouth, R.C., Turnbull, J.J., Welford, R.W., Clifton, I. J., Prescott, A.G. and Schofield, C.J. (2002) Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. Structure. **10**, 93-103.

World Market for Anthurium RAP Market Information Bulletin No.11, (1997) http://www.marketag.com/ma/bulletins/market/anthur.stm

Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Reiger, J. L., Triezenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J. and Greenbalt, J. (1994) Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. Mol. Cell. Biol. **14**, 7013-7024.

Yamazaki, M., Yamagishi, E., Gong, Z., Fukuchi-Mizutani, M., Fukui, Y., Tanaka, Y., Kusumi, T., Yamaguchi M. and Saito, K. (2002) Two flavonoid glucosyltransferases from *Petunia hybrida*: molecular cloning, biochemical properties and developmentally regulated expression. Plant Mol. Biol. **48**, 401-411.

Yang, Y. and Klessig, D.F. (1996) Isolation and characterisation of a tobacco mosaic virus-inducible Myb oncogene homolog from tobacco. Proc. Natl. Acad. Sci. USA. **93**, 14972-14977.

Zawel, L., Kumar, K.P., and Reinberg, D. (1995) Recycling of the general transcription factors during RNA polymerase II transcription. Genes and Dev. **9**, 1479-1490.

APPENDICES

Appendix I.

List of primers used for various experiments during this research

Degeneracy code: R=A+G; Y=C+T; S=G+C; D=A+G+T; N=A+G+C+T

CHI 1: GTNACNGGNCCNTTYGARAARTT
CHI 1-1: GTIACIGGNCCNTTYGARAARTT
F3'5'H-L: CGGAATTCAYGCICARGAYATGGTITTY
F3'5'H-R: CGGGATCCACCRAAIGGDATIARYTCRAA
UFGT 3-1 ext'd: AGCTTTAGATAAAGGGCCATCAGTTG
MYB C1: GGCAATTCTTNAYNGCRTTRTCNGT
MYB C2: CGGAATTCTTDAYYTCRTTRTCNGT
MYB C3: GGCAATTCTTNAYNTKRTTRTCNGT
MYB N1: CGGAATTCDKNAARAGYTGYAG
MYB N2: CGGAATTCDKNAARAGYTGYCG
MYB N3: CGGAATTCDKNAARTCNTGYAG
MYB N4: CGGAATTCDKNAARTCNTGYCG
MYB N5: CCCGGGTGYGGNAARTCNTG
MYB N6: GGAATTCTGYGGNAARAGYTG
MYB 20-5/18A: GATATGATCATTCGACTCCACAACCTC
MYB 20-5/18B: ATCAAGAGAGGCAACATAACCGAGGCT
MYB 20-3A: AATCGGTGGTCTCTCATCGCTGGGCGG
MYB 20-3B: AAGTTCACAGAGGACGAAGATGACCTA
ROSEA SP: TGCTTTTCGAAAGATGAACTGGACCTC
ADAPTOR dT: GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT
MYBP7: CGGAATTCGAGGAGATGACGGCGAAGGCAAAG
MYBP8: CCATCGATTGGAGGACAGATCACATAATCCAT
DFRGSP1: CAGCCTCATGATCAGCCATGAGCCAAC
DFRGSP2: GCACACGGTGCCCTTGTGCATCATCCT

DFRP27: CCCAAGCTTACTATAGGGCACGCGTGGT

DFRP28: GTACCCATGGCCTCGCAGATTATATATAACCAGC

PSeqA: CCAGCAAATTTATGAGCTACCCACGT

Appendix II. Nucleotide sequence encoding AaCHS1. The ATG start codon is shaded in red.

ATGGTGAGCACGAGCTTGGAAGCCATCCGCAAGGCGCAGAGAGCCGATGGGCCAGCCACC 70

TCGGCACGGCCGTCCCACCCAACGCCGTCGACCAGAGCACCTACCCAGACTACTACTTCCG 140

CAGTGAGCACCAGGTTGAGCTGAAGGAGAAGTTCAGGCGCATGTGTGAGAAGTCGATGAT 210

CACATGTACCTGACGGAGGAGATACTGAAGGAGAACCCGAGCATGTGCGCGTACATGGCG 280

ACGCGCGGCAGGACATGGTGGTGGTGGAGGTGCCCAGGCTCGGCAAGGAGGCGGCCACCC 350

GGAGTGGGGGCAGCCCAAGTCCAAGATCACCCACCTCGTCTTCTGCACCACCAGCGGCGTC 420

GGCGCCGACTACCAGCTCGCCAAGCTCCTCGGCCTCCGCCCCTCCGTCAAGCGCCTCATGA 490

AGGGCTGCTTCGCCGGCGGCACCGTCCTCCGCCTCGCCAAGGACCTCGCCGAGAACAACCG 560

TGTGCTCGTCATCTGCTCCGAGGTCACCGCCGTCACCTTCCGCGGCCCCAGCGAGTCCCACC 630

CTCGTCGGCCAGGCCCTCTTCGGCGACGGCGCGTCCGCCCTCGTCGTGGGGGCCGACCCCG 700

TCGAGCGGCCCATCTTCCAGGTGGTGTCGGCGGCTCAGACCATCCTCCCGGACAGCCACGG 770

CGGCCACCTCCGGGAGGTGGGGCTCACCTTCCACCTCCTCAAGGACGTGCCCGGCCTCATC 840

ATTGACAAGTCTCTTGTGGAGGCCTTCGAGCCGCTGGGCATCTCCGACTGGAACTCCCTCTT 910

CCCACCCCGGGGGCCCCGCCATCCTGGACCAGGTGGAGGACAAGCTCCGGCTCCAGTCGG 980

GGCCACCCGGAGGGTGCTGAGCGAGTACGGCAACATGTCCAGCGCCTGCGTCCACTTCATC 1050

ATGCGGAAGAGGTCGGCCGAGGAAGGGAGGGGCACCACCGGCGAGGGCCTCGAGTGGGG 1120

GGTTCGGGCCCGGCCTCACCGTGGAGACCCTCGTCCTCCACAGCGTCGC 1176

Appendix III. Nucleotide sequence encoding AaF3H1. The ATG start codon is shaded in red.

[redacted]GTTCCCGCGGCAACACCCTTCCTGCCGACGACCGCGGAGGAGGCGACGCTGCGCCCCA 70

GCGACGAGGACGAGCGCCCCAAGGTGCCCCACAACCAGTTCAGCGACGCCATCCCCGTCCT 140

CGGCATCGACGACTCCGACGGGTCGCCCCGCAGGGCGGAGCTGTGCCGCGAGATCGTGGA 210

GGGTGGGGCATCTTCCAGGTGGTGGACCACGGCGTCGACCCAGGCCTGGTCGCCGACATGA 280

CCACGGAGTTCTTCGCCCTCCCGCCCGAGGACAAGCTCCGCTACGACATGTCCGGGGGCAA 350

CTTCATCGTCTCCAGCCACCTCCAGGGGGAGGCCGTGCAGGACTGGAGGGAGATCGTGACC 420

TACCCAGTACGGGCGCGGGACTACACGAGGTGGCCCGACAAGCCGGAGGGGTGGCGGGCG 490

CGTACAGCGAGGGGTTGATGGGGCTCGCCTGCAAGCTGCTGGGGGTGCTGTCCGAGGCCAT 560

CAAGGAGGCGCTCGCCAAGGCCTGCGTCGACATGGACCAGAAGGTGGTGGTGAACTTCTA 630

CCCCAGCCCGACCTCACCCTCGGGGTCAAGCGCCACACCGACCCCGGCACCATCACCCTCC 700

ACCAGGTCGGCGGCCTCCAGGCCACCAGAGACGGCGGCAAGACCTGGATCACCGTCCAGC 770

CGCCTTCGTCGTCAACCTCGGCGACCACGGTCACCTCCTGAGCAACGGGCGGTTCAAGAAC 840

CAGGCGGTGGTGAACTCGGAGCGCAGCCGGCTGTCGATCGCGACGTTCCAGAACCCGGCG 910

TCGTGTACCCGCTGGCCGTCCGGGAGGGGGAGAAGCCGGTGCTCGACGAGCCCATCTCCTT 980

GTACCGGAGGAAGATGAGCCGAGACCTGGAGCTCGCCAGGCTCAAGAAACAGGCCAAGCT 1050

GACGTGCTGGCCAAGGCCGAGGATGCCAACAAACACAAAGCCCTCGATGAGATTCT 1115

Appendix IV. Nucleotide sequence encoding AaDFR1. The ATG start codon is shaded in red.

ATGATGCACAAGGGCACCGTGTGCGTGACGGGCGCTGCCGGGTTCGTTGGCTCATGGCTGA 70

TCCTCGAGCAGGGTTACTCTGTCAAGGCCACCGTCCGCGACCCCAGCAACATGAAGAAGGT 140

GCTCGACCTTCCAGGGGCCGCGAACCGCCTGACCCTCTGGAAGGCCGACCTCGTCGATGAG 210

GACGAGCCCATCCAAGGCTGCACCGGCGTCTTCCACGTCGCGACGCCCATGGACTTCGAAT 280

CAGAGAGTGAGATGATCAAGCCGACGATCGAGGGGATGCTGAACGTGTTGCGGTCGTGTG 350

CAGCACCGTCCGGCGAGTCGTCTTCACATCCTCGGCCGGCACGGTCTCCATACACGAGGGC 420

CTCTACGACGAGACCTCCTGGAGCGACGTCGACTTCTGCAGGGCCAAGAAGATGACCGGAT 490

TTGTGTCGAAGACGTTGGCCGAGAAGGCCGCGTGGGACTTTGCGGAGAAGAACAACATCG 560

CATTATCCCCACCCTAGTCAACGGCCCCTTCGTCATGCCCACCATGCCGCCCAGCATGCTCT 630

GCCCTCATCACAAGGAACGAGCCGCACTATTCGATCCTGAACCCGGTGCAGTTCGTCCACC 700

TCTGCAACGCCCACATCTTCCTGTTCGAGTGCCCCGACGCCAAGGGAAGGTACATCTGCTC 770

CGTCACCATCGCCGGCCTCGCCCAGATACTCCGGCAGCGCTACCCTGAGTTCGACGTCCCC 840

GGAGAGATGGAGGTCTTCGACATCATAAGCTACTCGTCCAAGAAGCTCACGGACCTAGGCT 910

AGTACAGCTTAGAGGATATGTTTGATGGGGCGATTCAGTCCTGCAGAGAGAAGGGCTTGCT 980

CACCAAGGAGCCATCATATGCCACTGAACAATTAATTGCCACCGGGCAAGACAACGG 1046

Appendix V. Nucleotide sequence encoding AaANS1. The ATG start codon is shaded in red.

ATGACGGCAGTGCCGGCGGGGTCGAGGGTGGAGAGCCTGGCGAGCAGCGGAATCCAGGCC 70

AGTACGTCCGGCCCGCGGAGGAGCGAGTCAGCCTCACCGACGCGCTCGAGGCGGCCAGGA 140

CGGCCCGCAGATCCCCACGGTGGACGTCGCGGGGTTCTCGTCCGGGGACGAGGCGGCGCG 210

GCGGAGGCGGTGCGGCGGGCCGCGACGGACTGGGGGGGTGATGCACGTCGTCAACCACGGC 280

AGCTGATCCGGCGCATGCAGGCGGCCGGGGAGGCCTTCTTCGCCCTCCCCATCGAGGAGAA 350

CGCCAACGACCAGTCGTCCGGGAACATCCAGGGCTACGGGAGCAAGCTCGCCAACAATGC 420

CTCGAGTGGCAGGACTACTTCTTCCACCTCATCTTCCCCGAGGACAAGGCCAACTTCTCTAT 490

AGCAGCCCGCCAACTACGTTGAAGAGACGAGGGAGTTCGGGCGGCAGCTGCGGGTGGTGG 560

GCTGGCGATGCTGTCGCTGGGGCTGGGGGTGGAGGAAGGGAAGCTGGAAGCGGAGGTGGG 630

GACCTGCTGCTGCAGATGAAGATAAACTACTACCCGCGGTGCCCGCAGCCGGAGCTGGCG 700

AGGCGCACACGGACGTCAGCGCCCTCAGCTTCATCCTCCACAACATGGTGCCCGGGCTGCA 770

CGAGGGCCAGTGGGTGACGGCGCGCTGCGTCCCCGACTCCATCATCATGCACGTCGGCGAC 840

ATCCTCAGCAACGGGCAGTACAAGAGCATCCTCCACCGCGGCCTCGTCAACAAGGAGAAG 910

CCTGGGCCGTCTTCTGCGAGCCCCCGCGCGACAAGATCCTGCTCCAGCCCCTGCCCGACAT 980

GGGCCACCCCGCGCAGTTCCCCCCGCGCACCTTCTCCCAGCACATCCAGCACAAGCTCTTC 1050

CAAGGTGACTTCGCCACCCCCAAC 1078

Appendix VI. Nucleotide sequences encoding anthurium MYB groups A-C

Group A

TGCGGAAAGAGCTGCAGATTGAGGTGGATAAACTACCTGAGGCCTGATCTCAAGAGAGGC 70

AGCAGGAGGAGAACCAGATCATTGAGCTACATGCAGTCCTGGGCAACAGGTGGTCTCAGA 140

ATTGCCCGGGAGGACTGACAATGAAATTAAG 176

Group B

TGTGGGAAGAGCTGCAGGCTCAGGTGGACCAACTACCTGAGACCCGACATCAAGAGGGGC 70

TGCAGGAGGAGCAGACCATAATTCAACTCCATGCTCTGCTGGGCAACCGGTGGTCGGCCAT 140

CCTGCCGAAGAGGACAGACAATGAAATCAA 176

Group C

TGTGGGAAGAGTTGCAGGCTGAGGTGGATCAACTACCTTAGGCCCAATCTCAAGAGAGGG 70

AGGACGAAGATGACCTAATAATCAAGCTCCATGCACTCCTAGGCAATCGGTGGTCTCTCAT 140

GTTGCCAGGCCGAACGGACAATGAAATAAAGA 178

Appendix VII. Nucleotide sequences encoding anthurium MYB groups D-F

Group D

TGTGGTAAGAGCTGCCGTCTCCGTTGGATAAACTACCTCCGGCCCGACCTCAAACGGG 70

AGGAAGAAGACGAGCTCATCATCAAGCTCCGTGAGCTGCTAGGCAACAAGTGGTCTC 140

CTTGCCGGGGAGGACGGACAACGAAATTAAG 179

Group E

TGTGGCAAGAGCTGCAGGCTAAGGTGGACCAACTATCTCAGGCCGGACCTGAAGAGA 70

AATCTGAGGAACAGATGGTCATCGACCTCCATGCCAGACTCGGTAACAGGTGGTCCA 140

TCTACCAGGTAGAACGGACAACGAGATCAA 177

Group F

TGCGGGAAGAGCTGCAGGCTAAGGTGGCTCAACTACCTCAGGCCGGGCATCAAGAGA 70

AGGCTGAGGAGGATATGATCATTCGACTCCACAACCTCATTGGCAACAGGTGGTCTCT 140

ATTGCCCGGTCGAACTGACAACGAGATTAA 177

Appendix VIII. Nucleotide sequence encoding the C-terminus of anthurium MYB group C

AATCGGTGGTCTCTCATCGCTGGGCGGCTGCCAGGCCGAACTGACAATGAGATCAAGAACTACTGGAACT 70

CTCATCTGAGGAAGAAGCTCATCCACATGGGGGTGAACCCAAACTCAAGAAATGCAGCACTTCCGGCTCC 140

AATGAGCATCCACAGGTGCAGCAGTTGTGCACTACGTCCATACCACAGAGCCAACAGCATTAAGTTCGAA 210

GACGAAGCCAGCTGCCTTGAAGACAATGCTCGCGTGTTGCCCGATTTGAATATTGATCTCAACATCTCAT 280

TTGCATCCCCTCCCCATGGTTTGGTGGATCAGAAGCAACATCCGGAAGACGTTGAGATGAAGGGTGACCC 350

TGGGGCCACTTATCCTACGCTCGTCCTATTCAGATGA 387

Appendix IX. Nucleotide sequence encoding AaMYB1. The ATG start codon is shaded in red.

ATGACGGCGAAGGCAAAGAGGGGAGCAGCAGTAGCTGCAGAAGCAAGAGTGGCCAAGAAGAGGGTGGTGA 70

AGTTGAACAAGGGGCCCTGGACTGCCGAAGAAGACCAGAAGCTTGTGGAGTATGTCGATGCCCATGGCGA 140

CAGGAAGTGGACGAGCCTCCCCACCAAAGCTGGTTTAAATAGGTGTGGGAAGAGCTGCAGGCTAAGGTGG 210

CTCAACTACCTCAGGCCGGGCATCAAGAGAGGCAACATATCCGAGGCTGAGGAGGATATGATCATTCGAC 280

TCCACAACCTCATTGGCAACAGGTGGTCTCTGATTGCGGGTAGATTGCCCGGTCGAACAGACAACGAGAT 350

CAAGAACTACTGGAACACCCATCTGAGCAAGAAGCCATTAACCATCAGTGATCTCAACGACAAGCTCAAC 420

GACGACGGCGGAAGCCGCTCCGACGTCGACGAACCGCCTAGCTCCGGGCAGTCACAGGGTCCACTAGATT 490

CCCGGCGGCCAGCAGAGCCACGGATCAGTGAGGCATCTAACGAGGTGGCATTCAACGTGGATGAACTCGT 560

CAACCTCCCAGCAATCCTAGATTTTGATGATCTCCTGGACAGCGGCAGCTGCAGCGATGGCATCATCACG 630

AGCAGCCAGATTTTTTGGACACCAAACTTCTCTGCTCATGATCAGCAGGATATTTCGGACTTCGATGACC 700

TCGGGATGTTGATGGACATAATGGATTATGTGATCTGTCCTCCATGA 747

Appendix X. *cis*-regulatory elements of anthurium DFR promoter fragment generated from a PLACE database search.

Signal Scan Program

ıse searched: PLACE

```
is the sequence you submitted
/signalseq.9436  [Unknown form], 905 bases, 75E3FF62 checksum.
ATATAGTTACTCAACTTGTTGAGAGGTCTCTTAGACCAGAAGTGT
AAGGTTGGGGAAGGTGATCCAGATGGGAATCGAGGTCAGACTATC
CCAATCGAAGTTTTGGGATCCATTTTTGCAAAACAAAGGGATGCC
ATGCCCCAGAACCCTTCCAATGCCGGGACCATATCTTCCTCAAAT
AATCTAGTAATAATATAGGTTTTGTGATTTTGTATTGAGAGGAGG
GCTTTTGAAGAGAAAGGGAGCCCGACCTCCCACCTTCTCTAGGAA
TCCATCCGAGTCAAACTCGCGGTTCGGAGAGGTGCAACGAGAGGA
CCTAGAGGATCCAAATTCCAGAGTCCGCAATCCGAAGGTTCGGAA
CGAGCTACCAGTCTTCGCGCGTCCGTCTCCTTGTACCAACGGTAA
ACTCTTCCCGCCACGGGGAcTACCAAACGACGAACCACTACACGA
ATGCATATCCGTACAAAATGACGAATTACTCTCGCACATCGCCTA
CCTCTCCTCATGCAGCGAAGAGTGTTAGGTGGTTAGTTTGAACAC
TACAGCAGGAGGTAGATGACGGCCGGGGCCCTGAGACTGAACGCT
TGTCATATGTGACATGTCCCAGCAAATTTATGAGCTACCCACGTC
AAATAGAGAGAGAAGAGCGTAGGGTAGACACAGTTGCCTCCGATT
CTCCATCGATCTCTCCTTCCCACACACACACACACACACACACAC
CACACACATATATATATAGAGAGAGAGAGAGAGAGAGGGAGAGAG
ATCAGTTTCTGCCTCTATAAACAACAGCTGGTTATATATAATCTG
```

TS OF YOUR SIGNAL SCAN SEARCH REQUEST
signalseqdone.9436: 905 base pairs
l Database File:

Factor or Site Name		Loc.	(Str.)	Signal Sequence	SITE #
ESIIUDCRNI	site	166	(-)	GGWAGGGT	S000374
OTIFCAMV	site	524	(+)	TGACG	S000024
OTIFCAMV	site	622	(+)	TGACG	S000024
SAS1	site	282	(+)	CTCCCAC	S000226
OX1	site	4	(+)	CAAT	S000028
OX1	site	107	(+)	CAAT	S000028
OX1	site	173	(+)	CAAT	S000028
OX1	site	383	(+)	CAAT	S000028
OX1	site	239	(-)	CAAT	S000028
OX1	site	551	(-)	CAAT	S000028
NAPA	site	577	(-)	CNAACAC	S000148
GGMSAUR	site	503	(+)	CATATG	S000370
GGMSAUR	site	659	(+)	CATATG	S000370
GGMSAUR	site	503	(-)	CATATG	S000370
GGMSAUR	site	659	(-)	CATATG	S000370
BOX1	site	3	(+)	CCAAT	S000030
BOX1	site	106	(+)	CCAAT	S000030
BOX1	site	172	(+)	CCAAT	S000030
OSAMY3	site	483	(+)	CGACG	S000205
DIANLELHC	site	253	(-)	CAANNNNATC	S000252
REZM	site	55	(+)	AAAG	S000265
REZM	site	140	(+)	AAAG	S000265
REZM	site	268	(+)	AAAG	S000265
REZM	site	257	(-)	AAAG	S000265
REZM	site	851	(-)	AAAG	S000265

OREDCDC3	site	597	(+)	ACACNNG	S000292
NNAPA	site	75	(+)	CANNTG	S000144
NNAPA	site	503	(+)	CANNTG	S000144
NNAPA	site	598	(+)	CANNTG	S000144
NNAPA	site	659	(+)	CANNTG	S000144
NNAPA	site	736	(+)	CANNTG	S000144
NNAPA	site	880	(+)	CANNTG	S000144
NNAPA	site	75	(-)	CANNTG	S000144
NNAPA	site	503	(-)	CANNTG	S000144
NNAPA	site	598	(-)	CANNTG	S000144
NNAPA	site	659	(-)	CANNTG	S000144
NNAPA	site	736	(-)	CANNTG	S000144
NNAPA	site	880	(-)	CANNTG	S000144
OX	site	97	(-)	GATA	S000039
OX	site	187	(-)	GATA	S000039
OX	site	511	(-)	GATA	S000039
OX	site	855	(-)	GATA	S000039
NSENSUS	site	446	(+)	GRWAAW	S000198
NSENSUS	site	853	(-)	GRWAAW	S000198
TG10	site	69	(+)	GTGA	S000378
TG10	site	229	(+)	GTGA	S000378
TG10	site	664	(+)	GTGA	S000378
CORE	site	854	(-)	GATAA	S000199
PSADB	site	639	(-)	YTCANTYY	S000395
COREATCOR15	site	277	(+)	CCGAC	S000153
RD22	site	585	(-)	CTAACCA	S000175
RE	site	736	(+)	CNGTTR	S000176
RE	site	442	(-)	CNGTTR	S000176
RE	site	877	(-)	CNGTTR	S000176
ANT	site	579	(-)	MACCWAMC	S000167
M	site	58	(-)	CCWACC	S000179
1	site	97	(-)	GGATA	S000180
1	site	511	(-)	GGATA	S000180
RD22	site	598	(+)	CACATG	S000174
IG3	site	214	(+)	AATAAT	S000088
N1LELAT52	site	266	(+)	AGAAA	S000245
N1LELAT52	site	861	(-)	AGAAA	S000245
MENTZMZM13	site	88	(+)	AGGTCA	S000254
AAT	site	877	(+)	CAACA	S000314
AAT	site	22	(-)	CAACA	S000314
ALGLHCB21	site	511	(-)	CGGATA	S000363
OTIFTAPOX1	site	5	(-)	ATATT	S000098
OTIFTAPOX1	site	217	(-)	ATATT	S000098
PEATBNNAPA	site	564	(+)	CATGCA	S000264
OTIFGM7S	site	127	(+)	RTTTTTR	S000103
COREENHAN	site	584	(+)	GTGGWWHG	S000123
COREENHAN	site	486	(-)	GTGGWWHG	S000123
BOX2	site	703	(+)	TATAAAT	S000109
BOX4	site	701	(+)	TATATAA	S000111
BOX4	site	890	(+)	TATATAA	S000111
BOX4	site	887	(-)	TATATAA	S000111
PVTRNALEU	site	701	(-)	TTTATATA	S000340
XPCWRKY1	site	314	(-)	TTTGACT	S000310
ATNPR1	site	315	(-)	TTGAC	S000390

Appendix X1. Crude pH measurements of anthurium spathe cell sap.

Material and Method:

Crude pH measurements were taken of anthurium spathe cell sap using the Shindengen Isfet pH meter KS701 (Shindengen Electric MFG, Japan). The pH meter was calibrated with the standard buffer (at pH 6.9), provided with the kit.

Freshly picked spathe tissue (0.5 g) was collected from each of the anthurium cultivars used in this study (Altar, Atlanta, Lido, Montana and Acropolis). Tissue was also collected from the *roseal* mutant and used as a standard positive control as it has a known pH level of 5.5 by similar measurements (Schwinn, 1999). Effort was made to ensure there was little time delay between collecting the tissue and measuring the pH.

Tissue for each sample was crushed in a polyurethane bag with a pestle to extract the sap A droplet of the cell sap was then immediately added to the tip of the pH sensor as instructed in the manual, and the display value recorded. Three measurements were taken for each tissue and the average of these recorded. These average results are given below.

Table I. Crude pH values of antirrhinum and anthurium spathe cell sap.

Cultivar	Average pH value
roseal	5.5
Altar	5.7
Atlanta	5.5
Lido	5.6
Montana	5.7
Acropolis	6.2