GENETIC ANALYSIS OF THE EVOLUTION OF PETALOID BRACTS IN DOGWOODS!

by

Nicole M. Maturen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biology)
in The University of Michigan
2008

Doctoral Committee:

Professor George F. Estabrook, Co-chair
Michael Frohlich, Royal Botanical Gardens Kew, Co-chair
Associate Professor Robyn J. Burnham
Associate Professor Steven E. Clark

GENETIC ANALYSIS OF THE EVOLUTION OF PETALOID BRACTS IN DOGWOODS!

by

Nicole M. Maturen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biology)
in The University of Michigan
2008

Doctoral Committee:

Professor George F. Estabrook, Co-chair
Michael Frohlich, Royal Botanical Gardens Kew, Co-chair
Associate Professor Robyn J. Burnham
Associate Professor Steven E. Clark

ACKNOWLEDGEMENTS

I give thanks to and for the following people and institutions:

my Ph.D. advisors Michael Frohlich and George Estabrook
my committee members Robyn Burnham and Steve Clark
the Department of Ecology and Evolutionary Biology
the Jer-Ming Hu lab at National Taiwan University, Taipei
the Department of Botany at the Natural History Museum, London
the Royal Botanical Gardens, Kew
the Chelsea Physic Garden, London
The Regents of the University of Michigan
The Hanes Fund
The National Science Foundation
The Organization for Tropical Studies
The Cole Award
The Youngman Award

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

LIST OF FIGURES vi

ABSTRACT viii

CHAPTER I BRACT DEVELOPMENT IN *CORNUS*

ABSTRACT 1

INTRODUCTION 1

Relationships within the genus.....2

Close relatives of *Cornus* 4

Petaloidy in *Cornus* 4

METHODS AND MATERIALS 5

Materials 5

Scanning Electron Microscopy 5

RESULTS 5

Bract development 5

Epidermal morphology of bracts 8

Teratologous forms 9

DISCUSSION 9

CHAPTER II PETAL ORGAN IDENTITY GENES IN *CORNUS*

ABSTRACT 23

INTRODUCTION 23

Floral organ identity genes are MADS-box genes 24

Divergence and duplication among floral organ identity genes 25

METHODS AND MATERIALS 28

Plant materials and gene cloning 28

Sequence identification 29

RESULTS 30

Sequences of MADS-box genes identified 30

The B-class genes 31

The A-class and E-class genes 32

Other MADS-box genes 33

DISCUSSION 33

Duplication and divergence in *Cornus* MADS-box genes 34

Relationships within the genus *Cornus* 35

CHAPTER III EXPRESSION OF PETAL ORGAN IDENTITY GENES IN *CORNUS*

ABSTRACT 38

INTRODUCTION 38

Expression patterns and functions of petal organ identity genes 39

Dimerization of petal organ identity genes 41

Ternary and quaternary interactions among A, B, E genes 42

Regulation of petal organ identity genes in model organisms 43

METHODS AND MATERIALS 46

Plant materials 46

Quantitative real-time PCR 46

RESULTS 48

DISCUSSION 49

FUTURE WORK 56

APPENDIX A: Degenerate primers used to isolate MADS-box genes 58

LITERATURE CITED 59

LIST OF FIGURES

Figure I.1. Putative relationships among subgroups of *Cornus* 11
Figure I. 2. Showy-bracted *C. florida* f. *pluribracteata* at RBG Kew 12
Figure I. 3. Showy-bracted *C. nuttallii* x *florida* at RBG Kew and *C. disciflora* at Cuereci, Costa Rica 13
Figure I. 4. Showy-bracted *C. kousa* var. *chinensis* at RBG Kew 14
Figure I. 5. Showy-bracted *C. kousa* at RBG Kew and showy-bracted dwarf *C. canadensis* at The Chelsea Physic Garden 15
Figure I. 6. Small-bracted *C. officinalis* and *C. mas* at RBG Kew..... 16
Figure I. 7. Small-bracted *C. chinensis* at RBG Kew..... 17
Figure I. 8. Bractless dogwoods at RBG Kew..... 18
Figure I.9. *C. kousa* var. *chinensis*, SEM 19
Figure I. 10. *C. florida* f. *pluribracteata*, SEM 19
Figure I.11. *C. canadensis*, SEM 20
Figure I.12. *C. officinalis* and *C. chinensis*, SEM 20
Figure I. 13.Unusual forms observed in flowering dogwoods 21

Figure II. 1. Relationships among major clades of MADS-box genes families in plants, redrawn from Becker and Theissen, 2003 36
Figure II. 2. Phylogeny of newly-identified *Cornus* MADS-box genes and reference genes 37

Figure III. 1. Relative abundance of floral organ identity genes in the showy-bracted C. florida 52
Figure III. 2. Relative abundance of floral organ identity genes in the showy bracted dwarf C. canadensis..... 53

Figure III. 3. Relative abundance of floral organ identity genes in the small, non-showy-bracted C. mas 54
Figure III. 4. Relative abundance of floral organ identity genes in the bractless C. alba 55

ABSTRACT

The genus *Cornus* is best known for beautiful flowering dogwood trees that have large petaloid bracts. Another group of species within *Cornus*, the dwarf dogwoods, also bear petaloid bracts, and whether there were one or two origins of petaloid bracts in *Cornus* is debated. Since the discovery of MADS-box floral organ identity genes in model organisms, the molecular evolution of this gene family has been investigated and implicated in floral evolution, especially in origins of petals/petaloidy and resulting perianth diversity. We hypothesize that ectopic expression of petal organ identity genes may play a key role in the transition of inflorescence bracts from small and leaf-like to large and petal-like in two clades of *Cornus*. Here we identify A, B, and E class MADS-box genes from across the genus and investigate their expression in bracts, flowers, receptacles and leaves of four species of dogwood representing the four clades of the genus. Our results of real-time quantitative PCR show that A, B and E class genes are significantly expressed in bracts of the flowering dogwood *C. florida* and are not significantly expressed in bracts of the dwarf dogwood *C. canadensis* or the non-showybracted *C. mas*; this difference in genetic formulas for petaloidy indicates that it may have originated twice in the genus. Our results also show that A and E class genes are highly expressed in receptacle tissue of bractless and non-showy-bracted dogwoods; this expression may have been a pre-adaptation for petaloidy.

CHAPTER I
BRACT DEVELOPMENT IN *CORNUS*

ABSTRACT

The genus *Cornus* is best known for beautiful flowering dogwood trees that bear large petaloid bracts. Another group of species within *Cornus*, the dwarf dogwoods, and a close relative, *Davidia involucrata*, also bear petaloid bracts, raising the possibility that the complex to which these species belong might be genetically pre-adapted for petaloidy in bracts. I studied the morphological and anatomical development of both petaloid and non-petaloid bracts in the phylogenetic context of the genus *Cornus* in order to identify 1) any proto-petaloid states in non-petaloid taxa and 2) when and where genes responsible for petaloidy may be expressed. This information is used to design experiments to investigate the genetic basis of petaloidy in bracts of dogwoods.

INTRODUCTION

The family Cornaceae has been described as a prime case of taxonomic chaos" and "a most unruly family" (Eyde, 1987). Though the genus *Cornus* is obviously a natural group, the chaos has stemmed from unresolved relationships among the four clades within the genus and wide-ranging disputes as to the identity of *Cornus'* closest relatives at the family and ordinal levels. Much splitting by various authors, at times ignorant of previous work, has left each of the four lineages and divergent members of

these lineages with one or more generic names, for a total of at least 10 genera which correspond to *Cornus sensu lato* (I will use *Cornus* to refer to the genus *sensu lato*).

The four clades within the genus are the showy-bracted dogwoods (7 species of trees), the dwarf dogwoods (3 species of perennial sub-shrubs which also have showy bracts), the small-bracted cornelian cherries (6 species of trees/shrubs with small, non-showy bracts subtending the inflorescence) and the bractless dogwoods (40-50 species of shrubs with miniature bracts, if any, present on inflorescence branches). The first three clades, the bracted dogwoods, all have red fruit and may collectively be referred to as the red-fruited dogwoods, while the bractless dogwoods have blue or white fruits and are also referred to as the blue-fruited dogwoods.

Relationships within the genus

Showy-bracted dogwoods, small-bracted dogwoods, two lineages of the bractless dogwoods, the alternate-leaved dogwoods (sister taxa *C. alternifolia* and *C. controversa*) and the distinctive, "enigmatic" *C. oblonga*, have all been proposed at some time as primitive within the genus *Cornus* (reviewed in Xiang, 1996). Three authors have proposed comprehensive hypotheses of phylogenetic relationships of the genus based on 1) a cladistic analysis of morphological, chemical, and cytological characters (Murrell, 1993); 2) morphological, chemical, cytological and fossil evidence (Eyde, 1988); and 3) molecular data from chloroplast DNA restriction sites and chloroplast and nuclear gene sequences (Xiang, 1993; Xiang, 1996; Xiang, 1998; Fan, 2001).

The latter two analyses indicate a sister relationship between the bractless and the bracted dogwoods, and, within the bracted dogwoods, a sister relationship between the small-bracted dogwoods and the clade of showy-bracted and dwarf dogwoods (Figure I.1.A). Murrell's analysis reconstructs the small-bracted dogwoods and showy-bracted dogwoods as sister taxa, which together are sister to the dwarf dogwoods; all of these bracted dogwoods are sister to the bractless dogwoods, save one, the unusual *C. oblonga,* which appears as the most basal member of the genus, sister to all other dogwoods (Figure I.1.B).

The phylogenies of Eyde/Xiang and Murrell differ in the sister relationships of the bracted dogwoods and the identification of the most primitive members of the genus. Of most significance for this study, Eyde and Xiang's analyses unite the showy-bracted and dwarf dogwoods with the synapomorphy of showy bracts, thus positing that petaloid bracts evolved once in the genus, while Murrell's analysis unites the small-bracted dogwoods to the showy-bracted dogwoods by five synapomorphic inflorescence characters and posits that petaloid bracts evolved independently in the dwarf and showy-bracted lineages.

In the fossil record, the four clades of dogwoods are reliably distinguishable by their fruit-stones, and the order of earliest appearances of the distinct fruit-stones supports the hypothesis that the blue-line is ancestral within the genus and that small-bracted dogwoods gave rise to the showy-bracted dogwoods (trees). However, fossil evidence does not inform the debate about relationships among the bracted dogwoods because the fruit-stones of the dwarf dogwoods do not appear in the fossil record. Nor have definitive fossil showy-bracts have been identified in the fossil record, though Eyde cites Gregor's claim that impressions known as *Diospyros* calyxes from the Miocene may be the inflorescence bracts of dogwoods (Eyde, 1988).

Molecular data supports a single clade of showy-bracted dogwoods and single blue-line clade; in the combined nuclear and chloroplast parsimony analysis, bootstrap support values are 95% for the clade of showy-bracted plus dwarf dogwoods, 90% for the blue-line dogwoods minus *C. oblonga*, 80% for the entire lineage of blue-line dogwoods, and 100% for the entire genus. The sister relationship between the showy-bracted and dwarf dogwoods is strongly supported, and, as Murrell himself points out, the five inflorescence synapomorphies in his analysis may be developmentally correlated, thus weakening his phylogenetic hypothesis. Overall, evidence favors a single origin of petaloid bracts in the common ancestor of the showy-bracted and dwarf dogwoods, sometime before the first appearance of showy-bracted species in the Oligocene and likely in the Eocene, after the first appearance of the sister group, the small-bracted dogwoods.

The position of *C. oblonga* as the basal member of the genus (Murrell) or the basal member of the blue-line (Eyde) is equivocal. *C. oblonga* is not distinguishable from

other blue-line species in the fossil record, and molecular support for *C. oblonga* as a member of the blue-line clade is moderate. Both Eyde and Murrell employ adaptive reasoning to argue that its traits are primitive, and their arguments depend upon the outgroup chosen.

Close relatives of *Cornus*

Both Eyde and Murrell suggest the genus *Mastixia* as sister to *Cornus* based on one synapomorphy, two-armed hairs (malpigheaceous trichomes). Xiang tested 29 genera which have been at some time suggested as close relatives of *Cornus*, and her results suggest *Alangium* as the sister taxon to *Cornus* (99% bootstrap support in the *rbcL-matK* tree) and show that *Alangium, Curtisia,* the hydrangeoids and the nyssoids-mastixiods complex are all closely related to *Cornus* (100% bootstrap support for this clade in the *rbcL-matK* tree) (Xiang, 1998).

Notably, the nyssoid *Davidia* has a pair of white involucral petaloid bracts, and inflorescence bract *scars* are seen in the fossil nyssoid genus *Amersinia obtrullata*, though fossil inflorescence bracts have not been found (Manchester, 1999). The scars resemble those of *Davidia*, and four or five are found in a whorl subtending an inflorescence, just as in modern day showy-bracted dogwoods. Keeping in mind the evidence from fossils and outgroups that petaloid bracts evolved from non-petaloid bracts in *Cornus,* the occurrence of petaloid bracts in a modern-day relative and some type of bracts in a fossil relative begs an inquiry into a possible tendency toward petaloidy of bracts in these related taxa.

Petaloidy in *Cornus*

In order to approach the phenomenon of petaloidy, petals themselves must be considered. It is widely accepted that petals have evolved independently multiple times in angiosperms. In this regard, petals are generally divided into two groups, those believed to have been derived from sepals (bracteopetals) and those believed to have been derived from stamens (andropetals) (Takhtajan, 1991). While petals often resemble

stamens rather than sepals in their color, thickness, deciduousness and venation, Ronse de Craene argues for a bracteolar origin for the majority of eudicots petals, and considers gene expression to be a determiner of petaloidy but not of petal homology (Ronse de Craene, 2007). We hypothesize that the same may be true for petaloid bracts in Cornus, and examine the bracts of the species in this light.

METHODS AND MATERIALS

Materials

Seventeen species of *Cornus,* representing all four clades and including one hybrid and two sub-species, were observed, photographed and collected over the course of inflorescence development from bud to flower/fruit (Table I.1).

Scanning Electron Microscopy

Specimens for SEM were collected in 70% ethanol and gradually transitioned into pure acetone with ethanol:acetone washes. The materials were critical-point dried using liquid CO2 in a Balzers CPD 030 and mounted onto aluminum stubs with slow-drying araldite. They were coated in a Cressington sputter coater with 20nm of gold-palladium and observed under high-vacuum with a Philips XL30 at The Natural History Museum, London.

RESULTS

Bract Development

Photographic series of inflorescence development of five showy-bracted and two small-bracted and various bractless species are shown in Figures I.2-I.8.

In the showy-bracted dogwoods, bracts are adnate to the floral receptacle. *C. florida* f. *pluribracteata* exhibits bract morphology characteristic of the American showy-bracted dogwoods (*C. florida*, *C. nuttallii, C. disciflora)* in which the bracts serve a protective function over winter. In *C. florida* and *C. nuttallii*, the bracts expand from their bases in the spring, while the in *C. disciflora* they fall off soon after opening. The portion of the bract that is exposed over winter retains its curved shape and dark color (Figure I.2.A-D) while new growth occurs at the base of the bract and turns from green (Figure I.2.A-E) to cream-colored (Figure I.2.F-G) to white (Figure I.2.H). In the case of *C. florida*, this growth pattern results in the iconic notched appearance of the bracts (Figure I.2.F-G), which number four. In the case of cultivated 'doubles' such as *C. florida* f. *pluribracteata*, additional inflorescence bracts may be present (Figure I.2.H), though they are often much smaller than the four primary bracts. Mature bracts of *C. florida* are UV dark, that is, they absorb rather than reflect UV rays, as is typical for petals.

In *C. nuttallii*, the six to eight protective inflorescence bracts do not fully enclose the flower buds over winter and are less curved than those of *C. florida*. Consequently, the transition from new-growth to old-growth in the mature bracts is less pronounced, as seen in the hybrid *C, nuttallii* x *florida*, which shows more characteristics of *C. nuttallii* than *C. florida* (Figure I.3.A-H).

Though the four caducous bracts of *C. disciflora* are not showy, they have been reported as off-white in color, possibly indicating some degree of petaloidy (Gonzalez Villarreal, 1996). I observed *C. disciflora* in flower in Costa Rica in July 2003 and observed no petaloid features in bracts, which were green with a whitish bloom, at best (Figure I.3.J).

Unlike the inflorescence bracts of American showy-bracted dogwoods, inflorescence bracts of Asian showy-bracted dogwoods are not exposed over winter. In *C. kousa* var. *chinensis*, two pairs of external bracts cover the inflorescence bracts and two pairs of leaves. The outer pair of external bracts separate from the branch as unit (Figures I.4.B). The inner pair become inflated and turn from brown to yellowish before falling off (Figures I.4.C, D). As the inflorescence bracts expand, they turn color at the tips from pale yellow with red margins to white, and they turn color in the body from

green to cream-colored to white (Figures I.4.E-H). A similar pattern of inflorescence bract development is seen in *C. kousa* (Figures I.5.A-F).

Of the three species of dwarf dogwoods, *C. canadensis* and *C. suecica* differ with respect to petal color, sepal shape, and leaf characteristics but not bract characters (Murrell, 1994), and *C. unalaschenkis* is a hybrid of the other two. *C. canadensis* overwinters underground, and when new shoots appear in the spring, the inflorescences are not apparent until the leaves unfurl, revealing four leaf-like inflorescence bracts which turn from green to cream-colored to white as they expand (FigureI.5.G-J). The flowers are born on minute peduncles inserted at the base of each bract; there is no fused receptacle like that found in the showy-bracted and small-bracted dogwoods.

In the showy-bracted dogwoods and in the dwarf dogwoods, bracts open while flowers are in bud, and maturation of bracts coincides with the beginning of the blooming of flowers, which open sequentially over several weeks. In the small-bracted dogwoods, flowers and bracts open concurrently. The small-bracted dogwoods are represented in this study by two species, *C. officinalis* and *C. chinensis*. In both species four inflorescence bracts protect the flower buds over winter (Figures I.6.A and I.7.A). As in *C. florida*, they open in the spring before the leaves emerge, and two stripes of lighter color corresponding to the parts of the bract that were covered are apparent on the abaxial surface of each of the inner bracts (Figures I.6.B and I.7.B). The adaxial surfaces of the inflorescence bracts are light green in color (Figures I.6.E and I.7.D), and may contribute to attraction of pollinators along with petals. The bracts reflex as the yellow flower buds open and remain on the tree as the leaves emerge (Figures I.6.E and I.7.D) and often until fruit set (Figure I.6.E). As in the showy-bracted dogwoods (not shown), the following year's inflorescence buds are visible in the preceding autumn (Figures I.6.H and I.7.E).

In "bractless" dogwoods, miniscule bracts may be present on inflorescence branches, as shown in *C. rugosa*, *C. purpusii*, and *C. oblonga*, a putative basal lineage in the genus. In *C. rugosa* the tiny bracts are often ephemeral (Figure I.8.A) while in *C. purpsii* they are large and strikingly leaf-like (Figures I.8.H), and *C. oblonga* they are slightly larger than in most bractless species and are persistent (Figures I.8.J-K).

Epidermal morphology of bracts

As showy inflorescence bracts develop, their epidermal cells undergo changes in morphology. Young bracts of small-bracted species are covered in trichomes on their abaxial side, as seen in *C. chinensis* (Figure I.12.C), indicative of their protective function. The same is true for young bracts of showy-bracted species, even when the bracts are covered by outer bracts in the inflorescence bud, as is the case of *C. kousa* var. *chinensis* (Figure I.9.A). The adaxial surfaces of young bracts in both groups have fewer trichomes than their abaxial surfaces, although the degree of adaxial trichome coverage in the small-bracted species (I.12.B) is much greater than that in the showy-bracted species (Figure I.9.B).

Epidermal cells visible on the adaxial surface of the young bract of *C. kousa* var. *chinensis* are flat and simple in shape (Figure 1.9.B), in contrast to epidermal cells on the adaxial surface of mature bracts which are papillose (Figure I.10.A, foreground, and I.10.B). On the abaxial surfaces of mature bracts the epidermal cells are slightly inflated but not papillose and have relatively complicated cell borders, in a jigsaw style typical of leaf epidermal cells (Figure 1.10.A., midground). The abaxial surfaces of the tips of the bracts, however, are an exception and have slightly papillose epidermal cells. Trichome cover continues to be greater in the abaxial than adaxial surface on mature bracts.

The epidermal morphology of inflorescence bracts in the dwarf dogwoods is similar to that of the showy-bracted dogwoods. Young bracts have a light cover of trichomes on the abaxial surface (Figure I.11.A) and the mature bracts have papillose cells on the adaxial surface and flatter cells with relatively simple borders and stomata on the abaxial surface (Figure I.11.B-C).

The adaxial surface of petals are generally papillose, but cells at the base tend to be flatter (Figure I.10.D-E). Ridges are seen on the surfaces of many cells, both papillose and flat, from bract, petal, and anther (Figures I.10.B, D-E, I.11.B-C, and I.12.A-B,D).

Teratologous forms

While studying various *Cornus* species at The Royal Botanical Gardens, Kew, two types of teratologies relevant to bract development were observed. In *Cornus nuttallii* x *florida* in 2005 a number of inflorescences exhibited leaves inserted into the floral head (Figure I.13.A-E). The leaves appear to be surrounded by sepals (Figure I.13.A, E). Also, in *C. florida* f. *pluribracteata,* one petaloid inflorescence bract with a green leaf-like patch several cm square was observed (Figure I.13.F, G).

DISCUSSION

Species representing all four clades of the genus *Cornus* were observed in light of the hypothesis that the genus may have been pre-adapted for petaloidy in bracts. We did not find clear morphological evidence for pre-adaption; most notably, papillose cells, a common characterstic of petals, were not found on the small, non-petaloid bracts of the cornelian cherries examined. However, we did find numerous large, leafy bracts in one species of bractless dogwoods, *C. purpusii*, which may represent a parallel adaptation of bracts subtending the inflorescence head.

Pre-adaptation for petaloidy is most likely to be detected at the genetic level. The morphological and anatomical observations made in this study will aid in the investigation of the genetic basis of petaloid bracts in several ways. The first issue is that of the timing of gene expression. Inflorescence bracts of both showy-bracted and small-bracted dogwoods form in the autumn preceeding spring flowering; genes responsible for petaloidy may be expressed anytime between formation and maturation of bracts, and RNAs or, more likely protein products, may be present in bracts months before changes become visible. Presumably one or more genes link petaloid development of bracts to environmental cues and/or floral development such that bracts reach the mature petaloid state when flowers begin to bloom. Where genes are expressed within bracts is also significant. In the American showy-bracted dogwoods like *C. florida,* gene expression directing petaloidy would seem to be restricted to the base of the bract, where growth

occurs, whereas in the Asian showy-bracted dogwoods like *C. kousa*, gene expression would seem to be uniform throughout the bract. *C. disciflora* is a natural experiment within the showy-bracted clade; it has apparently lost petaloidy and expression of the petal identity genes may have been lost in bracts, or bract-specific copies, if they exist, may have become pseudogenes.

When studying showy-bracted species for gene expression, it will be important to collect entire bracts, which are clearly demarcated from receptacle tissue, and to collect bracts early in the season. Ideally bracts should be sampled when they are first recognizable in the autumn and also in late winter, before they have undergone any visible transition to petaloidy. Bracts in the small-bracted group should be collected at similar times for comparison. Since bracts in *C. canadensis* form only in the spring, they should be collected when green and large enough to handle. There is no clear boundary between bract and receptacle tissue in *C. canadensis*, so care must be taken not to include any floral tissue in the collection of bracts.

Comparing gene expression in bracts of the showy-bracted and dwarf dogwoods to each other and to homologous genes and gene expression in outgroups should illuminate both the genetic basis/bases for petaloidy and the number of evolutionary origins of petaloid bracts in the genus. In the study of petaloid bracts, the small-bracted dogwoods serve as the outgroup to the showy-bracted + dwarf dogwoods, and the bractless dogwoods serve as an outgroup to the bracted dogwoods. It may be that some of the genetic changes associated with petaloid bracts actually occurred in the bractless dogwoods or the common ancestor of dogwoods and its petaloid-bracted relative the nyssoid *Davidia*. If the transition to petaloidy were cumulative, that is if several genes were turned on independently over evolutionary time, then expression of some genes may be maintained today in taxa without petaloid bracts. On the other hand, if a gene is acting as a master switch, we would not expect to see ectopic expression of petal-identity genes in non-petaloid taxa; however, we might discover pre-adaptation for switching to a petaloid state in the gene acting as a master switch.

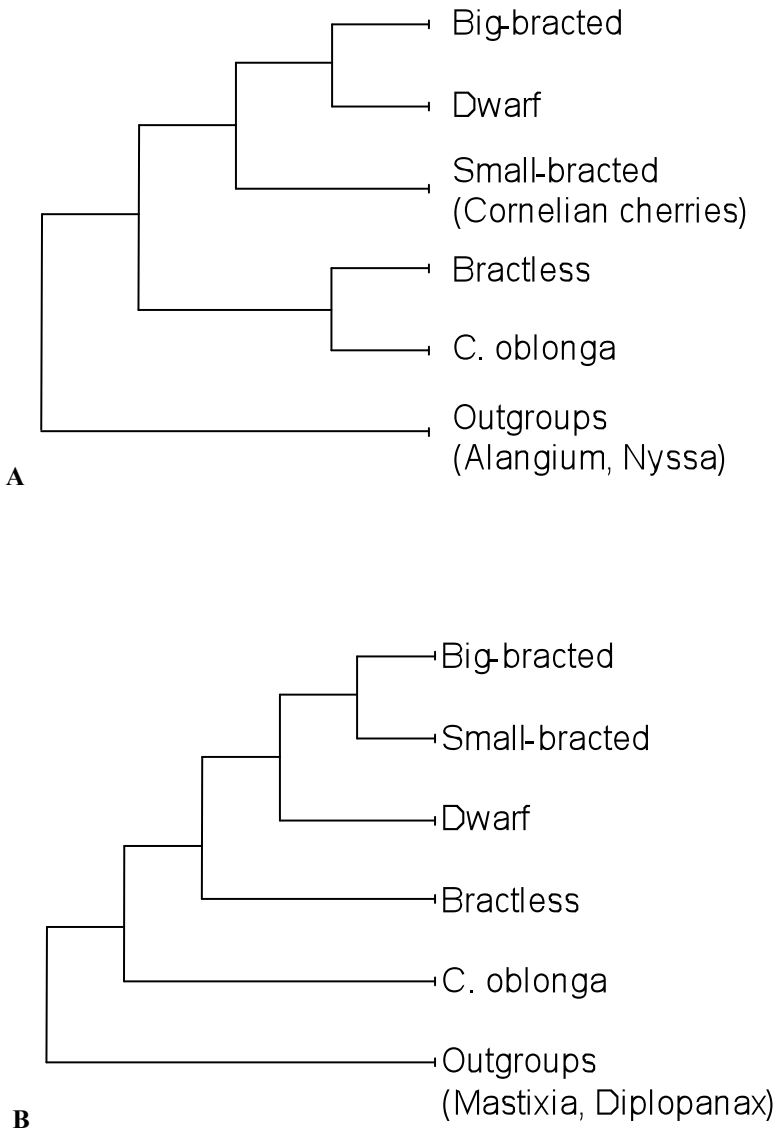


Figure I.1 (redrawn from Fan 2001): Putative relationships among subgroups of *Cornus*. A) The most parsimonious tree from analysis of 26S rDNA and cpDNA by Fan et al. (2001). B) the phylogenetic tree derived from cladistic analysis of 28 morphological, chemical, and cytological characters by Murrell (1993).

Figure I.2. Showy-bracted *C. florida* f. *pluribracteata* at RBG Kew. After winter inflorescence bracts have a grey bloom (A). In spring, bracts expand from the base while the tips retain their curved over-wintering shape (A-D). The bracts turn from green (B-E) to cream-colored (F, G) to white, and wilt after flowering (H). Habit is a small tree (I).

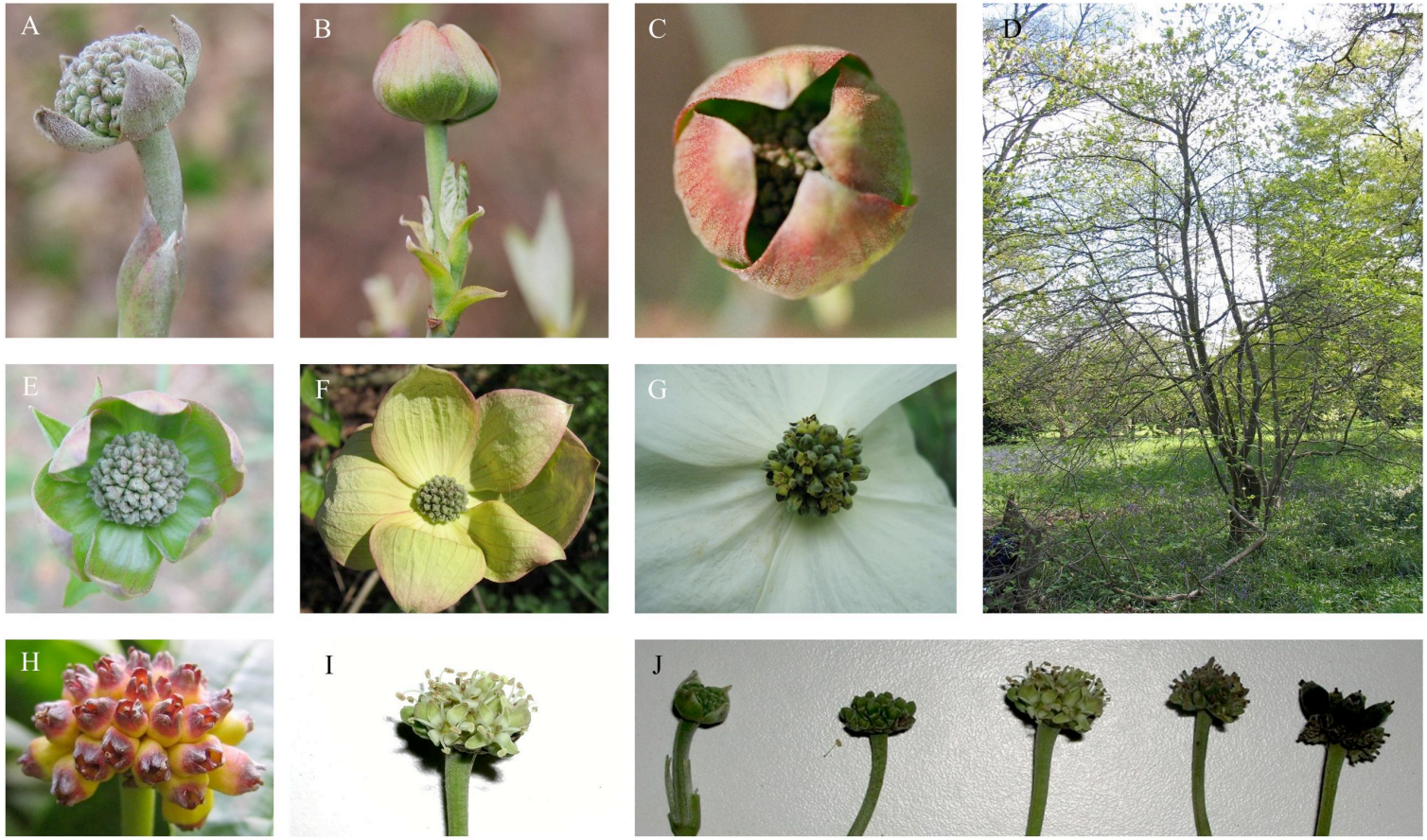


Figure I.3. Showy-bracted *C. nuttallii* x *florida* (A-H) at RBG Kew and *C. disciflora* (I, J) at Cuereci, Costa Rica. *Cornus nuttallii* x *florida* plants are said to display more characteristics of *C. nuttallii* than of *C. florida.* Habit is a medium sized tree, larger than that of *C. florida* (D). Bracts usually number 4-6 and do not enclose the flower buds over winter (A). In spring, bracts expand from the base (B). The tips remain red but do not display the curved shape of *C. florida* bracts. Like in *C. florida*, bracts turn from green (B-E) to cream–colored (F) and then to white (G) as flowers bloom. Flower buds swell and turn yellow after flowering, but no fruits were observed. Inflorescence development (I) and flowers (G) of *C. disciflora.* *C. disciflora* belongs to the clade of showy-bracted dogwoods but bracts do not expand and fall off before flowers bloom. Habit (not shown) is a large tree.

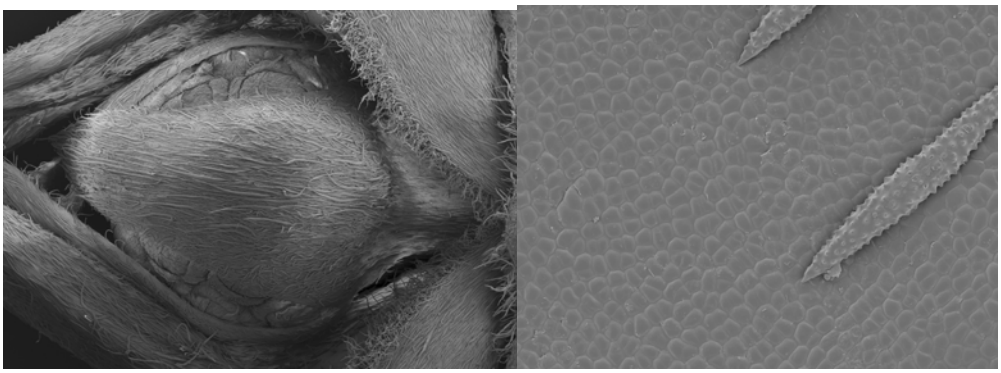
Figure I.4. Showy-bracted *C. kousa* var. *chinensis* at RBG Kew. Inflorescence buds are first visible in autumn (A). In spring, two pairs of protective bracts fall off. The outer pair dry and remain fused at the tips (B). The inner pair (white arrow, C) open and swell (white arrow. D) before falling off. Two of the inflorescence bracts are exposed slightly after the first pair of outer bracts falls off (black arrow, C). Inflorescence bracts expand and turn from green to cream-colored to white (black arrow, D, and E-H). Dried inflorescence head (I); no fruit was observed on this tree from 2004-2007 . Habit is a low growing tree (J).

Figure I.5. Showy-bracted *C. kousa* at RBG Kew and showy-bracted dwarf *C. canadensis* at The Chelsea Physic Garden,. Inflorescence development of *C. kousa* (A-F) is very similar to that of *C. kousa* var. *chinensis* described in Figure I.4. Inflorescence bracts of *C. canadensis* are enclosed in rolled leaves underground and do not emerge until late spring (when growing in London, U.K.). Bracts turn from leaf-like green (arrow, G) to cream (H, I) to white as flowers bloom (J). Habit is creeping groundcover, shown in a garden environment (I).

Figure I.6. Small-bracted *C. officinalis* and *C. mas* at RBG Kew. Inflorescence buds of *C. officinalis* overwinter (A) and open in early spring, revealing a cyme of bright yellow flowers (B-D). Inflorescence bracts retract but do not fall off, often remaining on the plant until after fruit set (arrows, E, F). Habit is a shrub (I). *C. mas* exhibits very similar morphology and phenology. Fruit set (G) and autumnal inflorescence buds (H) are shown.

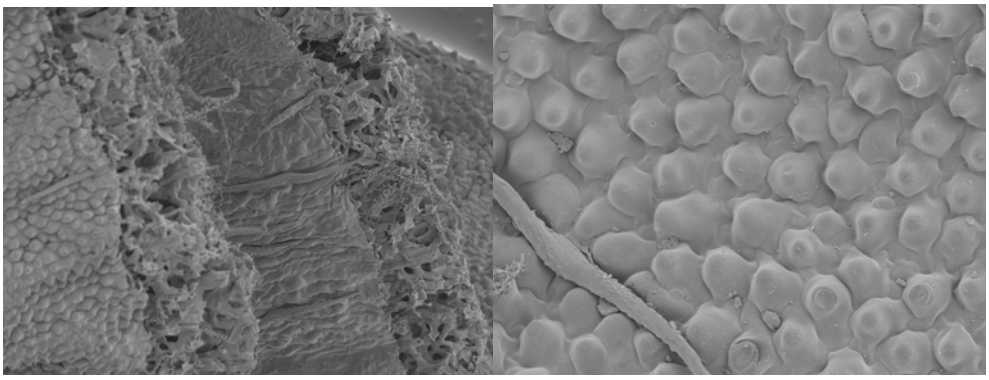
Figure I.7. Small-bracted *C. chinensis* at RBG Kew. Inflorescence buds are resupinate as compared to other bracted dogwoods. Bracts bear a grey bloom after winter (A) and open in spring (B, C), revealing a cyme of bright-yellow flowers. Bracts remain on the plant until after flowering (D). New inflorescence buds are visible in autumn (E). Habit is a shrub/small tree (F).

Figure I.8. Bractless dogwoods at RBG Kew. Rudimentary bracts are sometimes observed, as in *C. rugosa* (arrow, A) and *C. oblonga* (arrows I, J). *C. purpusii* bears 6-8 leafy green bracts at the base of the inflorescence (H). Inflorescences are panicles of white flowers: *C. rugosa* (B), *C. ammomum* 'Grandiflora' (C), *C stolonifera* (D, E) and *C. oblonga* (K). Habits are shrubs, *C. racemosa* (F), *C. purpusii* (G) and *C. oblonga* (L), which bears leathery evergreen leaves.

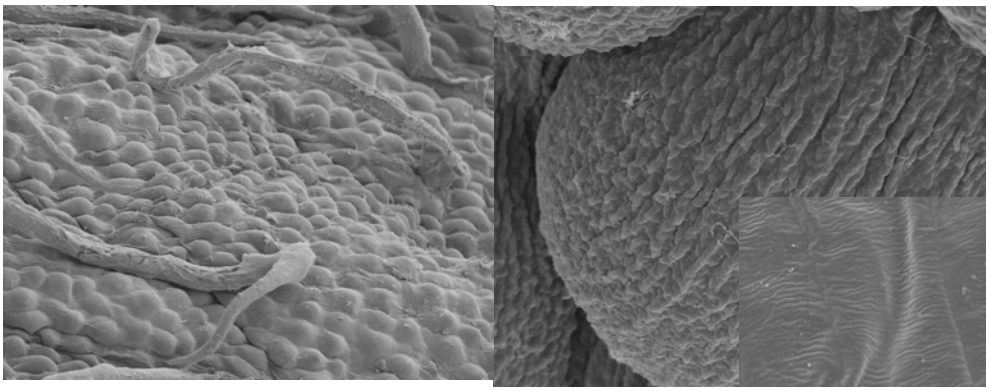


A **B**

Figure I.9. *C. kousa* var. *chinensis*, SEM. A) young inflorescence 5x. B) young bract adaxial surface 100x.

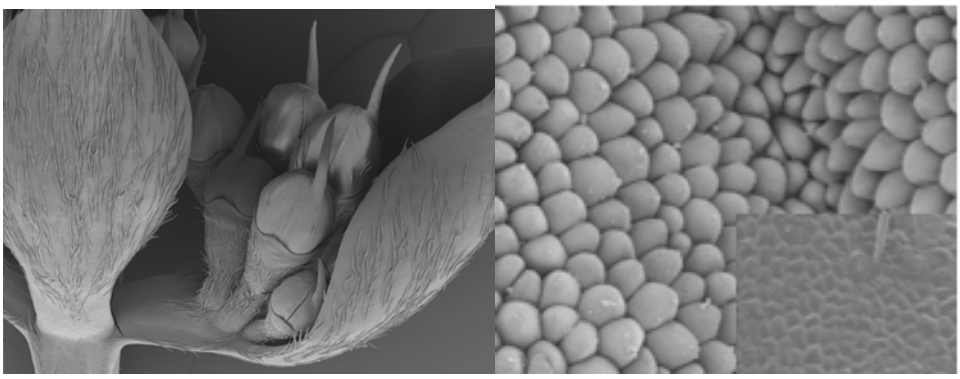


A **B**



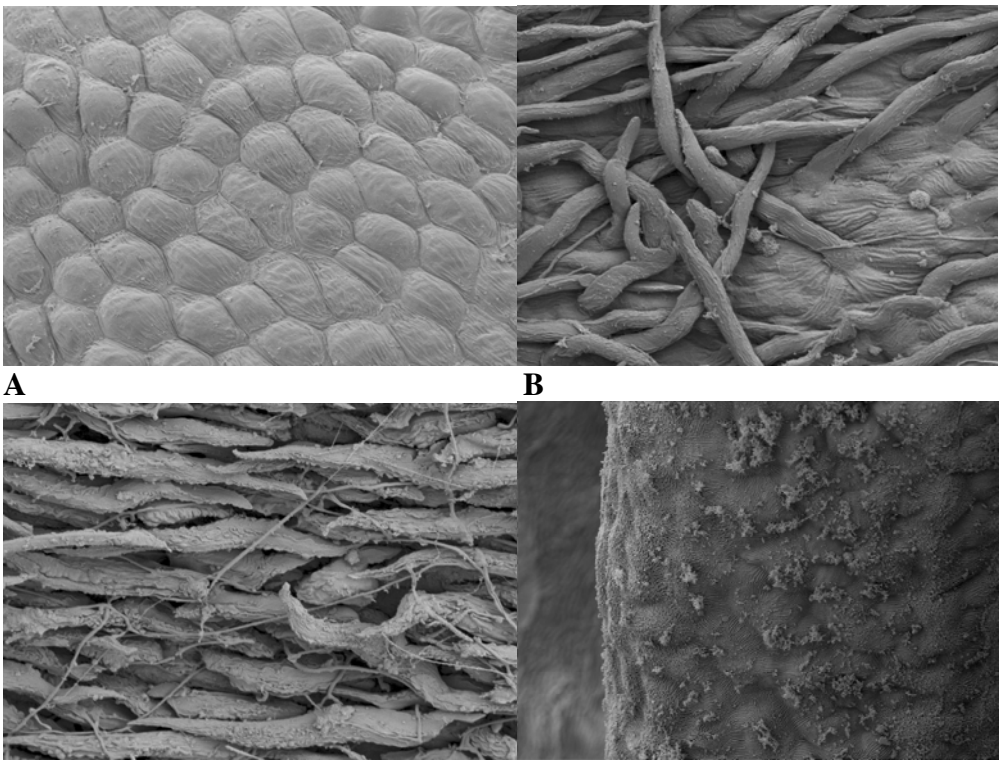
C **D** **E**

Figure I.10. *C. florida* f. *pluribracteata*, SEM. A) torn mature bract showing adaxial surface in foreground and background and abaxial surface in midground, 30x. B) mature bract adaxial surface 100x. C) mature bract adaxial surface at tip 100x. D) Petal adaxial surface, papilose cell, 1000x. E) Petal, abaxial suface, cell at base of petal,1000x.



A **B** **C**

Figure I.11. *C. canadensis*, SEM. A) young inflorescence, 5x. B) mature bract, adaxial surface 100x. C) mature bract abaxial surface 100x.



A **B**

C **D**

Figure I.12. *C. officinalis* and *C. chinensis*, SEM. A) *C. officinalis* petal 100x. B-D) *C. chinenesis* B) bract adaxial surface 100x. C) bract abaxial surface 100x. D) anther 130x.

Fig I.13. Unusual forms observed in flowering dogwoods. Approximately ten percent of inflorescences born on *C. nuttallii* x *florida* at RBG Kew during 2005 had leaves inserted in the floral head (A-E). Leaves appear to arise from among sepals (A, E). One case of partially green mature inflorescence bract was observed in *C. florida* var. *pluribracteata* in 2007 (F, G).

Species	Group	Native Range	Collection Locale	Specimen
C. kousa	showy	Japan, China	Taipei, Taiwan	n/a
C. kousa	showy	Japan, China	RBG Kew	1989-8288
C. kousa var. *chinensis*	showy	China	RBG Kew	1969-12600
C. florida	showy	Eastern North America	Ann Arbor, MI	n/a
C. florida f. *pluribracteata*	showy	horticultural form	RBG Kew	1969-13684
C. nuttallii x *florida*	showy	horticultural hybrid	RBG Kew	1925-74102
C. disciflora	showy	Central America	Costa Rica	voucher at MICH
C. canadensis	dwarf	Northern North America	Chelsea Physic Garden, London	none
C. chinensis	small	China	RBG Kew	1988-8692
C. officinalis	small	Japan	RBG Kew	1992-1476
C. mas	small	Southern Europe	UM Arboretum	(checking)
C. mas	small	Southern Europe	RGG Kew	1979-4521
C. alba	bractless	Siberia, Northern China	UM Botanical Gardens	n/a
C. amomum 'Grandiflora'	bractless	Eastern North America	RBG Kew	1895-1502
C. oblonga	bractless	Himalya	RBG Kew	unavailable
C. purpusii	bractless	Eastern North America	RBG Kew	1985-8414
C. racemosa	bractless	Eastern North America	RBG Kew	1968-23506
C. rugosa	bractless	Northern North America	RBG Kew	1987-647
C. stolonifera	bractless	North America	RBG Kew	1914-58102

Table I.1. Seventeen species of *Cornus* were studied. The four groups within the genus are the showy-bracted dogwoods, the dwarf dogwoods, the small-bracted dogwoods, also called cornelian cherries, and the bractless dogwoods. Collections were made between 2003 and 2007. The Chelsea Physic Garden and The Royal Botanical Gardens, Kew are in the U.K.

CHAPTER II
PETAL ORGAN IDENTITY GENES IN *CORNUS*

ABSTRACT

Since the discovery of MADS-box floral organ identity genes in model organisms, the molecular evolution of this gene family has been investigated and implicated in floral evolution, especially in perianth diversity. While the genus *Cornus* exhibits fairly standard asterid floral morphology, two lineages with in the genus and a close relative outside of the genus bear petaloid inflorescence bracts. Here we identify A, B, and E class MADS-box genes from across the genus *Cornus* and investigate their evolution in light of their possible roles in the evolution of these petaloid organs.

INTRODUCTION

Research on floral development in two angiosperm species possessing typical, classically defined flowers, *Arabidopsis thaliana* and *Antirrhinum majus*, inspired the ABC model of floral development (Bowman et al., 1991; Coen and Meyerowitz, 1991, Meyerowitz et al., 1991) in which the expression of three classes of genes, A, B, and C, determine organ identity of the four major floral organ whorls within the flower. Class A genes specify sepals, class A + class B specify petals, class B + class C specify stamens, and class C genes specify carpels. The original ABC model has been revised to include Class D and E genes; Class D genes control ovule development (Angenent et al., 1995; Rounsley et al., 1995; Colombo et al., 1995) and Class E genes are required for organ identity of all floral organs (Pelaz et al., 2000; Honma and Goto, 2001; Ditta et al., 2004).

In addition, the original model is under reconsideration with regards to A function. In plants other than *Arabidopsis*, class A genes are present but may not fulfill the canonical A function of specifying sepals and petals; A class genes may be best understood as genes whose principal role in floral meristem identity sometimes leads to a role in sepal organ identity and very rarely to a role in petal organ identity (Litt, 2007).

The A class gene are represented by*SQUAMOSA (SQUA)* and *LIPLESS* (*LIP1/2*)in *Antirrhinum* and *APETALA1 (AP1)* and *APETALA 2 (AP2)*in *Arabidopsis*. B class floral organ identity genes are represented by *DEFICIENS(DEF)* and *GLOBOSA (GLO)* from *Antirrhinum* and *APETALA (AP3)* and *PISTILLATA(PI)* from *Arabidopsis*, and class E genes are known from *Arabidopsis* as *SEPALLATA (SEP1/2/3/4)* and by homology to *SEP* as *DEFH200* and *DEFH72* from *Antirrhinum.*

Floral organ identity genes are MADS-box genes

With the exception of *AP2*, the A, B, C, and E class genes belong to the pan-eukaryotic MADS-box gene family of transcription factors. These floral organ identity genes belong to the large majority of plant MADS-box genes which are classified as Type II based on their phylogenetic position (Alvarez-Buylla et al., 2000) and have a MIKC protein structure (Mandel et al., 1992); the gene structure consists of a DNA-binding MADS-box (M), an intervening domain (I), a Keratin-like domain (K) and a C-terminal C-domain.

The N-terminal region encodes the highly conserved MADS-box domain of about 60 amino acids which is involved in DNA binding and protein dimerization (Pollock, 1991). The MADS-box domain binds a conserved DNA sequence called a CArG box ($CC(A/T)_6GG$) (Schwarz-Sommer et al. 1992), and the proteins bind DNA as dimers, or as higher-order complexes (Theissen and Saedler, 2001).

The I domain is moderately conserved and is involved in protein dimerization (Riechmann et al., 1996). In the B class genes, the I region is about 30 amino acids, while in the A, C and E class floral organ identity genes it is approximately 35 to 37 amino acids long. The highly conserved K domain of approximately 70 amino acids is named for its significant similarity to a portion of the sequence of keratin and is predicted to

form two to three amphiphatic αhelices which facilitate dimerization. This domain is unique to plant Type II MADS-box proteins, first appearing in Charophytes (Kaufmann et al., 2005).

The C domain of about 80 amino acids is the least conserved domain overall but contains small stretches of sequence that are highly conserved (Johansen et al., 2002; Vandenbussche et al., 2003). The C domain is involved in quaternary protein interactions (Egea-Cortines et al., 1999) and transcriptional activation (Riechmann et al., 1997).

As is typical for regulatory genes, the distinct domains of plant MADS-box genes experience different rates of nonsynonomous substitution. The rate for the MADS-box itself is ~3 x 10^{-10} substitutions/ site/ year while the rate for the K box is 1 x 10^{-9} substitutions/ site/ year.The C domain rate is the greatest of all at 3.7 x 10^{-9} substitutions/ site/ year, and accordingly, diversification in the K-box and C-terminal domains accounts for significant functional divergence between and within the major groups of MADS-box proteins (Purugganan et al., 1995).

Divergence and duplication among the floral organ identity genes

Phylogenetic analyses of plant Type II MADS-box genes show frequent gene duplication, both recent and in the distant past (Kramer and Jaramillo, 2005). Ancient duplication and divergence has produced fourteen major clades of Type II MADS-box genes in plants, and clades generally reflect the developmental roles of their members (Theissen et al., 2000; Becker and Theissen, 2003) (Figure II.1). The monophyletic *DEF/AP3* and *GLO/PI* clades together constitute the B gene clade, which forms a larger clade with a group of genes known as B-sister and relatives. The *AP1/ SQUA* (A gene) clade is sister to the *SEP* (E gene) + *AGL6* clade. MADS-box genes belonging to other clades control the transition to flowering (*FLC, AGL27* in the FLC clade), lateral root formation (*ANR1* in the AGL17 clade), and fruit dehiscence zone development (*SHP* in the *AG* clade), for example (cited in Becker and Theissen, 2003).

MICK MADS-box genes have been isolated from non-seed plants including the fern *Ceratopteris richardii* (Hasebe et al., 1998), the clubmoss *Lycopodium annotinum* (Svensson et al., 2000) and the bryophyte *Physcomitrella patens* (Krogan and Ashton,

2000), but none of these genes appear to be orthologs of any of the recognized seed plant MADS-box gene subfamilies. Homologs of the A, B, and E class genes have been isolated from hundreds of taxa representing major groups of angiosperms and, in the case of the B genes, from gymnosperms. The presence of B genes in gymnosperms and their absence in free-sporing plants, (e.g. ferns, lycopods, and mosses) supports the hypothesis that the origin of the B gene clade occurred in a common ancestor of seed plants after its separation from lineages leading to free-sporing plants. A rough estimated divergence time of 340 mya for the major MADS-box gene subfamilies in angiosperms (based on calculated mean substitution rate of 9 $x10^{-10}$ substitutions/ site/ year and mean level of nucleotide substitution between groups of 0.6105) (Purugganan et al., 1995) also places the origin of the B genes in the ancestor to seed plants.

Phylogenetic analyses of representative gene sequences show multiple duplication and divergence events within the B gene lineage (Kramer et al., 1998; Kim et al., 2004; Stellari et al., 2004). Many duplication events are relatively recent, but two ancestral duplication events stand out. The first gave rise to the *PI* and *AP3* lineages sometime after the split between extant gymnosperms and angiosperms and before the diversification of extant angiosperms; the second occurred within the *AP3* lineage in the "fringe" eudicots and resulted in two paralogous lineages found widely among the core eudicots. These lineages have been named eu*AP3* and *TM6,* the latter for the tomato *AP3* paralog *TOMATO MADS 6*, of which no ortholog is present in the complete genome of *Arabidopsis.* The function of *TM6* in Tomato is unknown; in *Petunia hybrida* it is involved in the specification of stamens but not petals (Rijpkema et al., 2006).

Like the B genes, the A genes comprise three major lineages, called *AP1, FRUITFUL (FUL),* and *AGL79*, which resulted from two ancient duplications (Litt and Irish, 2003; Shan et al., 2007). No orthologs of A genes are known outside of angiosperms, and both duplication events date to the period shortly before the diversification of the core eudicots, but the relationships among the three lineages are not resolved (Shan et al., 2007). *Arabidopsis FUL* is involved in floral meristem identity, like *Arabidopsis AP1*, and also in fruit and leaf development (Gu et al., 1998). The function of *Arabidopsis AGL79* is unknown.

The E genes appear to be restricted to angiosperms, though the sister lineage to the *SEP* genes, the *AGL6* lineage, is found in gymnosperms (Becker and Theissen, 2003). Two major clades of E genes have been identified, the *AGL9 (SEP3)* group and the *AGL2/3/4 (SEP1/4/2)* group, which originate before the diversification of extant angiosperms. Additional duplications occurred within the monocots and within the dicots. At least two duplications occurred in the MADS-box gene duplication hotspot of the fringe eudicots, resulting in the *AGL2* clade (containing *Arabidopsis SEP1* and *SEP2*), the *AGL3* clade (containing *ArabidopsisSEP4*) and the *FBP9* clade, named after a gene from *Petunia* (Zhan et al., 2004).

The lineages of A, B and E genes, and also C genes, bear two distinct sets of signature C-terminal amino acid motifs which contain hydrophobic and polar residues and are found in homologous positions along the protein sequences, indicating their possible functional importance. In B genes the paleoAP3 motif is found in *TM6* genes, non-eudicot *AP3* genes, and even pre-angiosperm B and B-sister class genes (Münster et al., 1997, Sundström et al., 1999, Becker et al., 2002). In the position of the paleoAP3 motif, *euAP3* genes from the core eudicots have a modified "euAP3" motif, which appears to be a frameshift mutation resulting from a one basepair deletion (Kramer et al., 2006). The paleo/euAP3 motif is entirely absent in *PI* genes. *PI* genes have the separate "PI motif", a strongly hydrophobic domain of 16 amino acids. The PI motif is also present, albeit less conserved, as the "PI-derived" motif in *AP3* genes.

The two C-terminal motifs of the *AP1* gene clade are referred to as the FUL motif (Shan et al., 2007) and the paleoAP1/euAP1 motif (Litt and Irish, 2003; Vandenbussche et al., 2003). The former is present in all A class genes while the latter is lineage-specific; *euAP1* (i.e. core eudicot *AP1*) genes bear the euAP1 motif while *FUL-like* (i.e. *paleoFUL*), *euFUL* and (*eu)AGL79* genes bear the paleoFUL motif. Again, the two distinct versions of a motif seem to be related to each other by a translational frameshift mutation, in this case a deletion prior to the paleoAP1 motif in a *FUL* or *euFUL* gene (Shan et al., 2007).

In E genes the two motifs are simply referred to as SEP I and SEP II motifs, both of which are present in all clades of the lineage (Zhan et al., 2004), with the exception of

some clades of grasses which have lost the second motif due to a frameshift mutation (Vandenbussche et al., 2003).

The divergence and duplication events in early eudicots resulted in lineages that retained ancestral C-terminal motifs (i.e., *TM6, euFUL*) and lineages with novel motifs (i.e., *euAP3, euAP1*). Because these frameshift mutations appear to have been almost immediately conserved, it seems that they endow a gene with divergent function, which may lie in a novel capacity for interaction with other molecules. It is widely noted that the origins of divergent gene lineages in pre-angiosperms and in the early eudicots correspond to the appearance of petals in the first angiosperms and to *de novo* petal evolution at the base of the higher eudicots, respectively; novel MADS-box protein-protein interactions, due to changes in M, I and K domains as well as changes in C-terminal motifs, are thought to play a key role in the specification of these novel organs (Kramer et al., 1998; Kaufmann et al., 2005; Soltis et al., 2007).

The genus *Cornus* can also be considered to exhibit *de novo* petal evolution, of a sort. Petals and the genetic architecture which underlies them are present in the genus and its relatives; yet petaloid bracts of showy and dwarf dogwoods are clearly morphologically novel organs and may also be genetically novel. Here we have identified expression of A, B and E class candidate genes in *Cornus* as a first step toward investigating the genetic basis of petaloid bracts in dogwoods.

METHODS AND MATERIALS

Plant materials and gene cloning

Inflorescences of *C. florida*, *C. kousa*, *C. canadensis*, *C. mas*, *C. officinalis*, *C. alba* and *C. stolonifera*, representing all four groups of dogwoods were collected over the course of inflorescence development from locations listed in Table I.1. Inflorescences were separated into bracts, receptacles and flowers/flower buds before being immersed in liquid nitrogen. These tissues were stored at –80ºC for periods up to four years before isolation of RNA. Tissues were also collected in RNAlater (Ambion of Applied

Biosystems, Foster City, CA, USA). These tissues were stored at –20ºC and used within one year of collection.

Total RNA was extracted separately from bracts and flowers using one or more of three methods: the Pine tree method (Chang et al., 1993), TRI reagent (Sigma), or the RNAqueous Midi kit (Ambion). RNA quantity and quality was assessed on the Bioanalyzer (Agilent, Santa Clara, CA, USA) to determine the efficacy of the RNA isolation method used and integrity of the RNA. RNA quality was inferred from the ratio of peak heights of 18S and 28S ribosomal RNA, with ratios between 1.5 and 2 being acceptable. rRNA peak shape was also used to asses quality, with narrow peaks indicating high integrity of RNA and peaks with wide bases indicating degradation of RNA. First strand cDNA was synthesized by Superscript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) with the Poly (T) primer (5'-CCGGATCTCTAGACGGCCGC$(T)_{17}$-3'). MADS-box genes were initially amplified from *C. florida* by PCR using nested degenerate primers, as given in Appendix II.A. In other species genes were amplified with both degenerate primers and primers designed from sequences previously obtained.

PCR products were screened on agarose gels and directly sequenced or cloned with the TA cloning kit (Invitrogen). Between five and ten colonies were screened for each PCR product cloned, and all positive clones were sequenced. Products of at least two PCR reactions were sequenced for targeted genes. 5’ RACE (rapid amplification of cDNA ends) was used to obtain the 5’ coding sequences of *CkPI, CkAP3, CkTM6, CkFUL-A, CkAP1,* and *CfPI*.

Sequence identification

Sequences were identified by conserved 3' motifs (Vandenbussche et al., 2003), by BlastN (Altschul, 1990), and by phylogenetic analysis. For phylogenetic analysis, experimentally derived cDNA sequences, sequences of a close relative *Hydrangea macrophylla* and reference cDNA sequenes representing the major clades of plant MADS-box genes were first aligned by MUSCLE (Edgar, 2004) and adjusted manually in MacClade (Maddison, 2000) according to amino acid sequence. The divergent C-

terminal ends were not alignable. Phylogenetic analysis was performed on the M-I-K and intial C regions of the genes using maximum parsimony in PAUP version 4.0 Beta (Swofford, 2002). *Arabidopsis AGL24* was designated as the outgroup (Becker and Theissen, 2003). Support for each branch was assesed by bootstrap analysis with 280 replicates, each with 30 random additions of taxa per bootstrap replicate.

Genes names consist of the name of the homolog in *Arabidopsis* preceeded by the *Cornus* species name (*Ck-* (*C. kousa*), *Cf-* (*C. florida*), *Ccan-* (*C. canadensis*),*Cmas-* (*C. mas*), *Coff* (*C. officinalis*), *Calb* (*C. alba*), *Cstol* (*C. stolonifera*). Sequence similarity was calculated for the available sequences of genes according to the formula: 1- ((# bp differences + # indels)/(total # nucleotides –# gaps). Calculations are approximate as they are not based on full coding sequences for both genes.

RESULTS

Sequences of MADS-box genes identified

Thirty-eight new MADS-box genes from *Cornus* were identified, including homologs of *PI, AP3, TM6, AP1* and *SEP* from each of the four clades within the genus, with the exceptions of *AP3* in the bractless clade, which was not identified, and *PI* in the dwarf dogwoods, which was identified by collaborators (Zhang et al., 2008. Throughout the genus *Cornus*, we detected two classes of *PI*-like sequences, two classes of *AP3*-like sequences, one class of *TM6*-like sequence, one class of *AP1*-like sequence, two classes of *FUL*-like sequences, three classes of *SEP*-like sequences, one *AGAMOUS(AG)*-like sequence and several MADS-box sequences which could not be assigned to a plant particular MADS-box gene family. A phylogeny of all *Cornus* genes identified and reference genes is shown in Figure II.2.

The B-class genes

We identified two distinct copies of *PI* in the genus *Cornus*, and the same two copies were also identified independently by our collaborators (Zhang et al., 2008) who named them *CorPI-A* and *CorPI-B*. The two copies are approximately 85% similar at the coding nucelotide level within a species, as compared to the 94% similarity found at the coding nucleotide level within each copy across the genus. We identified transcripts of *CorPIA* in showy-bracted dogwoods, small-bracted dogwoods and bractless dogwoods; these sequences form a clade with 81% bootstrap support. We identified transcripts of *CorPIB* only in the small-bracted and bractless dogwoods; these sequences did not form a clade. Together the two clades are supported by a bootstrap of 54%.

We detected two different transcripts of *AP3* in the showy bracted and dwarf dogwoods. The longer transcript encodes a typical AP3-like protein while the shorter transcript contains a 76bp deletion and is truncated by 83 or 86 bp compared to the normal version of the gene. The transcripts within a species are otherwise identical, and the truncated version appears to be due to alternative splicing. A standard splice site has the sequence: 5'-(A/C)AN:**GT**xxxC**AG**:G(G/T)-3', in which the bases in normal typeface are commonly found but not essential for splicing while the bases in bold are the splice donor (5') and splice acceptor (3'). Together they constitute a splice pair and are necessary for splicing. The putative splice site in the truncated version of *CkAP3* is 5'-TT(T/C):**GA**xxxC**AG:**GT-3'. It differs from the standard splice pair at the donor site (GA vs. GT) and is the second most frequent alternative splice pair found in *Arabidopsis* (www.tigr.org).

The sequences of *CoffAP3* and *CmasAP3* are identical to those of *CkAP3, CfAP3,* and *CcanAP3*at the putative alternative splice site, but no truncated transcripts were found when amplifying *AP3* in these species. The clade of *CornusAP3* has 99% bootstrap support and is sister to *Hydrangea AP3* with 83% support. *Hydrangea AP3* differs at the putative alternative splice site5'-TT(T/C):**GAxxx**C**AA:**AT-3' from the known *Cornus* sequences by one base; the same alternative donor site, GA, is present, but in the place of the standard receptor site, AG, are two adenosine residues, AA, which are not known to form part of any alternative splice pair.

We identified one copy of *TM6* in *Cornus*,which is present in all species examined and forms a clade with 99% bootstrap support. In most cases, we amplified *TM6* more readily than *AP3*, for example in *C. alba* and *C. stolonifera* in which all of our primers for *AP3* amplified only *TM6*-like genes.

The A-class and E-class genes

We identified one copy of *AP1*, which is present in all species of *Cornus* examined and forms a clade with 100% bootstrap support. We identified *FUL* only in *C. kousa* and *C. canadensis*; we found two copies in the former (*CkFUL-A, CkFUL-B)* and one in the latter (*CcanFUL-B*). *CkFUL-B* is marked by an 87bp insertion relative to *CkFUL-A, CcanFUL-B* and all other *Cornus* MADS-box genes identified in this study. However, as we only sequenced one clone of *CkFUL-B,* this insertion could be due to pcr recombination or a splicing error. *CkFUL-A* and *CkFUL-B* share approximately 90% similarity at the coding nucleotide level, and *CkFUL-B* and *CcanFUL-B* also share approximately 90% similarity at the coding nucleotide level. *CkFUL-B* and *CcanFUL-B* are sister to eachother with 100% bootstrap support, but *CkFUL-A* is not sister to them. In the phylogeny, the position of *CkFUL-A* is only resolved to the the AP1-FUL gene family level (100% bootstrap support); however, *CkFUL-A* clearly belongs to the FUL gene family based on the conserved motif (MPPWLIRHVNH) at its C-terminus.

Using our degenerate primers, we detected sequences belonging to three classes of *SEP*-like genes: *SEP-A* (*CfSEP-A, CstolSEP-A), SEP-B (CkSEP-B, CcanSEP-B, CmasSEP-B, CoffSEP-B),* and *SEP3* (CkSEP3). The entire class of *SEP* genes is highly supported by bootstrap analysis (97%), and the distinct classes of *SEP* genes are also highly supported (98% for *SEP*-A and 100% for *SEP-B*). CkSEP3is so-named because it forms a clade with *Arabidopsis* and Tomato *SEP3* genes with 96% bootstrap support. SEP-*A* and *SEP-B* are not clearly identifiable as homologous to an *Arabidopsis SEP* gene, though the *Cornus SEP-B* genes form a moderately-supported clade with a Tomato *SEP* (*LeSEP1)*, which has been shown to be orthologous to *AtSEP1* (Litt 2003). However, in our study, *AtSEP1* does not belong to this clade.

Other MADS-box genes

Our degenerate A, B and E class primers occasionally amplified sequences belonging to other classes of MADS-box genes. We identified a single *AG*-like sequence from *C. kousa*, an *AGL6*-like (*AGAMOUS-LIKE-6-like)* sequence with homologs in *C. kousa* and *C. florida* (*CkAGL6*-like*, CfAGL6*-like), and two transcripts of a PTM5-like (*POTATO-TOMATO-MADS-5*-like) sequence from *C. canadensis*. The *AG*-like sequence forms a clade with a bootsrap of 100% with *Arabidopsis AtSHP1* and has the C-terminal motif characteristic of the AP1 MADS-box family. The *AGL6*-like genes form a clade with *AtAGL6* with 63% bootstrap support. *CkAGL6*-like has the conserved C-terminal motif characteristic of the *AGL6* clade (NMHGWVL). *CfAGL6* appears to have an insertion of one base pair, causing it to be truncated by 334 bp compared to *CkAGL6*-like; however, because only one sequence was obtained, this could be due to PCR error. Of the two transcripts of a *PTM5*-like sequence from *C. canadensis,* one has a 98bp insertion and is truncated by 95bp at the C-terminal end compared to the other. The 98bp indel does not appear to be due to a splicing event as the nucleotides in at the ends of the indel do not match those of any known splice pair. These genes do not form a clade with any other genes in the phylogeny.

DISCUSSION

In this study we have identified thirty-eight new MADS-box homologs from *Cornus*, most of which belong to recognized families of floral organ identity genes. We identified the partial or full sequences of the A, B and E class genes chosen as candidates for involvement in the evolution of petaloid bracts in dogwoods. Among the genes identified, we observed recent duplication events, putative alternative splicing, and general concordance of gene phylogeny with established species phylogeny.

Duplication and Divergence in *Cornus* MADS-box genes

At least two copies of *PI* exist in *Cornus*, *CorPI-A* and *CorPI-B*. Consistent with our findings, Zhang et al. detected genomic and RNA sequences of *CorPI-A* in all four clades of the genus. However, they detected genomic and RNA sequences of *CorPI-B* only in showy-bracted, dwarf, and small bracted dogwoods, and detected neither in bractless dogwoods (Zhang et al., 2008). In contrast, we did not detect transcripts of *CorPI-B* in showy-bracted or dwarf species but did detect transcripts in small-bracted and bractless species. Our combined results show that both copies of PI are present and expressed in all clades. As such, we know that the duplication event in this lineage is at least as old as the origin of the genus and that both copies continue to function in flowers.

A duplication was also detected in the *FUL* lineage. As orthologous copies are present in *C. canadensis* and in *C. kousa*, it occurred before the divergence of showy-bracted and dwarf dogwoods. Relatively recent duplications are commonly found in studies of MADS-box genes; in *Arabidopsis* two recent duplications are *AP1/CAL* and *SEP1/SEP2*. Surveys of closely related genera or plant families also show frequent intragenic or intrafamilial MADS-box gene duplications (e.g. Kramer et al., 2003; Hileman, 2006). In some cases expression patterns of paralogs have been determined, and in some cases they indicate subfunctionalization (e.g. Yamaguchi et al., 2006), although functional studies are required in order to demonstrate different gene functions.

Along with recent duplications, alternative splicing of MADS-box genes has been detected in many studies (e.g., Shan et al., 2007; Kramer et al., 2003). Alternative splicing generates distinct versions of genes without duplication and divergence. We found at least one gene that appears to be alternatively spliced in some species of *Cornus*. Truncated transcripts of *AP3* were detected by RT-PCR in the two showy-bracted and one dwarf dogwood studied. By comparing *Cornus AP3* sequences with *AP3* from *Hydrangea*, it appears that a single substitution of A→G created a novel occasional splice site consisting of a nonstandard donor site + standard acceptor site in *AP3* of *Cornus*. Because we have not identified *AP3* from bractless dogwoods, we do not know if this substitution occurred within *Cornus* or in an ancestor of the genus. The truncated AP3 proteins generated by alternative splicing lack C-terminal motifs which are thought to

protein-protein interactions, and as such, could possibly interact with MADS-box partners with which an *euAP3* gene normally does not. If this is the case, the presence of a petal organ identity gene with such flexibility might have been a preadaptation in dogwoods for *de novo* evolution of a petaloid structure.

Relationships within the genus *Cornus*

As expected, *CornusPI, AP3, TM6, AP1,* and *SEPB* genesfrom all four clades of the genus form monophyletic groups with 95 to 100% bootstrap support. Support is also very high for orthologs from species in the same clade, e.g. 100% support for *CkTM6* with *CfTM6* and 99% support for *CmasAP3* with *CoffAP3*. The disputed relationships within *Cornus* are those among the showy-bracted, dwarf, and small-bracted dogwoods. The phylogeny of the MADS-box genes studied here generally supports the phylogeny of *Cornus* as determined by Xiang et al. (1996, 1998, 2002) in which showy-bracted dogwoods and dwarf dogwoods form a petaloid clade which is sister to the small bracted-dogwoods, while the bractless dogwoods are basal within the genus; for example, among *AP1* genes there is 97% support for the clade of *C, kousa*, *C. florida* and *C. canadensis* which falls to 70% with the addition of *CoffAP1*. Two clades of showy-bracted and small-bracted genes as sister exist among our gene trees, for *SEPB* and *AP3*; however, in both cases there is only 66% support for the clade of the *Ck* gene, *Cf* gene, and *Cmas* gene.

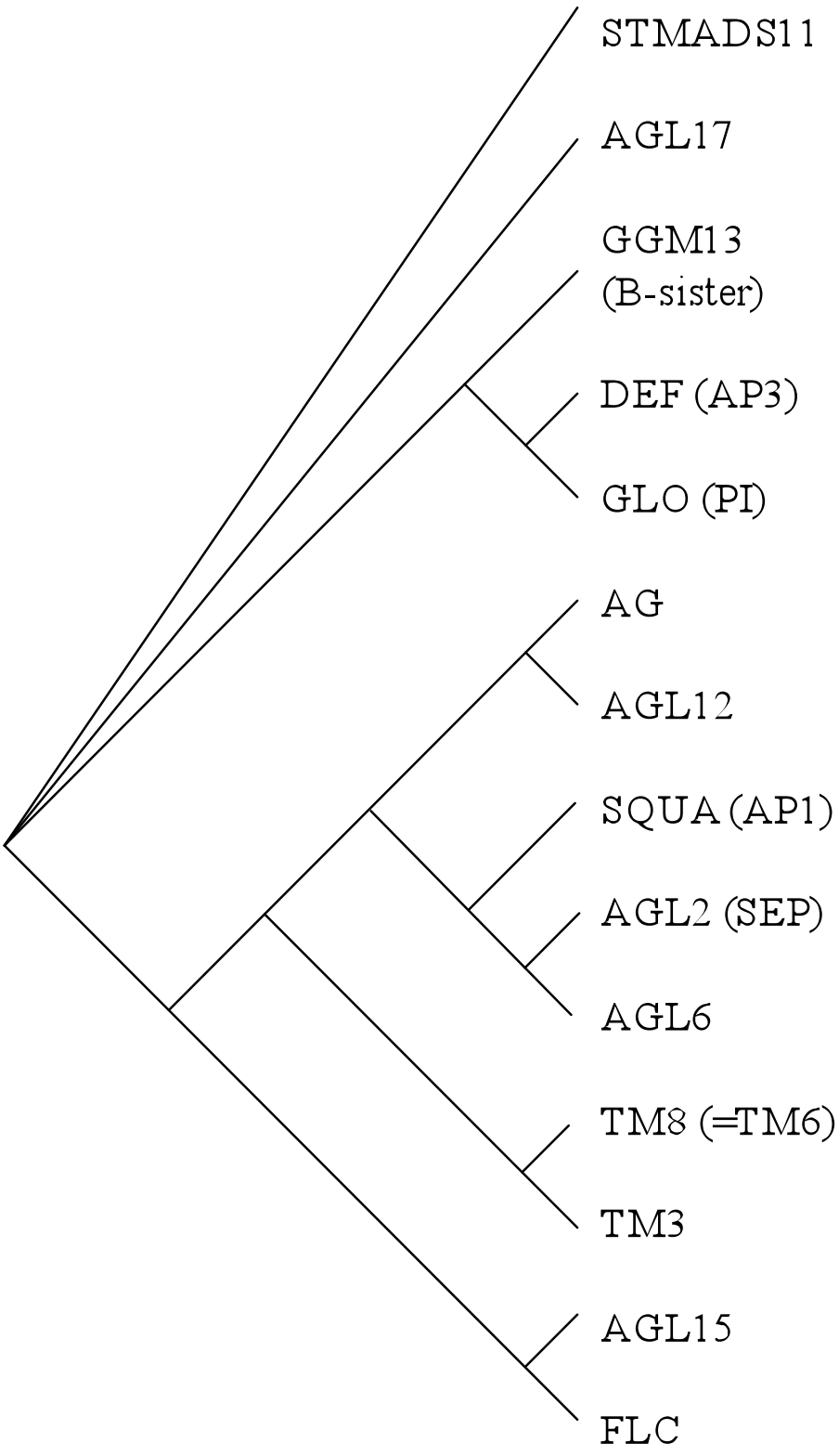


Figure II.1. Relationships among major clades of MADS-box genes families in plants, redrawn from Becker and Theissen, 2003.

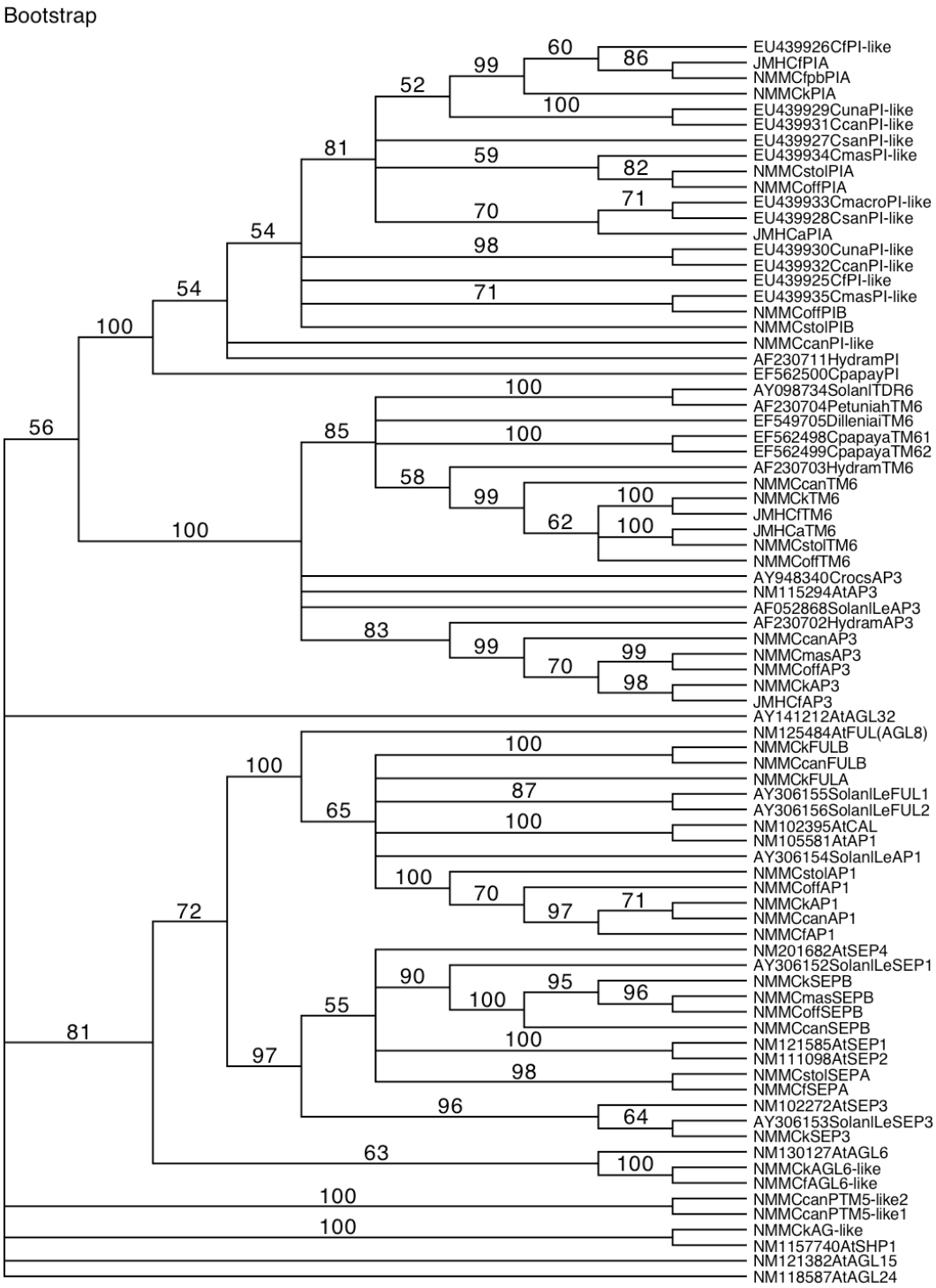


Figure II. 2. Phylogeny of newly-identified *Cornus* MADS-box genes and reference genes.

CHAPTER III
EXPRESSION OF PETAL ORGAN IDENTITY GENES IN *CORNUS*

ABSTRACT

Three out of four clades of dogwoods bear inflorescences subtended by prominent bracts. In two of these clades, the big-bracted and dwarf dogwoods, inflorescence bracts undergo a transition from small and leaf-like to large and petal-like before flowering. We hypothesize that ectopic expression of petal organ identity genes may play a key role in this transition. We examine the expression patterns of homologs of known petal-identity genes in bracts, flowers, receptacles and leaves of four species of *Cornus*, one representing each clade in the genus, by real-time quantitative PCR. Our results show that A, B and E class genes are significantly expressed in bracts of *C. florida* and are not significantly expressed in bracts of *C. canadensis* or *C. mas.*

INTRODUCTION

Petal organ identity gene expression, function, interaction, and regulation all contribute to our understanding of the evolution of petals, and changes in any of these phenomena could contribute to the origin and development of petaloid bracts in *Cornus.* According to the ABC(DE) model of floral development, three gene functions, A, B, and E are necessary for the specification of petals. The B and E elements of this model have been validated in a range of angiosperm flowers, though specifics of gene expression and function vary. The A function, however, has not been widely corroborated and is debated.

Expression patterns and functions of petal organ identity genes

In *Arabidopsis* and *Antirrhinum*, B gene transcripts initially appear independently of one another (stages 3-5 of *Arabidopsis* floral development) in somewhat different region-specific patterns, which generally hold true for orthologs. Early expression of *DEF*/*AP3* is normal in *glo*/*pi* mutants(Jack et al. 1992; Trobner et al. 1992), and vice versa (Goto and Meyerowitz, 1994), but late expression of both genes depends upon the presence of both B gene proteins (Jack et al., 1992;Schwarz-Sommer et al. 1992, Trobner et al. 1992; Goto and Meyerowitz, 1994) and is restricted to second and third whorls where the protein products occur together due to early expression. An exception to maintenance of late expression by autoregulation is expression of *AP3* at the base of first whorl sepals in wild type and in second whorl sepals of *pi* mutants(Jack et al., 1992; Weigel and Meyerowitz, 1993).

Expression patterns of *AP3* and *PI* genes in other core eudicots are similar to those described above, and both genes are required for petal formation in eudicots. Expression and function of *TM6* genes is less widely known. In *Petunia hybrida, PhTM6* is expressed in developing stamens and ovaries. Loss-of-function mutants do not have an obvious mutant phenotype (Rijpkema et al., 2006).

Expression patterns of B genes in basal angiosperms and monocots are more varied than in eudicots and *AP3* and *PI* genes are sometimes seem to function independently of each other (Kramer and Irish, 2000; Kim et al.,2005). For example, in the basal angiosperm *Saruma*, which has a classical flower in four whorls, B gene expression seems to follow the ABC model; however, in a close relative *Aristolochia*, which has an unusual fused perianth, B genes are expressed only at late stages and do not seem to have any role in specifying organ identity (Jaramillo 2004). In basal lineages of angiosperms including *Amborella, Nuphar*, and *Magnolia*, *AP3* and *PI* expression is generally detected in all floral organs, though both B genes are not always present in the same organ (Kim et al., 2005). In monocots, B genes are often expressed in the first three whorls, consistent with the lack of sepal/petal distinction and petaloid character of the first two whorls (Kanno et al., 2003; Nakamura et al., 2005). Such a pattern of gene expression can be incorporated into the ABC model by a modification called “shifting

borders" in which the domain of expression of a class of gene may grow or shrink to include more or fewer whorls. However, the shifting borders model cannot account for all diversity of perianth morphology, such as adjacent whorls of distinctly different petaloid organs seen in Ranunculaceae (Kramer et al., 2003).

The *SEPALLATA (SEP)* genes have been shown in *Arabidopsis* to be the flower-specific factor required for B and C organ identity functions and for a determinate floral meristem (Pelaz et al., 2000). In *Arabidopsis*, the *SEP1/2/3/4* are expressed very early in whorls one through four of the flower (Flanagan and Ma, 1994; Mandel et al., 1998) and are largely redundant. The triple mutant *sep1 sep2 sep3* produces indeterminate flowers with three whorls of sepals and in the fourth whorl a new flower repeating the same pattern (Pelaz et al., 2000) while the quadruple mutant *sep1 sep2 sep3 sep4* produces whorls of leaf-like organs (Ditta et al., 2004). In several basal angiosperms and in a basal eudicots, California poppy, *SEP* genes are also widely expressed throughout the flower (Kim et al., 2005; Zhan et al., 2004).

Expression of *AP1* in floral whorls one and two of *Arabidopsis* is not typical of eudicots, and in no plant other than *Arabidopsis* does loss of *AP1* function lead to loss of both sepal and petal identity (Litt, 2007). Expression patterns are wide and varied in eudicots, with principal expression occurring in floral meristems, concurrent with the primary function of *AP1* genes in specifying floral meristem identity (Litt 2007; Shan et al., 2007). Like *AP3, PI,* and *SEP* genes, *AP1* genes are expressed throughout floral organs in basal angiosperms (Kim et al., 2005; Shan et al., 2007).

The general pattern of broad expression of A, B, and E class genes in basal angiosperms and more specific expression patterns (even in the case of redundant genes) of such genes in eudicots is linked to the duplication and divergence which characterize these gene lineages (see Introduction, Chapter II). In most cases paralogous genes have a conserved structure but distinct regulation, supporting the observation that much evolution of regulatory genes may be due to changes in gene expression as opposed to changes in gene products. For example, *Petunia hybrida TM6,* which is expressed in developing stamens and ovaries, does not appear to be involved in petal development (Vandenbussche et al., 2004); however, 35S-driven *PhTM6* expression can rescue petal development in a *phdef* mutant background (Rijpkema et al., 2006). While it may be true

that a majority of the evolution of regulatory genes involves regulatory rather than structural changes, the dimerization and multimerization properties of plant MADS-box genes highlight the role that structural changes of genes play in evolution.

Dimerization of petal-organ identity genes

MADS-box transcription factors must dimerize in order to bind DNA. Generally speaking, A and E class genes form homodimers and heterodimers, while B genes form DNA-binding heterodimers strictly with each other (Davies et al., 1996; Riechmann et al., 1996). However, strict heterodimerization of B class genes in eudicots clearly evolved from ancestral homodimerization (Winter et al., 2002), and several instances of DNA-binding B-class gene homodimers have been found in monocots (Winter et al., 2002; Kanno et al., 2003; Tsai et al., 2008).

Dimerization properties are determined by various domains in plant MADS-box proteins. The I domain and a portion of the K domain of *Arabidopsis* AP3 are required for dimerization and allow dimerization only with PI (Purugganan et al.,1995; Riechmann et al., 1996). Similarly in *Antirrhinum,* the K box of DEF is required for dimerization with GLO (Davies et al., 1996). In transgenic experiments between *Arabidopsis* and *Chloranthus spicatus,* the C-terminal region of *Arabidopsis* AP3 was substituted by the C-terminal region of *Chloranthus spicatus* AP3, and it was not able to homodimerize. However, when the MIK region was substituted, the protein was able to homodimerize (Su et al., 2008). In contrast, motifs necessary for homodimerization have been identified from the C-terminal region of a monocot *AP3* homolog, *LMADS1*, and when the C-terminal region of *Arabidopsis AP3* was replaced by that of *LMADS1*, it too was able to homodimerize (Tzeng et al., 2004).

In *Arabidopsis* PI, the I domain and a portion of the K domain required for dimerization may form either heterodimers with AP3 or homodimers, and it is the MADS-box which prevents formation of PI homodimers (Reichmann et al., 1996; Su et al., 2008). Since MADS-box proteins bind to DNA as dimers, the I domain, as well as a portion of the K domain, is also required for AP3-PI DNA binding, as confirmed by tests of DNA-binding activity with truncated proteins. Unlike the class B genes, the A and C

genes require only the MADS-box and I domains for dimerization and DNA binding (Davies et al.,1996; Riechmann et al., 1996).

Ternary and quaternary interactions among A, B, E genes

Protein-protein interactions among DEF-GLO and SQUA constructs in yeast occur through the C-termini, which are not required for DEF-GLO heterodimerization or SQUA-SQUA homodimerization (Egea-Cortines et al., 1999). In *Arabidopsis* proteins AP1 and SEP, the C-terminal half of the K domain plus the C domain (K2+C) is sufficient for interaction of these proteins with PI-AP3 and provides transcriptional activation, which PI and AP3 themselves lack (Honma and Goto, 2001).

Using a yeast two-hybrid system, SEP3 has been shown to interact with the PI-AP3 complex and independently with AP1 and with AG (Fan et al., 1997; Honma and Goto, 2001). AG does not interact directly with PI-AP3 but with the addition of SEP, all four together do interact, consistent with the requirement for all four genes to specify stamens (Honma and Goto, 2001).

Evidence suggests that the MADS-box proteins function as tetramers in vivo. SQUA, DEF, and GLO also form ternary complexes which show increased binding to CArG motifs in yeast compared with DEF-GLO or SQUA homodimers (Egea-Cortines et al., 1999), and AP1 and SEP3 interact with AP3-PI and provide the B gene complex with transcriptional activator domains. The ‘quartet model’ has been proposed to explain determination of the four floral whorls by combinations of homeotic genes: AP1-AP1 interacts with SEP-SEP to specify sepals, AP3-PI interacts with AP1-SEP in whorl 2 to specify petals and AG-SEP in whorl 3 to specify stamens, and two AG-SEP dimers interact to specify carpels (Theissen and Saedler, 2001). As A function is debated, the same quartet model can also be framed in a two-gene-function model of floral development in which AP1-AP1 interacts with SEP-SEP to specify floral meristem and default organs (sepals), and the B and C functions are added to this background to produce additional types of organs (Litt, 2007).

Another important element of organ specification and development is cell-cell communication and the issue of cell-autonomous vs. cell-non-autonomous control of

development. Autonomous vs. non-autonomous control of development by the B genes has been studied by localizing B transgene expression to epidermal cells and observing whether cells in all layers develop as expected for petals and stamens. In *Antirrhinum*, epidermal DEF and GLO can influence petal development in lower cell layers by activating endogenous genes in those layers. Strikingly, in *Arabidopsis*, epidermal AP3, PI and even DEF are sufficient to individually control petal and stamen development in an *ap3* background and at ectopic positions without influencing transcription of endogenous B genes or RNA trafficking (Efremova et al., 2001). In *Antirrhinum* DEF and GLO proteins are trafficked from subdermal to epidermal layers but not vice versa (Perbal et al., 1996) and in *Arabidopsis* B gene protein does not traffic between cells (Jenik and Irish, 2001). Thus the mechanism of cell-cell communication is unknown, but it seems that in *Arabidopsis* target genes can be activated indirectly, by signaling, whereas in *Antirrhinum* the physical presence of intracellular class B proteins is required (Efremova et al., 2001).

Regulation of petal organ identity genes in model organisms

As the mechanisms of gene regulation are less accessible than gene expression, or even function, little is known about regulation outside of model organisms. B genes in *Arabidopsis* are the best-studied of petal organ identity genes and are reviewed here.

Arabidopsis AP3 and *PI* have similar function, their *cis*-acting regulatory elements have diverged greatly. The regulatory sequence in *PI* is unique because unlike promoters of *AP3, DEF,* and *GLO*, its promoter contains no CArG boxes to which MADS-box proteins typically bind. In addition, *PI* does not autoregulate directly (Chen et al., 2000). Thus, although the *PI* amino acid sequence is more conserved than that of *AP3*, it appears to have experienced more divergence in regulation.

AP3 also seems to have undergone changes in regulation. Promoter dissection of *AP3* has identified an early-and-late petal specific sequence which is regulated independently of PI (Irish and Yamamoto, 1995). It is possible that the petal-specific *AP3* element evolved at the time of the duplication and divergence of the *AP3* lineage at the base of the core eudicots. On the other hand, the petal specific element may have been

ancestral and the autoregulatory elements may have evolved at this juncture. The factor which regulates *PI*-independent expression of *AP3* is unlikely to include *AP3*, because AP3 must be bound to PI to be transported from the cytoplasm into the nucleus.

B gene transcripts must be translated into protein, and the proteins must move to the nucleus in order to fulfill their function as transcription factors. A unique nuclear co-localization system has been discovered by examining AP3 and PI reporter gene fusion constructs in onion cells and in transgenic *Arabidopsis*. Neither protein alone contains functional nuclear localization sequences, as is evidenced by their localization in the cytoplasm; however, co-expression of the fusion proteins shows localization to the nucleus. Truncated fusion proteins show that the amino terminal ends of both proteins, the first 69 amino acids of AP3 and the first 105 amino acids of PI, are responsible for co-localization. The authors hypothesize that these regions contain an intermolecular bipartite signal, or that because these domains are responsible for heterodimerization, a conformational change upon dimerization unmasks a signal present in one or both proteins (McGonigle et al., 1996).

Three genes are required to direct normal initial expression of *AP3* and *PI*, although their interactions and mechanisms of regulation are still not completely understood*;* they are the floral meristem identity genes *LEAFY (LFY)*, *AP1* and *UNIDENTIFIED FLORAL ORGANS (UFO)*. *LFY*, homolog of *FLORICAULA (FLO)* of *Antirrhinum* (Coen et al., 1990)*,* is both a flowering-time and floral meristem identity gene (Blazquez, 1997). *LFY is* transiently expressed in the very early stages of flower development (Coen et al., 1990; Weigel et al. ,1992), and the LFY protein product persists after mRNA expression has ceased. *LFY* itself is induced by genes in the daylength-dependent and daylength-independent flowering time pathways (Ruiz Garcia et al. 1997; Kardailsky et al. 1999; Blazquez and Weigel, 2000). In *lfy* mutants, *AP3* and *PI* expression is reduced (Weigel and Meyerowitz, 1993), and ectopic expression of both *AP3* and *PI* rescues the floral organ identity defects of *lfy* (Krizek and Meyerowitz, 1996). Thus *LFY* acts upstream of the B genes to stimulate B gene expression. Since LFY acting in a meristem is sufficient to cause the development of a flower, the three *SEP* genes act downstream of *LFY*, though their direct regulators are not known.

LFY acts in part through *AP1*, both directly and indirectly upstream of *AP1* (Parcy et al., 1998; Wagner et al., 1999). *AP1* also is induced by the gene *FLOWERING TIME (FT),* which acts in parallel to *LFY* in the daylength-dependent flowering time pathway. In *lfy ap1* double mutants, transformation of flowers into shoots is much more complete, and *AP3* and *PI* expression are dramatically reduced as compared to *lfy* single mutants, indicating a role for *AP1* in activation of the B class genes (Bowman et al., 1993, Weigel and Meyerowitz, 1993). AP1 does bind to the *AP3* promoter *in vitro,* supporting the hypothesis that *AP3* is a direct target of AP1 (Hill et al., 1998). In addition, *AP1* has also been shown to act upstream of *UFO* in early stages, making *AP3* an indirect as well as a possibly direct target of *AP1* (Ng and Yanofsky, 2001).

UFO, homolog of *FIMBRIATA (FIM)* of *Antirrhinum* (Simon et al., 1994), is expressed in a cup-shaped domain in the center of both vegetative and floral meristems. It acts downstream of the gene *SHOOTMERISTEMLESS (STM)* which maintains the shoot apical meristem, although the function of *UFO* in the vegetative meristem is unknown. In the floral meristem, it participates in influencing floral meristem and organ identity through activation of the B genes in conjunction with *LFY* (Levin and Meyerowitz, 1995; Lee et al., 1997). *Fim* null mutants show no or little expression of *DEF* and *GLO,* and as is the case with *lfy* mutants, ectopic expression of both *PI* and *AP3* also rescues the floral organ identity defects of *ufo*, demonstrating that *UFO*/FIM acts upstream of the B genes (Krizek and Meyerowitz, 1996; Ingram et al., 1997). However, overexpression of *AP3* or *PI* independently shows that *AP3* can rescue a *ufo* mutant while *PI* can not, indicating that *UFO* activates transcription of *AP3* but is not required for expression of *PI* (Samach et al., 1999) *UFO/FIM* encodes a protein which contains the F-box motif, suggesting that UFO/FIM associates in a complex which targets specific proteins for ubiquitin-mediated degradation. Furthermore, UFO and FIM interact physically with other proteins (ASK1/FAP) that also appear to be involved in ubiquitin-mediated degradation (Ingram et al., 1997, Samach et al., 1999).

In this study we examine the expression patterns of only the petal-identity genes belonging to B and E classes, and *AP1*. Ultimately we seek to identify the genetic evolutionary origin of petaloid bracts in *Cornus*, which may lie upstream of the petal-organ identity genes, which we do find to be expressed in showy bracts.

METHODS AND MATERIALS

Plant Materials

Leaves and inflorescences of *C. florida*, *C. canadensis*, *C. mas*, and *C. alba* were collected from locations listed in Table I.1. *C. florida* inflorescences were sampled ten times during the year; of those, six samples yielded total RNA of sufficient quantity and quality for qPCR (October, February, March, early-April, mid-April, and late-April), as illustrated in Figure III.1. Bracts harvested in May had very low yields of RNA, as is typical for older tissues. *C. mas* and *C. canadensis* were both sampled at three stages of bract development: February, March and April in *C. mas* and green, cream-colored and white inflorescence bracts in *C. Canadensis*. *C. alba* was collected in a late-bud stage. Inflorescences were separated into bracts, receptacles and flowers/flower buds before being immersed in liquid nitrogen. These tissues were stored at –80ºC for periods up to four years before isolation of RNA. Leaves and inflorescences of *C. canadensis* were also collected in RNAlater (Ambion of Applied Biosystems); these tissues were stored at –20ºC and used within one year of collection. Total RNA was extracted separately from leaves, bracts, flowers and receptacles using the RNAqueous Midi kit (Ambion). RNA quality was assessed on the BioAnalyzer (Agilent), as described in Chapter II Methods and Materials.

Quantitative real-time PCR

Quantitative PCR was performed using SYBR green dye (SYBR Green Mastermix, Rovalab, Teltow Germany) on the Chromo4 real-time PCR machine (BioRad Laboratories, Hercules, CA, USA). After each cycle this machine illuminates the PCR reaction with light from an LED, filtered through an interference filter, with the fluorescence measured by a photodiode, through a second interference filter. Runs were done for 40 cycles (3 minutes at 95ºC and 40 cycles of 20 seconds at 95ºC and 30 seconds at melt-anneal temperature; melt-anneal temperatures were chosen to optimize all PCR reactions for a given species and were set at 58ºC for *C. florida*, 61ºC for *C.alba*,

and 62ºC for *C. canadensis* and *C. mas*). Gene specific primers were designed to amplify 100 to 200 bp of nucleotide sequence that was determined by the DINAMelt program (Markham, 2005) to be free from significant secondary structure. For amplification of *PI*, primers were designed to amplify both A and B copies in *C. florida, C. canadensis,* and *C. mas*. In *C. alba,* however, sequence divergence permitted the design of two pairs of primers to distinguish between copies A and B. Primers were tested in the absence of template to check for false product, and only primer pairs that did not yield any false product within the first 30 cycles of PCR were used. Primers were tested in the presence of template to ensure that they amplified product of a single size, as determined by melt curve analysis. Products were spot-checked by agarose gel and sequenced in the case of any doubt regarding authenticity.

Analysis of qPCR data was performed according to the Pfaffl method in which $RA = (E_{test}\text{^}\Delta Ct)/(E_{ref}\text{^}\Delta Ct)$. RA is Relative Abundance of the gene under amplification. E is the PCR efficiency of the PCR reaction which was calculated by linear regression of the log of fluorescence against cycle number in LinReg PCR (Ramakers et al., 2003). Critical threshold (Ct) values were calculated by the Chromo4 realtime PCR machine under the model of a global minimum value and subtracted baseline. ΔCt is the difference in (Ct) points between amplification of the gene in a calibrator tissue and in the test sample. Flowers were used as calibrators in each species, and *Cornus Ubiquitin10* was used as the reference gene to normalize comparisons among tissues.

Five replicates were performed for each gene and tissue combination. Reactions for which PCR efficiency could not be determined with an R^2 of 0.998 or more were thrown out. Reactions with a PCR efficiency less than 40% were also thrown out. In the case of five usable data points, the middle three in terms of PCR efficiency were used. In case of only four usable data points, the three most typical points were used. In the case of three usable data points all three were used, and in the rare (6/152 samples) case of only two usable data points, two were used. The mean PCR efficiency and Ct values for a set of replicates were used in the Pfaffl calculation. To estimate error, high and low values of PCR efficiency were substituted into the Pfaffl equation and a range of Relative Abundance was determined.

RESULTS

Figures III.1-4 show the results of quantitative PCR. General patterns of gene expression are as expected. A, B and E gene expression is low in leaves and high in flowers (not shown). Flowers are a suitable calibrator for expression of B and E genes because those genes are highly expressed in flowers. *AP1* is not as highly expressed in flowers as the other genes tested, which results in higher relative abundance of *AP1* than of other genes in *C. florida, C. mas,* and *C. alba* (e.g. Figure III.3). In *C. canadensis*, expression of *AP1* in flowers was below the threshold level, indicating a problem with primer design. Thus, though we have identified the gene, we were not able to assess expression of *AP1* in *C. Canadensis* (Figure III.2).

Most notably, the genes tested appear to be expressed in the showy bracts of *C. florida* (Figure III.1) and appear not to be expressed in showy bracts of dwarf *C. canadensis* (Figure III.2) in which the gene profiles of bracts are indistinguishable from that of leaf.

In *C. florida,* gene expression in bracts is lowest in stages 2 and 3, which correspond to winter months, February and March, and little obvious morphological change. Expression is moderate in stages 1 (October) and 4 (early April) and increases in stages 5 and 6, mid and late-April, respectively. During the later time bracts are expanding and becoming petal-like, although they still exhibit a green color.

API expression in bracts approximates *AP1* expression in flowers at three stages of development, 1, 5 and 6. Of B and E genes, only *AP3* expression approaches floral levels, in stage 6. Interestingly, expression of *AP3* and *AP3t* is approximately equal in bracts, except for stage 3, as in flowers (not shown).

In contrast, *PI* levels are about half that of *AP3* after stage 1 and only reach a significant level in stage 1 and stage 6 bracts. Those levels may be too low and/or fleeting to detect by normal PCR, as we failed to amplify *PI* from bracts despite many attempts previous to performing qualitative PCR. We have, however, amplified *SEP* and *TM6* from bracts of *C. florida,* and *SEP, TM6* and *PI* appear to be expressed at almost equal levels throughout bract stages. In receptacle tissue, however, all B gene expression is minimal while *SEP* and *AP1* expression is high.

The same pattern of high *SEP* and *AP1* expression is also seen in the peduncles of *C. alba.* In contrast, in *C. mas, SEP* expression in receptacle is low, and *C. canadensis* does not have significant receptacle tissue to harvest for isolation of RNA.

DISCUSSION

Our results show significant expression of petal-organ identity genes in the petaloid bracts of C. florida. This is the first demonstration of petaloidy outside the flower coincident with expression of petal identity genes. This result is not surprising if one takes the view that petal-organ identity genes specify petaloidy, not petals (Ronse de Craene, 2007). However, several elements of the observed gene expression pattern are surprising.

The most surprising to us is the fact that *PI* is expressed in bracts according to the qPCR results in the light of the previous failed attempts to isolate it by RT- PCR with both specific as well as degenerate primers. During our efforts to initially clone the A, B, and E gene, we isolated *AP1*, *AP3*, *TM6* and *SEP* from bracts, but not *PI*. The *PI* products generated in this study have been sequenced to verify their identity; the same sequence is amplified from bract, leaf and receptacle tissues.

Based on our results, a genetic program known from model organisms is clearly operating in the bracts of *C. florida*, but it is not a replicate of the genetic program of eudicot petals. Expression levels of *AP1* and *AP3* are much higher than those of *SEP* and *TM6*, and the truncated *AP3* (missing the C domain) is also highly expressed. A number of possibilities exist regarding dimerization and multimerization of the MADS-box genes studied. Such high and similar levels of expression of *AP1* and *AP3* indicate that they may be interacting, though typically a *euAP1* and a *euAP3* protein would not be able to dimerize. *AP1* should be able to homodimerize; it is possible that truncated *AP3t* may homodimerize, or may heterodimerize with full-length AP3, but studies in *Arabidopsis* show that the MIK region of the AP3, which is intact in *Cornus AP3t*, controls dimerization. *SEP* and *TM6* are expressed at similar levels which are relatively low levels compared to their expression in flowers and to *AP1*/*AP3*. In the case of *SEP,* it is

likely that several redundant copies of the gene are expressed, so the relatively low level of expression detected by real-time PCR may reflect the true level of expression of only one of several copies that normally constitute full expression. However, we detected only one copy of *TM6* in the genus during our efforts to clone the gene, so we expect that our real-time PCR results accurately reflect expression of *TM6.*

While specific conclusions can not be made about the genetic program for petaloidy in bracts, a number of useful components are clearly present. Considering the genes expressed in bracts of *C. florida*, we expect to find some, if not all, of the same genes expressed in the petaloid bracts of *C. canadensis*. However, the pattern of gene expression in bracts of *C. canadensis* is not distinguishable from that of leaves. It is possible that normalization by *UBIQUITIN10* does not give reliable results in *C. canadensis*, if the gene is expressed in significantly different amounts in flowers than in bracts and leaves. If the genetic formulas for petaloidy in the showy-bracted and dwarf dogwoods are truly different, petaloidy likely evolved independently in the two groups, which would be consistent with the phylogeny of *Cornus* proposed by Murrell.

In the small-bracted cornelian cherry, *C. mas*, we see a gene expression pattern that resembles that of leaves, except for the elevated level of *AP1* expression, which is higher than that of flowers. The same is seen in *C. florida*, raising the possibility that expression of *AP1* in bracts is a preadaptation for petaloidy.

AP1 is also highly expressed in receptacle tissue of *C. mas*. Expression in receptacles of both species reflects the role of *AP1* as a floral meristem identity gene, as floral meristematic tissue is likely to be present in the tissue that bears floral buds. Expression of *AP1* is also found in the floral heads of the Compositae, and considering its role as a meristem identity gene, expression of *AP1* may be related to the condensation of the shoot system into a head in the showy-bracted and small-bracted dogwoods. However, *AP1* is also expressed in the peduncles of the bractless dogwoods, so *AP1* expression by itself is not enough to generate a head-like inflorescence in dogwoods.

In *C. florida*, *SEP* is also expressed at high levels in receptacle tissue. Again, this may reflect the role of *SEP* as a floral-meristem identity gene. As high levels of *SEP* expression are not seen in *C. mas* receptacle, it may also be related to the expression of *SEP* in bracts (which are fused to receptacle tissue), or vice versa. However, the bractless

dogwood *C. alba* also displays high levels of expression of both AP1 and SEP in peduncle tissues, similar to those in flowers, so *AP1* and *SEP* expression in flower-bearing tissues may be the ancestral state in *Cornus*, with *SEP* expression having been lost in the small-bracted dogwoods or simply not detected in this study.

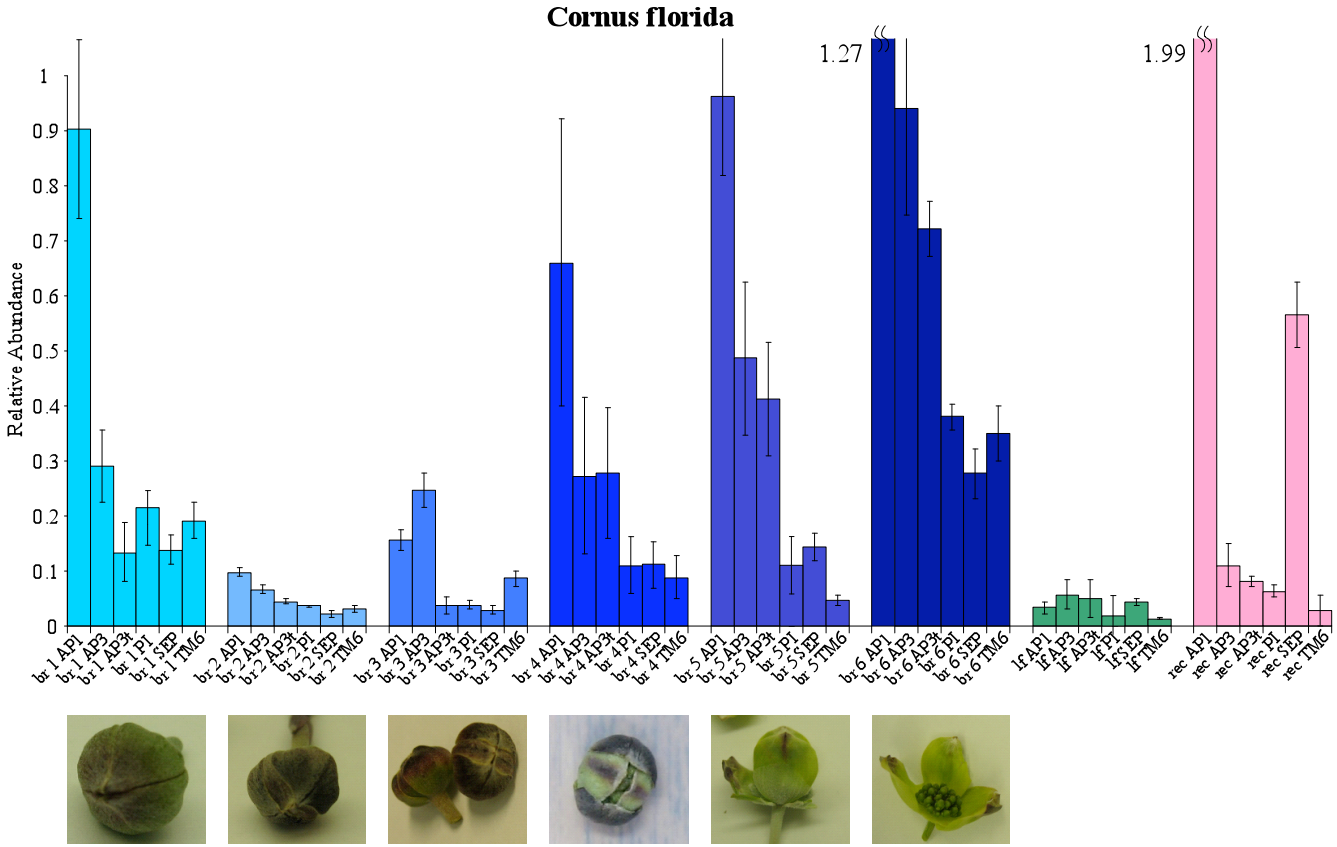


Figure III.1. Relative abundance of floral organ identity genes in the showy -bracted C. florida. Abundance of RNA in bracts, leaves and receptacle is shown relative to abundance in flowers and normalized against expression of Ubiquitin. Stages of bract development (bracts 1-6) are illustrated .

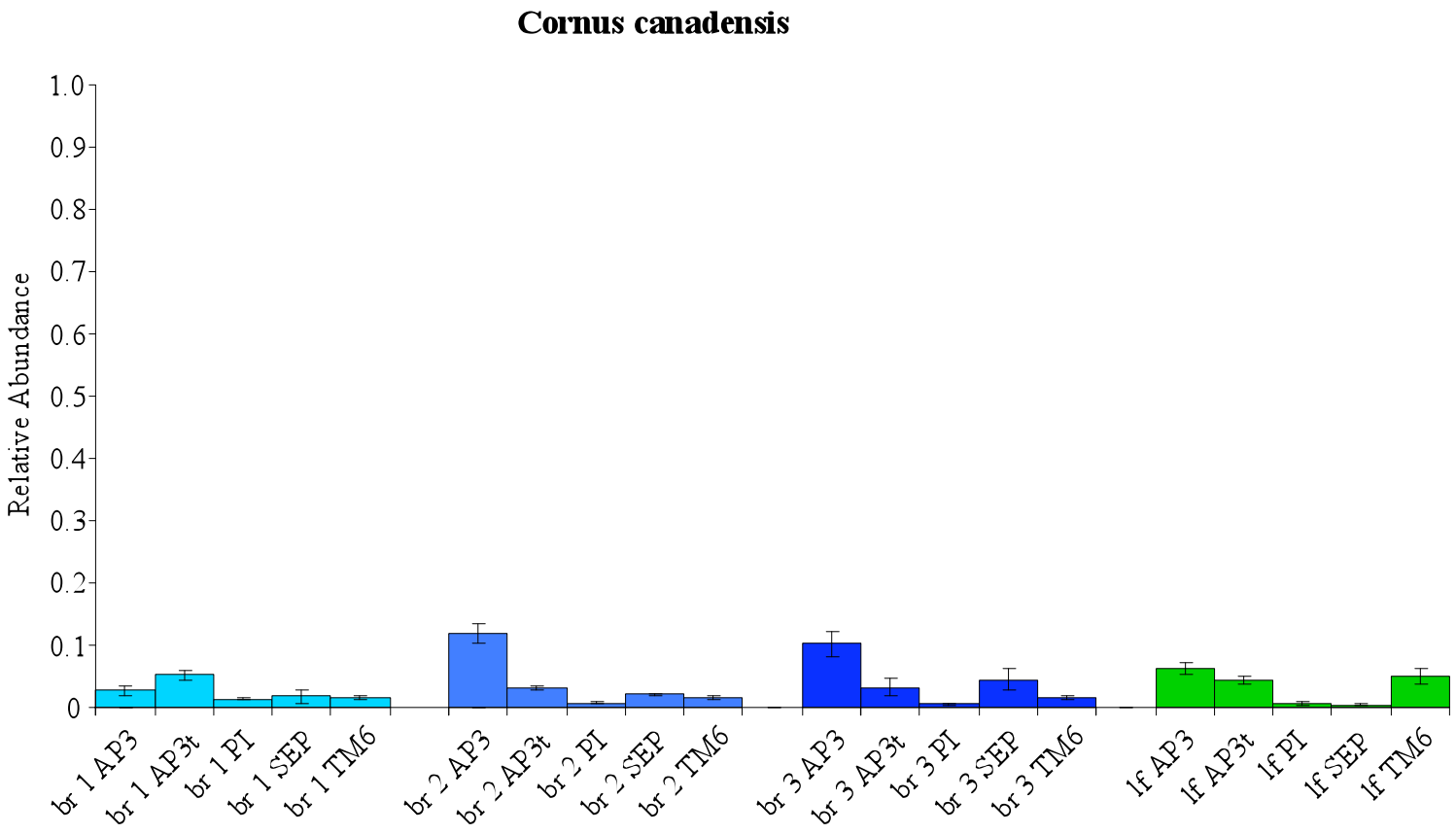


Figure III.2. Relative abundance of floral organ identity genes in the showy bracted dwarf C. canadensis. Abundance of RNA in bracts and leaves is shown relative to abundance in flowers and normalized against expression of Ubiquitin. Stages of bract development (bract 1, bract 2, bract 3) correspond to bracts illustrated in Figures I .5.G, I.5.H-I, and I.5.J, respectively.

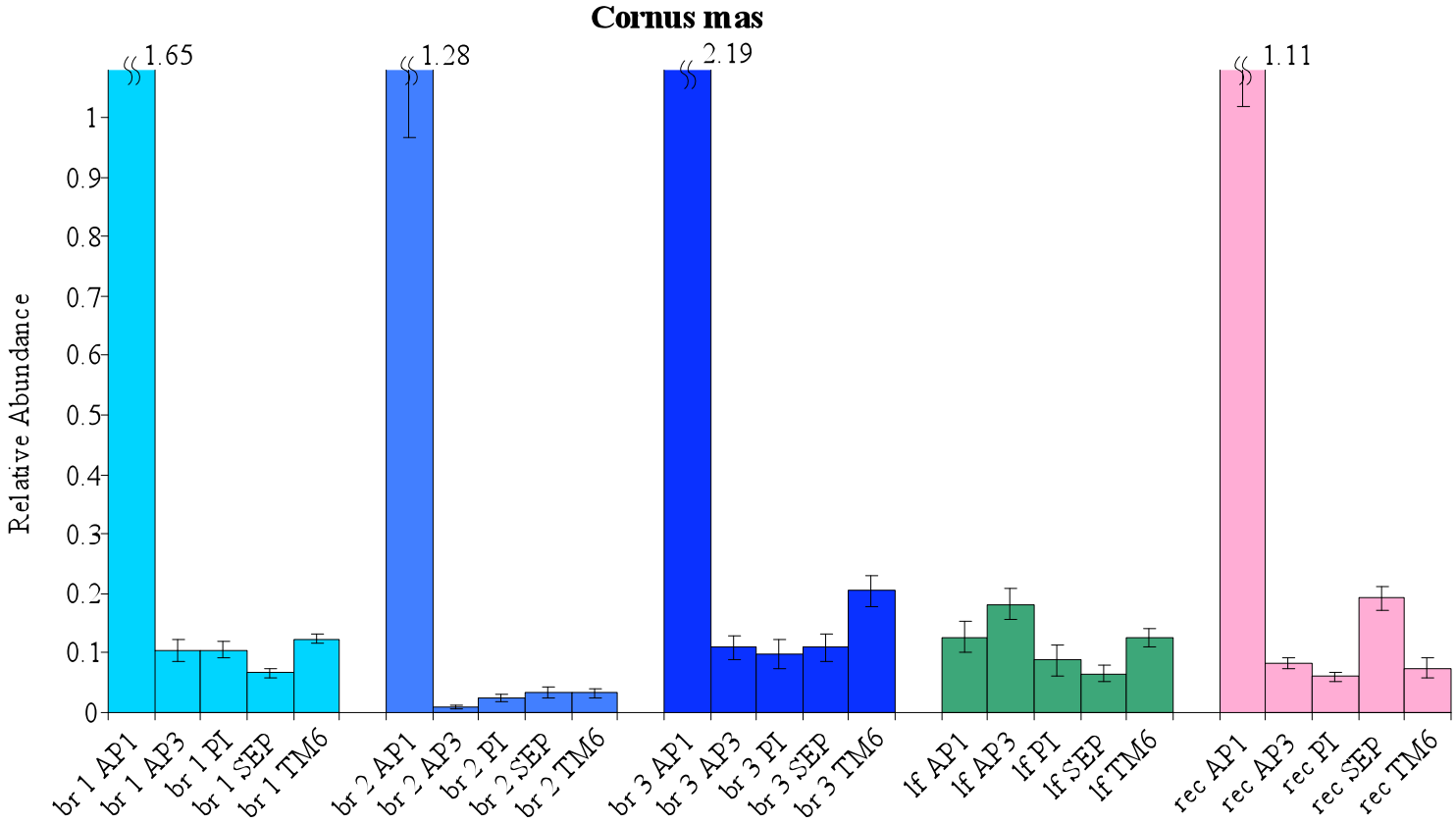


Figure III.3. Relative abundance of floral organ identity genes in the small , non-showy-bracted C. mas. Abundance of RNA in bracts, leaves and receptacle is shown relative to abundance in flowers and normalized against expression of Ubiquitin. Stages of bract development (bract 1, bract 2, bract 3) correspond to bracts illustrated in Figures I .6.A, I.6.B-C, and I.6.D, respectively.

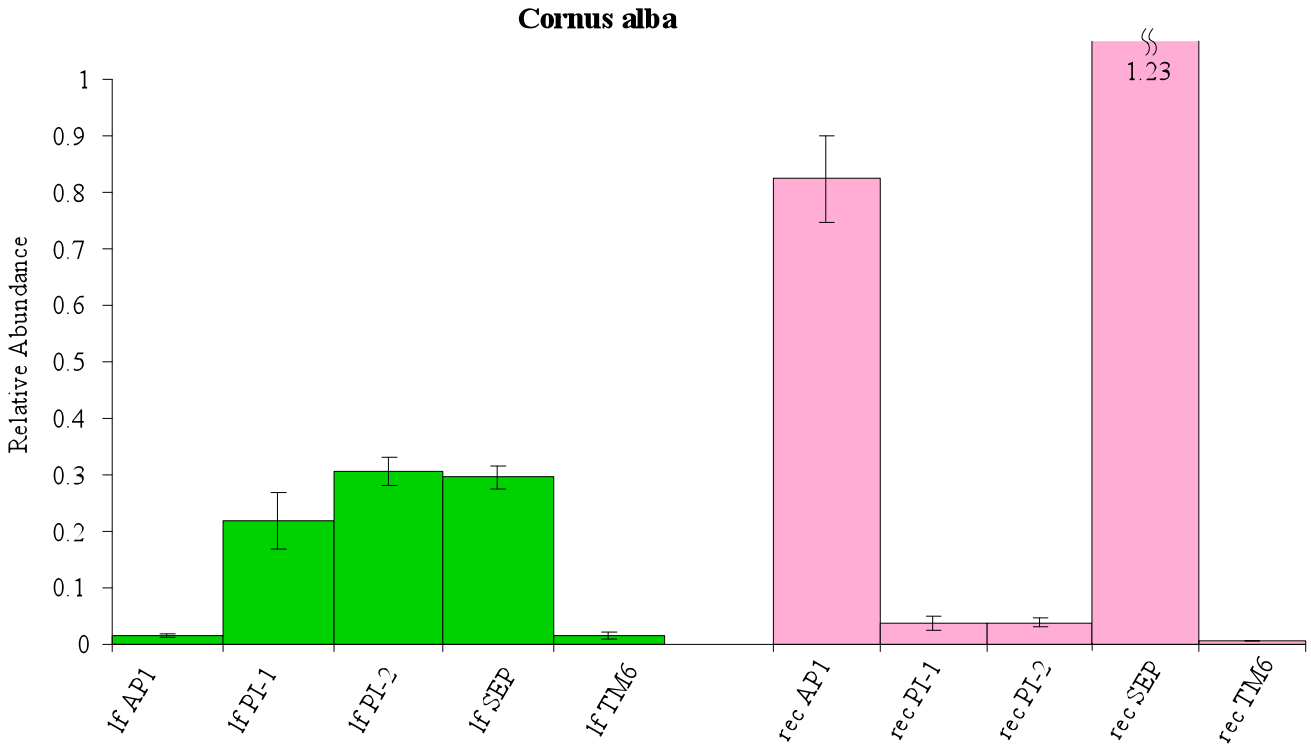


Figure III.4. Relative abundance of floral organ identity genes in the bractless C. alba. Abundance of RNA in leaves and receptacle is shown relative to abundance in flowers and normalized against expression of Ubiquitin. Tissue designated as 'receptacle' in C. alba is the peduncle, illustrated in Figure I.8.I for another species of bractless dogwood.

FUTURE WORK

In addition to qualitative real-time PCR, we have performed 454 sequencing of *C. florida* bract mRNAs as an independent measurement of expression of petal organ identity genes. The 454 sequencing has been done, and we are awaiting the assembled sequences. When the data analysis is complete, we expect to have a statistically valid sampling (over 400,000 sequences which are100-200 base pairs in length) of genes expressed in stages 1-6 of petaloid bracts. When sequenced in this way, mRNA of whole flowers typically yields more than a hundred of both AP3 and PI sequences. 454 sequencing is the best way to test for expression of *PI* in bracts. If PI is not detected 454 will provide a statistically valid upper bound to its expression level not influenced by the vagaries of PCR and primer design. If it is detected, the 454 results will provide an estimate of relative levels of expression compared to known genes (this study) as well as to not-yet-identified genes which may contribute to petaloidy of bracts.

Our results of real-time PCR for *C. canadensis* indicate that very different genetic programs specify petaloidy in the showy-bracted and small-bracted dogwoods. However, the lack of gene expression observed in *C. canadensis* may be due to flaws in our experimental design or execution. We are also performing 454 sequencing on bract RNAs of *C. canadensis*, as for *C. florida*, in order to independently examine the expression of petal organ identity genes in dwarf dogwoods and to identify genes which may specify petaloidy of bracts if the known petal organ identity genes do not do so. This will provide an independent test of our current results for *C. canadensis*, including tests of the expression level of UBIQUITIN10 and other UBQ gene paralogs.

454 sequencing is our method of choice to resolve the questions about the genetic program of petaloidy in *Cornus* that were raised in this study. Once a comprehensive picture of gene expression in bracts of *C. florida* and *C. canadensis* is available from 454,

the next step will be to elucidate the regulation of petal-organ identity genes in bracts in the hopes of pinpointing the genetic changes which were responsible for generating novel inflorescence morphology in two groups of *Cornus*.

APPENDIX A: Degenerate primers used to isolate MADS-box genes

PRIMER NAME	SEQUENCE 5' --> 3'	Gene
CAP3-KL	GCAAGARAAYYTGARRAAACTGAAAGA	*AP3*
CPI-KL	GAARGAGAAYGACAGYATGCAGATTGAGC	*PI*
CSEP-DL1	TBAAGAGRATAGAGAACAARATCA	*SEP*
CSEP-DL2	TYKCYCTYATYRTYTTCTCHAAYMG	*SEP*
CAP1-DL1	TKAAGMGSATAGARAAYAAGATCAA	*AP1*
CAP1-DL2	AARCTMTTTGAGTATKCCACWGATTC	*AP1*
Kram1 (MADS)	GGGGTACCAAYMGNCARGTNACNTAYTCNAAGMGNMG	*B*
Kram2 (PI)	TGNARRTTNGGNTGNAWKGGNTG	*PI*
Kram3 (AP3)	CNAGNCGNAGRTCRT	*AP3*
DayGLO-MZ	GSKMGIGGIAAGATCKAGAT	*PI*
DayGLO-MZ2	AACMGGCARGTGACGTAYTC	*PI*
DayGLO-Kim	GGGGTACCAAYMGICARGTIACITAYTCIAAGMGIMG	*PI*
DayGLO-MuGR	ATG GGIM GIGG IAAR ATH GAR	*PI*
DayGLO-Cryin	AAR MGIA THG ARA AYW SI	*PI*
DayGLO-TYSK	AAY MGIC ARG TIAC ITAY WSIA AR	*PI*
DayGRO-PIQ	ARR TTIG GYT GDA TIGG YTG	*PI*
DayGRO-FAF	GGY TGIA CICK RAA IGCR AAIG G	*PI*

LITERATURE CITED

Altschul, S. (1990) BLAST Reference Manual.

Alvarez-Buylla, E. R., Pelaz, S., Liljegren, S. J., Gold, S. E., Burgeff, C., Ditta, G. S., de Pouplana, L. R., Martinez-Castilla, L. and Yanofsky, M. F. (2000) An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proceedings of the National Academy of Sciences of the United States of America* 97, 5328-5333.

Angenent, G. C., J. Franken, Franken, J. Busscher, M., Vandijken, A., Vanwent, J. L., Dons, H. J. M. and Vantunen, A. J. (1995) A novel class of MADS box genes is involved in ovule development in Petunia. *Plant Cell* 7, 1569-1582.

Becker, A. and Theissen, G. (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution* 29, 464-489.

Becker, A., Kaufmann, K., Freialdenhoven, A., Vincent, C., Li, M.-A., Saedler, H. And Theissen, G. (2002) A novel MADS-box gene subfamily with a sister-group relationship to class B floral homeotic genes. *Molecular & General Genetics* 266, 942-950.

Blazquez, M. A. (1997) Illuminating flowers: CONSTANS induces LEAFY expression. *Bioessays* 19, 277-279.

Blazquez, M. A. and Weigel, D. (2000) Integration of floral inductive signals in Arabidopsis. *Nature* 404, 889-892.

Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M. and Smyth, D. R. (1993) Control of Flower Development in Arabidopsis thaliana by Apetala1 and Interacting Genes. *Development* 119, 721-743.

Bowman, J. L., Smyth, D. R. and Meyerowitz, E. M. (1991) Genetic interactions among floral homeotic genes of Arabidopsis. *Development* 112, 1-20.

Chang, S., J Puryear, J Chirney (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11, 113-116.

Chen, X. M., Riechmann, J. L., Jia, D. X. and Meyerowitz, E. (2000) Minimal regions in the Arabidopsis PISTILLATA promoter responsive to the APETALA3/PISTILLATA feed back control do not contain a CArG box. *Sexual Plant Reproductio*n 13, 85-94.

Coen, E. S. and Meyerowitz, E. M. (1991) The War of the Whorls - Genetic Interactions Controlling Flower Development. *Nature* 353, 31-37.

Coen, E. S., Romero, J. M., Doyle, S., Elliott, R., Murphy, G. and Carpenter, R. (1990) Floricaula: a Homeotic Gene Required for Flower Development in Antirrhinum-Majus. *Cell* 63, 1311-1322.

Colombo, L., Franken, J. Koetje, E., Vanwent, J., Dons, H. J. M., Angenent, G. C. And Vantunen, A. J. (1995) The Petunia MADS box gene Fbp11 determines ovule identity. *Plant Cell* 7, 1859-1868.

Davies, B., Egea-Cortines, M., Silva, E. D., Saedler, H. and Sommer, H. (1996) Multiple interactions amongst floral homeotic MADS box proteins. *Embo Journal* 15, 4330-4343.

Ditta, G., Pinyopich, A., Robles, P., Soraya Pelaz and Yanofsky, M. F. (2004) The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. *Current Biology* 14, 1935-1940.

Edgar (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32, 1792-7.

Efremova, N., Perbal, M. C., Yephremov, A., Hofmann, W. A., Saedler, H. and Schwarz-Sommer, Z. (2001) Epidermal control of floral organ identity by class B homeotic genes in Antirrhinum and Arabidopsis. *Development* 128, 2661-2671.

Egea-Cortines, Saedler, H. and Sommer, H. (1999) Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus. *Embo Journal* 18, 5370-5379.

Eyde, R. (1988) Comprehending Cornus, Puzzles and Progress in the Systematics of Dogwoods. *The Botanical Review* 54, 233-351.

Eyde, R. H. (1987) The case for keeping Cornus in the broad Linnaean Sense. *Systematic Botany* 12, 505-518.

Fan, C. and Xiang, Q. Y. (2001) Phylogenetic Relationships within Cornus (Cornaceae) based in 26S rDNA sequences. *American Journal of Botany* 88, 1131-1138.

Fan, H. Y., Hu, Y., Tudor, M. and Ma, H. (1997) Specific interactions between the K domains of AG and AGLs, members of the MADS domain family of DNA binding proteins. *The Plant Journal* 12, 999-1010.

Flanagan, C. A. and Ma, H. (1994) Spatially and Temporally Regulated Expression of the Mads-Box Gene Agl2 in Wild-Type and Mutant Arabidopsis Flowers. *Plant Molecular Biology* 26, 581-595.

Gonzalez Villarreal, L. M. (1996). La Familia Cornaceae en el estado de Jalisco, Mexico. Guadalajara, Universidad de Gualdalajara.

Goto, K. and Meyerowitz, E. M. (1994) Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. *Genes & Development* 8, 1548-1560.

Gu, Q., Ferrandiz, C., Yanofsky, M. F. and Martienssen, R. (1998) The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. *Development* 125, 1509-1517.

Hasebe, M., Wen, C.-K., Kato, M. and Banks, J. A. (1998) Characterization of MADS homeotic genes in the fern Ceratopteris richardii. *Proceedings of the National Academy of Sciences of the United States of America* 95, 6222-6227.

Hileman, L. (2006) Molecular and phylogenetic analyses of the MADS-box gene family in Tomato. *Molecular Biology and Evolution* 23, 2245-2258.

Hill, T. A., Day, C. D., Zondlo, S. C., Thackeray, A. G. and Irish, V. F. (1998) Discrete spatial and temporal cis-acting elements regulate transcription of the Arabidopsis floral homeotic gene APETALA3. *Development* 125, 1711-1721.

Honma, T. and Goto, K. (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* 409, 525-529.

Ingram, G. C., Doyle, S., Carpenter, R., Schultz, E. A., Simon, R. and Coen, E. S. (1997) Dual role for fimbriata in regulating floral homeotic genes and cell division in Antirrhinum. *Embo Journal* 16, 6521-6534.

Irish, V. F. and Yamamoto, Y. T. (1995) Conservation of floral homeotic gene-function between Arabidopsis and Antirrhinum. *Plant Cell* 7, 1635-1644.

Jack, T., Brockman, L. L. and Meyerowitz, E. M. (1992) The Homeotic Gene Apetala3 of Arabidopsis thaliana Encodes a MADS Box and Is Expressed in Petals and Stamens. *Cell* 68, 683-697.

Jaramillo, M. A. and Kramer, E. M. (2004) APETALA3 and PISTILLATA homologs exhibit novel expression patterns in the unique perianth of Aristolochia (Aristolochiaceae). *Evolution & Development* 6, 449-458.

Jenik, P. D. and Irish, V. F. (2001) The Arabidopsis floral homeotic gene APETALA3 differentially regulates intercellular signaling required for petal and stamen development. *Development* 128, 13-23.

Johansen, B., Pedersen, L. B., Skipper, M. and Frederiksen, S. (2002) MADS-box gene evolution-structure and transcription patterns. *Molecular Phylogenetics and Evolution* 23, 458-480.

Kanno, A., Saeki, H., Kameya, T., Saedler, H. and Theissen, G. (2003) Heterotopic expression of class B floral homeotic genes supports a modified ABC model for tulip (Tulipa gesneriana). *Plant Molecular Biology* 52, 831-841.

Kardailsky, I., Shukla, V. K., Ahn, J. H., Dagenais, N., Christensen, S. K., Nguyen, J. T., Chory, J., Harrison, M. J. and Weigel, D. (1999) Activation tagging of the floral inducer FT. *Science* 286, 1962-1965.

Kaufmann, K., Melzer, R. and Theissen, G. (2005) MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* 347, 183-198.

Kim, S., Yoo, M.J., Albert, V. A., Farris, J. S., Soltis, P. S. and Soltis, D. E. (2004) Phylogeny and diversification of B-function MADS-box genes in angiosperms: evolutionary and functional implications of a 260-million-year-old duplication. *American Journal of Botany* 91, 2102-2118.

Kramer, E. M. and Irish, V. F. (2000) Evolution of the petal and stamen developmental programs: evidence from comparative studies of the lower eudicots and basal angiosperms. *International Journal of Plant Sciences* 161, S29-S40.

Kramer, E. M. and Jaramillo, M. A. (2005) Genetic basis for innovations in floral organ identity. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution* 304B, 526-535.

Kramer, E. M., Di Stilio, V. S. and Schluter, P. M. (2003) Complex patterns of gene duplication in the APETALA3 and PISTILLATA lineages of the Ranunculaceae. *International Journal of Plant Sciences* 164, 1-11.

Kramer, E. M., Dorit, R. L. and Irish, V. F. (1998) Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the APETALA3 and PISTILLATA MADS-box gene lineages. *Genetics* 149, 765-783.

Kramer, E. M., Su, H. J., Wu, C. C. And Hu, J. M. (2006) A simplified explanation for the frameshift mutation that created a novel C-terminal motif in the APETALA3 gene lineage. *BMC Evolutionary Biology* 6, 30.

Krizek, B. A. and Meyerowitz, E. M. (1996) The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development* 122, 11-22.

Krogan, N. T. and Ashton, N. W. (2000) Ancestry of plant MADS-box genes revealed by bryophyte (Physcomitrella patens) homologues. *New Phytologist* 147, 505-517.

Lee, I., Wolfe, D. S., Nilsson, O. and Weigel, D. (1997) A LEAFY co-regulator encoded by UNUSUAL FLORAL ORGANS. *Current Biology* 7, 95-104.

Levin, J. Z. and Meyerowitz, E. M. (1995) UFO: an Arabidopsis Gene Involved in Both Floral Meristem and Floral Organ Development. *Plant Cell* 7, 529-548.

Litt, A. (2007). An evaluation of A-function: Evidence from the APETALA1 and APETALA2 gene lineages. *International Journal of Plant Sciences* 168, 73-91.

Litt, A. and Irish, V. F. (2003) Duplication and diversification in the APETALA1/FRUITFULL floral homeotic gene lineage: Implications for the evolution of floral development. *Genetics* 165, 821-833.

Maddison, D. R. and Maddison, W. P. (2000) MacClade 4: Analysis of phylogeny and character evolution. Sunderland, Massachusetts, Sinauer Associates.

Manchester, S. R., Crane, P.R. and Goloneva, L.B. (1999) An extinct genus with affinities to extant Davidia and Camptotheca (Cornales) from the paleocene of North America and Eastern Asia. *International Journal of Plant Science* 160, 188-207.

Mandel, M. A. and Yanofsky, M. F. (1998) The Arabidopsis AGL9 MADS box gene is expressed in young flower primordia. *Sexual Plant Reproduction* 11, 22-28.

Mandel, M. A., Gustafsonbrown, C., Savidge, B. and Yanofsky, M. F. (1992) Molecular characterization of the Arabidopsis gloral homeotic gene APETALA1. *Nature* 360, 273-277.

McGonigle, B., Bouhidel, K. and Irish, V. F. (1996) Nuclear localization of the Arabidopsis APETALA3 and PISTILLATA homeotic gene products depends on their simultaneous expression. *Genes & Development* 10, 1812-1821.

Meyerowitz, Bowman, J. L., Brockman, L. L., Drews, G. N., Jack, T., Sieburth, L. E. and Weigel, D. (1991) A Genetic and Molecular Model for Flower Development in Arabidopsis thaliana. *Development* 1, 157-167

Münster, T., Pahnke, J., Di Rosa, A., Kim, J. T., Martin, W., Saedler, H. and Theissen, G. (1997) Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. *Proceedings of National Academy of Sciences, USA*, 94, 2415-2420.

Murrell, Z. (1994) Dwarf Dogwoods: Intermediacy and the Morphological Landscape. *Systematic Botany* 19, 539-556.

Murrell, Z. E. (1993) Phylogenetic relationships in Cornus (Cornaceae). *Systematic Botany* 18, 469-495.

Nakamura, T., Fukuda, T., Nakano, M., Hasebe, M., Kameya, T. and Kanno, A. (2005) The modified ABC model explains the development of the petaloid perianth of Agapanthus praecox ssp orientalis (Agapanthaceae) flowers. *Plant Molecular Biology* 58, 435-445.

Ng, M. and Yanofsky, M. F. (2001) Activation of the Arabidopsis B class homeotic genes by APETALA1. *Plant Cell* 13, 739-753.

Parcy, F., Nilsson, O., Busch, M. A., Lee, I. and Weigel, D. (1998) A genetic framework for floral patterning. *Nature* 395, 561-566.

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E. and Yanofsky, M. F. (2000) B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 405, 200-203.

Perbal, M. C., Haughn, G., Saedler, H. and Schwarz-Sommer, Z. (1996) Non-cell-autonomous function of the Antirrhinum floral homeotic proteins DEFICIENS and GLOBOSA is exerted by their polar cell-to-cell trafficking. *Development* 122, 3433-3441.

Pollock, R. (1991) Human SRF-related proteins: DNA-binding properties and potential regulatory targets. . *Genes & Development* 5, 2327-2341.

Purugganan, M. D., Rounsley, S. D., Schmidt, R. J. and Yanofsky, M. F. (1995) Molecular Evolution of Flower Development - Diversification of the Plant Mads-Box Regulatory Gene Family. *Genetics* 140, 345-356.

Ramakers, C., Ruijter, J. M., Deprez, R. H. L. and Moorman, A. F. M. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* 339, 62-66.

Riechmann, J. L., Krizek, B. A. and Meyerowitz, E. M. (1996) Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proceedings of National Academy of Sciences*, USA 93, 4793-4798.

Rijpkema, A. S., Royaert, S., Zethof, J., van der Weerden, G., Gerats, T. and Vandenbussche, M. (2006) Analysis of the Petunia TM6 MADS box gene reveals functional divergence within the DEF/AP3 lineage. *Plant Cell* 18, 1819-1832.

Ronse de Craene, L. P. (2004) Floral development of Berberidopsis corallina: a crucial link in the evolution of flowers in the core Eudicots. *Annals of Botany* 94, 741-751.

Rounsley, S. D., Ditta, G. S. and Yanofsky, M. F. (1995) Diverse roles for MADS box genes in Arabidopsis development. *Plant Cell* 7, 1259-1269.

Ruiz Garcia, L., Madueno, F., Wilkinson, M., Haughn, G., Salinas, J. and Martinez Zapater, J. M. (1997) Different roles of flowering-time genes in the activation of floral initiation genes in Arabidopsis. *Plant Cell* 9, 1921-1934.

Samach, A., Klenz, J. E., Kohalmi, S. E., Risseeuw, E., Haughn, G. W. and Crosby, W. L. (1999) The UNUSUAL FLORAL ORGANS gene of Arabidopsis thaliana is an F-box protein required for normal patterning and growth in the floral meristem. *Plant Journal* 20, 433-445.

Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P. J., Hansen, R., Tetens, F., Lonnig, W. E., Saedler, H. and Sommer, H. (1992) Characterization of the Antirrhinum floral homeotic MADS-box gene deficiens - Evidence for DNA-binding and autoregulation of its persistent expression throughout flower development. *Embo Journal* 11, 251-263.

Shan, H. Y., Zhang, N., Lui, C. J., Xu, G. X., Zhang, J., Chen, Z. D., Kong, H. Z. (2007) Patterns of gene duplication and functional diversification during the evolution of the AP1/SQUA subfamily of plant MADS-box genes. *Molecular Phylogenetics and Evolution* 44, 26-41.

Simon, R., Carpenter, R., Doyle, S., and Coen, E. (1994) Fimbriata controls flower development by mediating between meristem and organ identity genes. *Cell* 78, 99-107.

Soltis, D. E. (2007) The floral genome: an evolutionary history of gene duplication and shifting patterns of gene expression. *Trends in Plant Science* 12, 358-367.

Stellari, G. M., Jaramillo, M. A. and Kramer, E. M. (2004) Evolution of the APETALA3 and PISTILLATA lineages of MADS-box-containing genes in the basal angiosperms. *Molecular Biology and Evolution* 21, 506-519.

Su, K. M., Zhao, S., Shan, H., Kong, H., Lu, W., Theissen, G., Chen, Z., Meng, Z. (2008) The MIK region rather than the C-terminal domains of AP3-like class B floral homeotic proteins determines functional specificity in the development and evolution of petals. *New Phytologist* 178, 544-558.

Sundström, J., Carlsbecker, A., Svensson, M. E., Svenson, M., Johanson, U., Theiben, C. and Engström, P. (1999). MADS-box genes active in developing pollen cones of Norway spruce (Picea abies) are homologous to the B-class floral homeotic genes in angiosperms. *Developmental Genetics* 25, 253-266.

Svensson, M. E., Johannesson, H. and Engström, P. (2000) The LAMB1 gene from the clubmoss, Lycopodium annotinum, is a divergent MADS-box gene, expressed specifically in sporogenic structures. *Gene* 253, 31-43.

Swofford, D. L. (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, Massachusetts, USA, Sinauer Associates, Inc.

Takhtajan, A. L. (1997) Diversity and classification of flowering plants. New York, Columbia University Press.

Theissen, G. and Saedler, H. (2001) Floral quartets. *Nature* 409, 469-471.

Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J. T., Munster, T., Winter, K. U. and Saedler, H. (2000) A short history of MADS-box genes in plants. *Plant Molecular Biology* 42, 115-149.

Trobner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P., Lonnig, W. E., Saedler, H., Sommer, H. and Schwarz-Sommer, Z. (1992) Globosa - a Homeotic Gene Which Interacts with Deficiens in the Control of Antirrhinum Floral Organogenesis. *Embo Journal* 11, 4693-4704.

Tzeng, T. Y., Liu, H. C. and Yang, C. H. (2004) The C-terminal sequence of LMADS1 is essential for the formation of homodimers for B function proteins. *Journal of Biological Chemistry* 279, 10747-10755.

Vandenbussche, M., Theissen, G., Van de Peer, Y. and Gerats, T. (2003) Structural diversification and neo-functionalization during floral MADS-box gene evolution by C-terminal frameshift mutations. *Nucleic Acids Research* 31, 4401-4409.

Vandenbussche, M., Zethof, J., Royaert, S., Weterings, K. and Gerats, T. (2004) The duplicated B-class heterodimer model: Whorl-specific effects and complex genetic interactions in Petunia hybrida flower development. *Plant Cell* 16, 741-754.

Wagner, D., Sablowski, R. W. M. and Meyerowitz, E. M. (1999) Transcriptional activation of APETALA1 by LEAFY. *Science* 285, 582-584.

Weigel, D. and Meyerowitz, E. M. (1993) Activation of floral homeotic genes in Arabidopsis. *Science* 261, 1723-1726.

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F. and Meyerowitz, E. M. (1992) LEAFY controls floral meristem identity in Arabidopsis. *Cell* 69, 843-859.

Winter, K. U., Weiser, C., Kaufmann, K., Bohne, A., Kirchner, C., Kanno, A., Saedler, H. and Theissen, G. (2002) Evolution of class B floral homeotic proteins: Obligate heterodimerization originated from homodimerization. *Molecular Biology and Evolution* 19, 587-596.

Xiang, Q. Y., Moody, M. L., Soltis, D. E., Fan, C. Z. and Soltis, P. S. (2002). Relationships within Cornales and circumscription of Cornaceae - matK and rbcL

sequence data and effects of outgroups and long branches. *Molecular Phylogenetics and Evolution* 24, 35-57.

Xiang, Q. Y., Soltis, D.E., Morgan, D.R. and Soltis, P.S. (1993). Phylogenetic Relationships of Cornus L. sensu ato and Putative Relatives Inferred from rbcL Sequence Data. *Annals of the Missouri Botanical Garden* 80, 723-734.

Xiang, Q.Y., Brunsfeld, S. J., Soltis, D. E. and Soltis, P. S. (1996). Phylogenetic relationships in Cornus based on chloroplast DNA restriction sites: implications for biogeography and character evolution. *Systematic Botany* 21, 515-534.

Xiang, Q.Y., Soltis, D. E. and Soltis, P. S. (1998) Phylogenetic relationships of Cornaceae and close relatives inferred from matK and rbcL sequences. *American Journal of Botany* 85, 285-297.

Yamaguchi, T., Lee, D. Y., Miyao, A., Hirochika, H., An, G. H. and Hirano, H. Y. (2006). Functional diversification of the two C-class MADS box genes OSMADS3 and OSMADS58 in Oryza sativa. *Plant Cell* 18, 15-28.

Zhan, L. (2005) The Evolution of the SEPALLATA Subfamily of MADS-box Genes: A Preangiospem Origin Wit Multiple Duplications Throughout Angiosperm History. *Genetics* 169, 2209-2223.

Zhang W., Xiang Q.Y., Thomas D.T., Wiegmann B.M., Frohlich M.W., Soltis D.E. (2008) Molecular evolution of PISTILLATA-like genes in the dogwood genus Cornus (Cornaceae). *Molecular Phylogenetics and Evolution* 47, 175-95.