ABSTRACT

FENG, CHUNMIAO. Evolutionary Studies of Alangiaceae and Cornaceae – Tracking Evolutionary Histories using Phylogenetic and Comparative Development Approaches. (Under the direction of Dr. Qiu-Yun (Jenny) Xiang and Dr. Robert G. Franks.)

Alangiaceae and Cornaceae are two sister families in the order Cornales, that represent the first branch of the Asterid clade (the largest clade in the Angiosperm tree of life). Alangiaceae consists of approximately 24 species in the single genus *Alangium* that are distributed mainly in tropical and subtropical regions of the Old World, but with many fossil records in the north temperate regions. Cornaceae contains the single dogwood genus *Cornus* distributed mostly in the north temperate regions of the northern hemisphere. The Cornaceae family exhibits four different inflorescence types including paniculate cymes, minidichasiums, umbels and heads that are associated with variations in bract morphology. Phylogeny-based comparative analyses have proven to be powerful tools to track evolutionary histories and are widely applied to test biogeographic, molecular, and morphological evolutionary hypotheses. In my dissertation study, a comparative approach was applied to the evolutionary studies of Alangiaceae and Cornaceae. Phylogenetic analyses were conducted using data from DNA sequences, fossils, and morphology to reconstruct the phylogeny of Alangiaceae. This phylogeny served as a basis to elucidate evolutionary relationships among extant species of Alangiaceae and to evaluate alternative biogeographic hypotheses for this family. For Cornaceae, comparative developmental and gene expression analyses were conducted to test hypotheses on the origin and evolution of umbels, heads and petaloid bracts at the developmental and molecular levels, using a previously established phylogeny of the family. Results of phylogenetic analyses of Alangiaceae suggested that section *Conostigma* is the first diverged lineage, followed by section *Rhytidandra*. Sections

Marlea and *Alangium* are the most advanced groups. Results of biogeographic analyses using the dispersal-vicariance analysis program (DIVA), combining data from divergence time estimation with the MULTIDIVTIME method, indicated that the ancestor of Alangiaceae arose in S.E. Asia in the Late Cretaceous. Several intercontinental migrations involving the margin of the Tethys seaway (TESW), the North Atlantic land bridge or the Bering land bridge and long-distance dispersals and extinction were suggested to be important to shape the current distribution of this family. The results support TESW for plant migration of thermophilic (including evergreen) taxa in the early Tertiary.

Our comparative developmental study of Cornaceae characterized 24 developmental events for six species of *Cornus* that represent the four major inflorescence types. Seven of these developmental events are identified as important to the development and evolution of umbels and heads. Ancestral character reconstruction by tracing these developmental events on the phylogenetic tree suggests that heads and umbels in *Cornus* evolved independently from paniculate cymes through an umbellate dichasium ancestor involving multiple evolutionary changes. We completed a comparative developmental analysis of the bract development and combined this analysis with evolutionary analyses of the gene sequence and expression patterns of MADS-box B class genes (*PI* and *AP3* homologs) known to control petal organ identify in model organisms. These analyses support the independent developmental origins of petaloid bracts in *C.canadensis* and *C.florida*. Furthermore, we detected divergence in selection, co-evolution partnership and expression patterns between the paralogs of duplicated *PI* homologs in *C.florida*. The data suggest divergent molecular mechanisms underlying the bract petaloidy in the two lineages and indicate that activities of B-class genes may have contributed to the development of petaloid bracts in *C.florida*. To

facilitate the validation of the functions of these genes, and other yet to be identified candidate genes in the future, an efficient regeneration system and a preliminary genetic transformation system have been established in *C. canadensis*.

Evolutionary Studies of Alangiaceae and Cornaceae – Tracking Evolutionary Histories using Phylogenetic and Comparative Development Approaches

by
Chunmiao Feng

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Plant Biology

Raleigh, North Carolina

2011

APPROVED BY:

_____	_____
Dr. Qiu-Yun (Jenny) Xiang Committee Chair	Dr. Robert G. Franks Committee Co-Chair
_____	_____
Dr. Dominique (Niki) Robertson	Dr. Deyu Xie

DEDICATION

To my family

BIOGRAPHY

Chunmiao Feng was born in Shaoxing, Zhejiang, a beautiful city in southeast China. She has been very interested in plant biology since childhood. She received her Bachelor's Degree from Fudan University in 2005. She then Joined Dr. Jenny Xiang's lab in the Department of Plant Biology at North Carolina State University in August 2005. Chunmiao learned various methods on computational evolutionary biology and conducted research on the phylogeny and biogeography of Alangiaceae under the direction of Dr. Jenny Xiang during the first year of her graduate study. She became interested in evolutionary developmental genetics of dogwood inflorescence diversity later. She studied the evolution of four inflorescence types and petaloidy bracts in dogwood groups at the developmental and molecular levels, and established an efficient regeneration system for *Cornus canadensis* for future functional study under the direction of Dr. Jenny Xiang and Dr. Robert G. Franks.

ACKNOWLEDGMENTS

I would to express my sincere gratitude to my committee: Dr. Jenny Xiang, Dr. Robert Franks, Dr. Deyu Xie and Dr. Niki Robertson, for their guidance during my time in graduate school here at North Carolina State University. I am extremely grateful to my advisor, Dr. Jenny Xiang, for providing me exciting and challenging projects and guiding my research with her knowledge and passion. I especially appreciate her encouraging me to conduct independent research and interact with researchers across the country. My gratitude also goes to my co-advisor Dr. Robert Franks, who is always patient and supportive, and gave me very helpful advices when I had difficulties. I am also grateful to Dr. Deyu Xie and Dr. Niki Robertson for their encouragement and insightful advices. My committee exemplified what a researcher should be with their rigorousness and enthusiasm.

I also would like to give special thanks to Dr. Ronda Qu for his support and advice on plant regeneration and transformation. Although he is not on my committee, he spent a lot of time and is always there to help me when I encounter difficulties. Dr. Steven R. Manchester has also helped me a lot in checking *Alangium* fossil data and improving my English for my first paper. I would like to thank Dr. David E. Boufford for his assistance on field collection and kindness.

Many thanks also to the current and former members of the Xiang lab, David Thomas, AJ Harris, Dr. Wenheng Zhang, Yuelong Guo, Ashley Brooks, Xiang Liu, Dr. Norman Douglas, Wade Wall, Amanda Saville, Yi Yu, Juan Liu and Dr. Jian Zhang, for their

friendship and sharing their expertise. I especially appreciate Wenheng for her help in study, research and life. I also appreciate the technical assistances from David Thomas and Xiang Liu, and helpful discussion with Dr. Norman Douglas and Wade Wall. I also would like to give special thanks to Sridevi Azhakanandam and Dr. Fang Bao from the Franks lab for their technical help on *in situ* hybridization and SEM study. My gratitude also goes to Dr. Lili Zhou from the Xie lab who taught me techniques on tissue culture, Hainian Zeng and Mingzhu Shi from the Xie lab for their valuable comments on my research, Dr. Shujie Dong, Dr. Ruyu Li, Dr. Bingwu Wang and Dr. Wanjun Zhang from the Qu lab for their technical help on plant transformation.

My thanks further extend to Dr. Alexander Krings for advising on handing herbarium specimens, Dr. John M. Mackenzie Jr. and Ms. Valerie Knowlton at Center for Electron Microscopy at NCSU for valuable advice on SEM analysis, Dr. Carole Saravitz and Dr. Janet Shurtleff at phytotron on advising on caring dogwood plants, Dr. Dennis Warner and Mark Weathington at JC Raulston Arboretum for providing dogwood materials, Dr. Becky Boston, Dr. Richard Blanton, Ms. Sue Vitello and Ms. Carol Apperson for keeping me on the right track in the administrative process. I also would like to say “thank you” to Ms. Donna Wright for giving me excellent teaching advice. Finally, my deepest gratitude goes to my parents and my husband.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY DATA

LIST OF ABBREVIATIONS

aa - amino acid
ANOVA - analysis of variance
AP3 - *APETALA3*
BAP - 6-Benzylaminopurine
BB - big-bracted dogwoods
bp - base pair
BW - blue- or white- fruited dogwoods
CC - cornelian cherry
CI - consistency index
CorPI - *PISTILLATA*-like genes in *Cornus*
CorPI-A - A copy of *CorPI* in *Cornus*
CorPI-B - B copy of *CorPI* in *Cornus*
cpDNA - chloroplast DNA
dN - nonsynonymous substitution
DE - developmental event
DIVA - dispersal-vicariance analysis
DMSO - dimethyl sulfoxide
DNA - deoxyribonucleic acid
dNTP - dinucleotide tri-phosphate
dS - synonymous substitution
DW - dwarf dogwoods
FAA - Formaldehyde
GA - Gibberellic acid, C19H22O6
GTR - general time reversible
IBA - Indole-3-butyric acid
IBM - inflorescence branch meristem
IM - inflorescence meristem
ITS - inter transcribed spacer
Ka - nonsynonymous substitution
kb - kilobase
Ks - synonymous substitution
LRT - likelihood-ratio test
matK- maturase K
MCMC - Markov Chain Monte Carlo
ML - maximum likelihood
MP - maximum parsimony
MS - Murashige and Skoog
MYA - million years ago

NAA -1-Naphthaleneacetic acid
PAML - phylogenetic analysis by maximum likelihood
PAUP - phylogenetic analysis using parsimony
PCR - polymerase chain reaction
PI - PISTILLTA
PVPP - Polyvinylpolypyrrolidone
rbcL - ribulose bisphosphate carboxylase large subunit
RI - retention index
RSA – relative surface accessibility
SD - standard deviation
SEM- Scanning Electron Microscopy
TBR - tree-bisection-reconnection

Chapter I

Phylogeny and biogeography of Alangiaceae (Cornales) inferred from DNA sequences, morphology, and fossils

Chun-Miao Feng[1]*, Steven R. Manchester[2], Qiu-Yun (Jenny) Xiang[1]

[1]Department of Plant Biology, North Carolina State University, Raleigh, NC 27695, USA

[2]Florida Museum of Natural History and Department of Botany, University of Florida, Gainesville, FL 32611, USA

Published in Molecular Phylogenetics and Evolution, 2009, 51: 201-214

Chunmiao did all of the work reported in this paper, Dr. Steven Manchester and Dr. Jenny Xiang provided scientific advice and guidance.

Molecular Phylogenetics and Evolution 51 (2009) 201–214

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogeny and biogeography of Alangiaceae (Cornales) inferred from DNA sequences, morphology, and fossils

Chun-Miao Feng [a,*], Steven R. Manchester [b], Qiu-Yun (Jenny) Xiang [a]

[a] Department of Plant Biology, North Carolina State University, Campus Box 7612, Raleigh, NC 27695, USA
[b] Florida Museum of Natural History and Department of Botany, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 21 April 2008
Revised 12 December 2008
Accepted 23 January 2009
Available online 6 February 2009

Keywords:
Alangiaceae
Biogeography
Divergence time
Morphology
Fossils
Phylogeny

ABSTRACT

Alangiaceae, in the basal Asterid clade Cornales, consists of only one genus, *Alangium*. The genus has approximately 24 species distributed mainly in tropical and subtropical regions of the Old World and is well represented in the Tertiary fossil record of the northern hemisphere. We conducted phylogenetic and biogeographic analyses for *Alangium* by integrating data from DNA sequences, morphology, and fossils to evaluate systematic and biogeographic hypotheses. The results largely agree with the traditional classification of four sections within the genus, and suggest sects. *Conostigma* and *Rhytidandra* are successive sister taxa to a clade containing sects. *Marlea* and *Alangium*. Our results also indicate that the widespread species *A. chinense* consists of at least two lineages meriting recognition as distinct species. Biogeographic analysis using DIVA and divergence time dating with the Bayesian method (MULTIDIVTIME) resolved the ancestor of *Alangium* as being in S.E. Asia in the Late Cretaceous. Several intercontinental migrations involving the margin of the Tethys seaway (TESW), the North Atlantic land bridge (NALB) or the Bering land bridge (BLB), and long-distance dispersals are suggested. The results support TESW for plant migration of thermophilic (including evergreen) taxa in the early Tertiary.
© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Alangiaceae, in the basal Asterid clade, consists of only the genus *Alangium*, with approximately 24 species mostly distributed in tropical and subtropical regions of eastern and southeastern Asia, extending to most of the western and southwestern Pacific Islands. Two species, *A. salviifolium* (L.) Wangerin and *A. chinense*, (Loureiro) Harms, extend from southeastern Asia (S.E. Asia) to Africa, while *A. salviifolium* and *A. villosum* (Blume) Wangerin extend to Australia, Fiji, and New Caledonia, and one species, *A. grisolleoides* Capuron, is endemic to Madagascar. *A. chinense* and *A. kurzii* Craib also extend to temperate eastern Asia. Only one species, *A. platanifolium* (Siebold and Zuccarini) Harms, occurs solely in the temperate regions of eastern China and Japan, extending northward to Russia (Table 1, Bloembergen, 1939; Cronquist, 1981; Fedina et al., 2002). All species are trees or shrubs, with the exception of the liana *A. kwangsiense* Melch. Some species, such as *A. salviifolium*, *A. chinense*, and *A. platanifolium*, have a long history of medicinal uses in eastern and southeastern Asia (Jain et al., 2005).

Alangium is characterized by the combination of alternately arranged simple leaves, 5-9-merous polypetalous epigynous flowers with stamens equal or half in number to the petals, a 1- or 2-carpellate ovary and a drupe-like fruit. The genus has usually been treated in or close to Cornaceae. It was included as a member of Cornaceae by Harms (1897), but since 1910 (Wangerin, 1910) the genus has been commonly recognized as a monogeneric family in the Cornales (Takhtajan, 1980, 1997; Cronquist, 1981; Thorne, 1992). Recent molecular phylogenetic studies confirmed that the genus is a member of Cornales and sister to the dogwood genus *Cornus* (Chase et al., 1993; Xiang et al., 1998, 2002; Fan and Xiang, 2003).

The intrageneric classification of Alangiaceae has been somewhat debated. Formerly, the family included the genera *Marlea* and *Rhytidandra* as well as *Alangium*, but the first two are now treated as sections within *Alangium*. In an early revision of *Alangium* (Bloembergen, 1939), the genus contained 18 species (including 16 subspecies and 6 varieties), divided into four sections: *Alangium* Bail., *Marlea* Baill., *Rhytidandra* Baill., and *Conostigma* Bloemb. The sectional classification was mainly based on the number of stamens and morphology of style, stigma, and embryo (Table 2). Based on pollen morphology, however, Reitsma (1970) questioned this scheme and regarded sects. *Marlea* and *Conostigma* as artificial groups. For example, pollen types of two species in sect. *Marlea*, *A. griffithii* and *A. scandens* were more similar with that of sect. *Conostigma.* The pollen type of *A. chinense* from Africa also treated in sect. *Marlea*, more closely resembled pollen of sect. *Alangium*.

Three additional species were described after Bloembergen's work: *Alangium qingchuanense* He, *A. grisolleoides* Capuron, and *A. circulare* Stone and Kochummen (He et al., 1996; Capuron, 1962;

* Corresponding author. Fax: +1 919 515 3436.
E-mail address: cfeng@ncsu.edu (C.-M. Feng).

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.01.017

Table 1
List of *Alangium* species indicating geographic distribution.

Section	Species	Distribution
Alangium Baillon	*Alangium salviifolium* (L.) Wangerin	Africa, India, Ceylon, China, Indochina, Siam, India, Andaman islands, Nicobar Islands, Philippines, Malay, Sumatra, Java, Kangean Islands, Lesser Sunda Islands, Celebes, Salajar, New Guinea (ABCF)
	Alangium longiflorum Merrill	Philippines, Borneo (B)
	Alangium hirsutum Bloembergen	Borneo, Sumatra (B)
	Alangium brachyanthum Merrill	Philippines (B)
Marlea Baillon	*Alangium platanifolium* (Siebold and Zuccarini) Harms	China, Korea, Japan, Formosa, Russian Far East (B)
	Alangium alpinum Smith and Cave	Asia (B)
	Alangium chinense (Loureiro) Harms	Africa, India, China, Indochina, Siam,, Japan, Formosa, Philippines, Java, Lesser Sunda Islands (ABC)
	Alangium rotundifolium (Hasskarl) Bloembergen	Malay, Sumatra, Borneo, Java and Lesser Sunda Islands (B)
	Alangium kurzii Craib	China, Burma, Indochina, Siam, Malay, Java (B)
	Alangium faberi Oliv.	China (B)
	Alangium kwangsiense Melch. In Notizbl	China (B)
	Alangium barbatum (C.B. Clarke) Harms	India, China, Indochina (BC)
	Alangium scandens Bloembergen	Sumatra, Borneo (B)
	Alangium griffithii (Clarke) Harms	Siam, Malay Peninsula, Sumatra, Borneo, Java, Celebes, Moluccas (B)
	Alangium yunhuiense R.H. Miau and W.B. Liao	China (B)
	Alangium qingchuanense M.Y. He	China (B)
Rhytidandra Baillon	*Alangium villosum* (Blume) Wangerin	Java, Queensland, Philippine, New Guinea, Solomon Islands, New Caledonia, Fiji, Eastern Australia (BF)
	Alangium grisolleoides Capuron	Madagascar (A)
Conostigma Bloembergen	*Alangium nobile* (Clarke) Harms	Malay Peninsula. Penang, Sumatra Palembang, Borneo Sarawak, Sumatra (B)
	Alangium havilandii Bloembergen	Borneo Sarawak (B)
	Alangium maliliense Bloembergen	Celebes (B)
	Alangium ridleyi King	Indochina, Malay Peninsula, Sumatra (B)
	Alangium. javanicum (Blume) Wangerin	Celebes, New Guinea (B)
	Alangium circulare B.C. Stone and K. M. Kochummen	Sarawak (B)

Geographic summary designations of each species, used in the DIVA analysis, are indicated in parentheses following the species names. A = Africa, Madagascar; B = eastern and southeast Asia (e.g., China, Japan, Korea, Vietnam, Java, Sumatra, Philippines, Indochina, Borneo, Malay, Sarawak); C = India; F = Australia, New Guinea, Fiji and New Caledonia).

Stone and Kochummen, 1975). *Alangium circulare* from Malaysia was placed in sect. *Conostigma* (Stone and Kochummen, 1975; Govindarajalu, 1979), while *A. qingchuanense* from China was placed in sect. *Marlea* (He et al., 1996). However, the position of the Madagascan species *A. grisolleoides* has been debated. Capuron (1962) and Eyde (1968) placed it in sect. *Conostigma* based on flower and fruit characteristics (papillose epidermis on anthers and style, massive unilocular endocarp with many fine longitudinal ridges on the surface), but this placement is not supported by pollen morphology (Reitsma, 1970). Govindarajalu (1979) placed the species in sect. *Rhytidandra* based on foliar and node and internode anatomic characters (e.g., presence of two layered palisade tissue, unicellular hair, monostelic petiolar vasculature, diffuse-in-aggregate xylem parenchyma, thick-walled libriform fibers, heterogeneous type I rays, short vessel members with simple oblique perforations and alternate crowded intravascular pits).

The wood, fruit, and pollen of *Alangium* are sufficiently distinctive to facilitate recognition of the genus in the fossil record. Hence, the past geographical distribution of *Alangium* can be traced with the help of paleobotanical data. Although extant species are now mostly found in the tropics, early Tertiary fossils from Europe and North America are documented. Two of the oldest records are from the Early Eocene of southern England based on fruit remains (Reid and Chandler, 1933) and from the Middle Eocene in western North America (Oregon) based on wood (Wheeler and Manchester, 2002) and fruit remains (Manchester, 1994). These and other occurrences, such as pollen and fruits from the Miocene of eastern North America (Vermont; Eyde et al., 1969; Traverse, 1994), and various Eocene to Miocene sites in Europe (Krutzsch, 1962; Reid and Chandler, 1933; Mai, 1976; Eyde et al., 1969) indicate a much wider geographic distribution in the past than today.

The fossil record shows that sect. *Marlea*, now distributed in eastern and southeastern Asia and Africa, occurred in Europe and North America during much of the Cenozoic era (Eyde et al., 1969; Krutzsch, 1962; Manchester, 1994; Traverse, 1955; Tables 3 and 4), and sect. *Conostigma*, now confined to Indochina, Sumatra and eastward, also occurred in eastern Asia (especially in China) during the Paleocene to Eocene (Morley, 1982; Song et al., 2004; Table 4). Records from fossil pollen indicate that sect. *Alangium* was already present in Australia as early as the Oligocene (Martin et al., 1996) and in India during the Miocene (Phadtare and Thakur, 1990). Fossils of sect. *Rhytidandra* were found in Australia in the Early Eocene (Martin et al., 1996).

Such a distributional pattern, with extant species absent and only fossils found in certain areas (Europe and North America), reflecting warmer climatic intervals of the Tertiary in these areas, also occurs in other groups of seed plants, including *Actinidia*, *Mastixia*, *Diplopanax*, *Amentotaxus*, *Corylopsis*, *Platycarya*, *Toricellia* (Manchester, 1999). Thus, understanding the biogeographic history of *Alangium* may shed light on the past floristic exchanges between temperate and tropical floras in general, as well as plant migrations between Gondwana and Laurasian lands.

A few hypotheses have been proposed regarding the biogeographic origin and migration of species in *Alangium* based on the study of fossils. Reitsma (1970), comparing the pollen morphology of modern and fossil *Alangium*, concluded that most fossil pollen types were similar to those of sect. *Marlea* and suggested that the fossils represented the ancestors of the genus as a whole. He placed the pollen of *Alangium* into 15 types and pollen of *A. chinense* into four types designated A–D. He considered the B type of *A. chinense* (3-, 4-zonocolporate; lalongate colpus or porus; costae faint, broad, or small; muri simplicolumellate; sexine thinner than nexine; capita rounded; columellae short and thin) to be more primitive than the other types, based on comparison with the fossil pollen grains. The A type pollen of *A. chinense* is restricted to India and Asia, while B type is restricted to India, C type to Africa, and D type from Philippines to Java. Based on this evidence, Reitsma (1970) suggested that *A. chinense* originated somewhere in India and subsequently migrated west to Africa either via the Lemurian

Table 2
Morphological comparison among four sections of *Alangium* following Bloembergen (1939) and Eyde (1968) and the outgroup.

Sections	Number of stamens	Stigma	Style	Inflorescences	Leaves	Radicle	Endosperm
Alangium	2–6× as many as petals	Capitate, not 4-lobed	Cylindrical, glabrous	Sessile or shortly pedunclulate	Symmetrical	1.5× as long as the cotyledons	Superficially grooved
Marlea	As many as petals	Capitate, 4-lobed	Cylindrical, glabrous or hairy	Distinctly pedunclulate	Asymmetrical	Shorter than half the length of the cotyledons	Smooth
Rhytidandra	As many as petals	Deeply split into 2 or 3 long lobes	Cylindrical, glabrous or hairy	distinct peduncle	Symmetrical	Shorter than half the length of the cotyledons	Smooth
Conostigma	As many as petals	Obtuse-conical, with four longitudinal stripes	Cylindrical or clavate, hairy	Sessile with few long primary branches or one long-peduncled branch	Nearly symmetrical	Shorter than half the length of the cotyledons	Smooth
Cornus (outgroup)	As many as petals	Single, not obviously lobed or split	Cylindric or clavate, hairy	Sessile or distinctly pedunclulate	Symmetrical	Shorter than half the length of cotyledons	Smooth

land bridge (a lost continent connecting India and South Africa, proposed before the theory of continental drift was accepted) or via western Asia, and to the east via the Asiatic mainland, Taiwan, the Philippines, the Moluccas and the Lesser Sunda Islands to Java (Fig. 1).

Morley (1982) reviewed the fossil record of *Alangium* and proposed another scenario for the evolutionary and geographic history of the genus. He suggested that sects. *Conostigma* and *Marlea* were more or less of equal antiquity. All reliable Tertiary *Alangium* records from outside S.E. Asia were derivatives of the *Marlea* type, and sections other than *Marlea* never extended beyond their present ranges. He considered that the persistence through much of the Malesian Tertiary suggested that these sections may have been separated for a long period of time (but did not specify a date). The past and present distributions of all sections in S.E. Asia and elsewhere led him to conclude that the genus evolved solely as a tropical megathermally adapted genus in S.E. Asia. Soon after its origin, it diverged into two separate evolutionary lines, one tropical, now represented by sect. *Conostigma*, and one subtropical to temperate, represented by sect. *Marlea* (Fig. 1). He did not make any comments regarding the origin and evolution of sects. *Alangium* and *Rhytidandra* due to the lack of fossils belonging to these two sections at that time. Several years later, Krutzsch (1989) briefly reviewed the pollen fossil record of *Alangium* and proposed a contrasting hypothesis with an origin of *Alangium* in North America during Paleocene followed by migration into India and Indonesia at the beginning of the Eocene. The genus became widespread and diversified in S.E. Asia in the Mid Tertiary. However, the North American origin advocated by Krutzsch was heavily based on Margocolporites lihokus, a fossil pollen species from the Paleocene (Srivastava, 1972) and Late Cretaceous (Farabee and Canright, 1986 of western North America, that Krutzsch identified, without justification, to Alangiaceae. (see Section 2). In 1996, two Tertiary pollen fossils from sect. *Rhytidandra* were found in east coast of Australia (Martin et al., 1996). Based on this evidence, Martin et al. (1996) proposed that sect. *Rhytidandra* originated in Australia and subsequently migrated to Java and the Philippines (Fig. 1).

These systematic and evolutionary hypotheses about *Alangium* have never been tested with phylogenetic analyses. In this study, we conducted a phylogenetic analyses based on morphological and DNA sequence data for the genus. We combined fossil, morphological, and molecular data to infer the biogeographic history of *Alangium*. The goals are to elucidate relationships among extant species of *Alangium*, to evaluate the classification schemes and inter-sectional relationships previously proposed for the genus, and to test the evolutionary hypotheses regarding the evolution and biogeography in *Alangium*.

2. Materials and methods

2.1. Sampling and data collected for the molecular study

We included 34 samples representing 13 species from all four sections of *Alangium* in the molecular study. This sample includes

Table 3
Morphological characters scored for fruit fossil and living species.

Species	Age	Locality	Section and similar living species	Characters			
				1	2	3	4
A. rotundicarpum (N15)	Middle Eocene	Oregon[c]	*Marlea: A. kurzii*	1	0	0	0
A. vermontanum	Middle Eocene	North America[b]	*Marlea: A. chinense*	0	0	1	02
A. eydei	Miocene	Oregon[c]	*Marlea: A. chinense*	0	0	1	0
A. begoniifolium	Pliocene	Japan[a]	*Marlea: A. chinense*	0	0	0	?
A. platanifolium	Pleistocene	Japan[a]	*Marlea: A. platanifolium*	0	0	2	?
A. chinense (sect. Marlea)	Living species	China, India, Africa, Java		0	0	1	01
A. platanifolium (sect. Marlea)	Living species	China, Japan		0	0	2	1
A. kurzii (sect. Marlea)	Living species	China, Burma, Indochina		0	0	0	01
Sect. *Alangium*	Living species	China, India, Africa, South Pacific Islands		1	1	2	0
Sect. *Rhytidandra*	Living species	Australia, Madagascar		1	1	2	2
Sect. *Conostigma*	Living species	South Pacific Islands		2	2	2	0

Characters, character states, and scores: (1) fruit base and apex-39: ovate and obtuse at the base, pointed at the apex, 0, elliptic and obtuse at the both base and apex, 1; elliptic and acute at both the base and apex, 2. (2) length of dried fruits-40: small (5–10 mm × 2–7 mm), 0; medium (10–30 mm × 7–14 mm), 1; large (23–30 mm × 11–20 mm), 2. (3) Number of locules-48: two, equal, 0; two, unequal, 1; (4). Ornamentation on fruits-49: with ridges, 0; grooved, 1; deep pitted 2. Question marks "?" indicate missing data, number behind each character corresponded to that in morphological data matrix (Supplementary material 1).
References: [a]Miki, 1956; [b]Eyde et al., 1969; [c]Manchester, 1994.

Table 4
Pollen fossils of *Alangium* and characters scored for pollen for phylogenetic analyses.

Species	Age	Locality (and reference)	Section, and similar living species	Characters						
				1	2	3	4	5	6	7
Alangiopollis barghoornianus	Oligocene to Miocene	North America[a]; German[b]; China[c]	*Marlea: A. kurzii*	0	01	0	0	0	0	0
Lanagiopollis javanicoides (N1)	Paleocene to Eocene	China[d]	*Conostigma: A. javanicum*	0	1	?	?	0	0	1
Clavastephanocolporites meleosus (N19)	Late Early Eocene (55 mya)	Australia[e]	*Rhytidandra: A. villosum*	1	1	?	?	0	2	1
Clavastephanocolporites blysus (N4)	Early Oligocene to Late Miocene	Australia[e]	*Alangium: A. salviifolium*	2	1	?	?	?	2	1
Alangiopollis eocaenicus (N9)	Early Eocene	Central Europe[f]	*Marlea: A. chinense A*	0	01	?	0	1	0	0
Lanagiopollis fossils	Middle to Late Eocene	Southeast Mankalihat Peninsula, Kalimantan[g]; Central Java[h,i]; India, Malesia[l]	*Conostigma: A. javanicum*	0	01	0	0	1	0	1
Lanagiopollis emarginatus	Early Pliocene	Brunei[j,k]	*Conostigma A. havilandii*	0	01	?	0	1	0	1
Alangiopollis sp.1	Pliocene	The northern part of the South China Sea[l]	*Marlea: A. kurzii*	0	01	?	1	1	1	0
Palaeoalangiumpollis zoniporatus	Miocene	India[m]	*Alangium: A. salviifolium*	2	1	?	?	1	0	1

Characters, character states, and scores: (1) Apertures-50: 3 0; 4 1; 5 2; 6 3; (2) Ectoapertures-51: colpate 0; porate 1; (3) shape of the ends of ectoapertures-54: acute 0; round 1; (4) margin of ends of ectoapertures-55: composed of sexinous muri of the reticulum with distinct margin 0; absent 1; (5) Costae of endoapertures-57: faint 0; heavy 1; (6) Ornamentation-58: consisting of lumina and muri (reticulate, rugulate and striate patterns) 0; consisting of lumina and muri (two different patterns) 1; consisting of detached warty elements (consisting of a hemispherical, verruca- or gemma-like sexine part, subtended by thin and short columellae) 2; (7) Sexine: nexine ratio-59: >1 0; <1 1; =1 2 Question marks "?" indicate missing data, number behind each character corresponded to that in morphological data matrix (Supplementary material 1).
References: [a]Traverse, 1955; [b]Krutzsch, 1962; [c]Liu, 1986; [d]Song et al., 2004; [e]Martin et al., 1996; [f]Krutzsch, 1962; [g]Muller, 1964; [h]Baumann et al., 1972; [i]Hartono, 1969; [j]Leichti et al., 1960; [k]Fui, 1978; [l]Morley, 1982; [m]Phadtare and Kulkarni, 1983.

nearly all widely accepted extant species and multiple populations of widely distributed species (Appendix 1). Species not sampled were mostly from regions to which we did not have easy access (Malay, Sumatra, Borneo, Java, Philippines, etc.). Herbarium specimens were usually old and rare, and attempts to extract DNA from some of the dry specimens were not successful. Three species of *Cornus*, the sister genus of *Alangium* (Xiang, 1999; Xiang et al., 1993, 1998, 2002, Fan et al., 2003), were chosen as the outgroup for the study.

Total genomic DNA was extracted from leaf materials of herbarium specimens or silica-dried leaves collected from the field using the DNeasy plant Mini kit. The nuclear internal transcribed spacer (ITS) of the ribosomal gene and the chloroplast intergenic spacer between *trn*L and *trn*F and *matK* gene were sequenced. Primers

Fig. 1. Distribution of *Alangium*, fossil (solid circle: Paleocene; open triangle: Eocene; open circle: Oligocene; square: Miocene; solid triangle: Pliocene) and Biogeographic hypotheses proposed for *A. chinensis* by Reitsma (1970) (dotted line with arrow), for sect. *Rhytidandra* (dashed line with arrow) by Martin et al. (1996) and Biogeographic hypothesis proposed for *Alangium* by Morley (1982) (solid line with arrow).

and polymerase chain reaction (PCR) for ITS and *matK* followed Xiang et al. (2006) and described in Sang et al. (1997 for *matK* primers) and in White et al. (1990 for ITS primers). For *trn*L-F, primers and PCR followed Sang et al. (1997).

Sequencing for all regions was done with an automated sequencer ABI377 (Applied Biosystems, Foster City, CA). Detailed methods for preparation of sequencing reactions are described in Fan and Xiang (2001).

2.2. Sampling of taxa and characters for morphological analysis

We included 23 living species in the morphological analysis (Supplementary material 1). The sister genus of *Alangium*, *Cornus*, was again used as the outgroup. *Cornus* was represented by four OTUs, corresponding to the four major groups of *Cornus* (Big-Bracted dogwoods [BB], Dwarf dogwoods [DW], Cornelian Cherries [CC], and Blue- or White-fruited dogwoods [BW]). Specimens were borrowed from the herbaria at Missouri Botanical Garden, Leiden, Harvard University, Kew, and the Smithsonian Institution. A total of 59 morphological characters were selected for the analysis (Supplementary material 1). Among these, 11 reproductive and vegetative characters were also scored for the fossil species. Character states were determined from combination of specimen observations and information from the literature. Species with more than one state for a character were scored as polymorphic for the character (i.e., all states were scored). Non-applicable characters (only in the outgroup) were scored as missing.

2.3. Phylogenetic analyses

Sequences of ITS, *trn*L-F, *matK* were first aligned using Clustal X and then manually adjusted in MacClade 4.0 (Maddison and Maddison, 2000). Phylogenetic analyses of molecular data were performed by using the maximum parsimony and Bayesian methods implemented in PAUP 4.0* (Swofford, 2002) and MrBayes 3.7 (Huelsenbeck and Ronquist, 2003) for ITS, combined *trn*L-F and *matK*, and combined ITS–*trn*L-F–*matK*. Due to the low variation of *trn*L-F and *matK* sequences, not all accessions of a species sequenced for ITS were sequenced for these two regions.

For parsimony analyses, heuristic searches were performed by using random taxon addition of 100 replicates, tree bisection–reconnection (TBR) branch swapping, and MulTrees on settings. Characters were equally weighted and character states were treated as unordered. For Bayesian analyses, Modeltest 3.06 (Posada and Crandall, 1998) was first used to determine the best model of sequence evolution for the ITS, *trn*L-F, and *matK* data. The GTR + I + G model was suggested as the best-fit model for ITS among the 56 models, TIM + G model for *trn*L-F, and TVM + I + G model for *matK* under AIC (Akaike Information Criterion). These models were used in the Bayesian analyses for the sequence data. In Bayesian analysis of morphological data the equal rate model was used (Lewis, 2001). Four heated Markov chains, each starting at a different randomly selected initial state, were run for 1,000,000 generations for all Bayesian analyses. Trees were sampled once every 500 generations. To assess convergence, the program TRACER (Rambaut and Drummond, 2004) was employed. Posterior approximations were derived by combining the output from the four different heated Markov Chain runs. Only trees after the burn-in were used in calculating the posterior approximations.

Phylogenetic analyses of morphological data of extant species were conducted by using Parsimony and Bayesian methods as well, as described above with the following exceptions. In parsimony analyses, taxa that were polymorphic for a character were coded as such in the heuristic searches. In Bayesian analysis, the equal rate model was used (Lewis, 2001).

Bootstrap analyses were conducted to estimate support for the MP trees. MP bootstrap analyses were performed with 1000 replicates with full heuristic searches. Each replicate consisted of 100 heuristic searches with random taxon addition, TBR branch swapping and MulTrees on.

2.4. Integrating fossils and reconstructing ancestral areas

Five fruit fossils (Table 3) and nine pollen fossils (Table 4) of *Alangium* with informative characters were included in this study. The questionable pollen fossil, *Margocolporites lihokus* (Farabee and Canright, 1986), described by Srivastava (1972) from eastern North America was excluded here for following reasons: (1) the sizes of *Alangium* pollen and this fossil are significantly different. All extant *Alangium* pollen grains (Reitsma, 1970) and other pollen fossils (Morley, 1982) are larger than 45 μm, typically larger than 60 μm, while the size of this fossil is only 34 to 45 μm; (2) the colpus membranes are smooth in this fossil, rather than nodular or granular.

To place the fossil species onto the phylogeny, four different approaches were taken: (1) parsimony analysis of morphological characters, (2) parsimony analysis of combined morphological and molecular characters, (3) Bayesian analysis of combined morphological and molecular data and, (4) parsimony analysis of morphological data of fossil and living species with 23 sequenced taxa constrained to the combined ITS–*trn*L-F–*matK* phylogeny (using the 'backbone constraints' PAUP* command) and the other fossil and extant taxa left unconstrained and allowed to insert into their optimal positions. Only the fourth method provided suitable resolution (see Section 4), and was used in the biogeographic analyses. The total evidence tree with the fossils included was therefore used as the framework for the subsequent biogeographic analyses using DIVA.

The dispersal-vicariance analysis program (DIVA, Ronquist, 1997) is an event-based method that optimizes ancestral distributions by assuming a vicariance explanation, but incorporating the potential contribution of dispersal and extinction in shaping the current distributional pattern (Ronquist, 1996, 1997). It minimizes dispersal and extinction events under a parsimony criterion. Because DIVA requires a fully resolved tree, we used a modified version of our total molecular tree in which multiple samples of the same species in a clade were combined into one branch. Terminal polytomous clades containing closely related species from the same geographic area were reduced to simplify the analysis. Living species without molecular data, mostly from sect. *Conostigma*, were placed onto the molecular tree for DIVA from results of the morphological analysis. Geographic areas were chosen and defined to cover the distributions of all extant and fossil species and the outgroup, as designated in Table 1, based on current paleogeographic reconstructions (Whitmore, 1973; Heaney, 2004; Sanmartin and Ronquist, 2004; Scotese, 2004; Lomolino et al., 2006). Landmasses that were separated from Gondwanaland and Laurasia super-continents before the Cretaceous were treated as a separate area. Small landmasses or oceanic archipelagos were included within a larger area when they shared land connections during the Cretaceous and Paleocene or originated after Paleocene. Six geographic endemic areas were defined in this study: Africa and Madagascar, eastern and southeastern Asia, India, Europe, North America, and Australia. Eastern and southeastern Asia includes China, Japan, Korea, Vietnam, Indochina, Borneo, Celebes, Java, Malaysian Peninsula, Philippines, Sumatra, Lesser Sundas (the main islands of the Indonesian region). This region was formed by numerous small terranes from the remnant of Gondwanaland and accreted on to Asia during the Mesozoic through Cenozoic (Metcalfe, 1999). India was recognized as a separate area from S.E. Asia and Africa because it separated from Africa about

160 mya, then moved northward and joined the Asian plate during the Early Eocene (ca. 50 mya, Lomolino et al., 2006; Scotese, 2004; Stanley and Steven, 2005). The Australian region includes Australia, New Guinea, the Solomon Islands, New Caledonia and Fiji. New Caledonia was connected to Australia until the Late Cretaceous. New Guinea joined Australia and then uplift occurred in the Oligocene (30 mya, Kroenke, 1996) or in the lower Miocene (15 mya, Keast, 1996) accompanied by accelerated biotic exchange between New Guinea and Australia (Keast, 1996). A species was coded for presence (1) or absence (0) in each unit area. For the outgroup *Cornus*, the ancestral area (Europe) of the genus inferred from DIVA in Xiang et al. (2006) was applied. DIVA was performed both without constraint of maxareas and with various constraints of maximum areas at each node equaling to 2, 3, 4, and 5 to infer potential vicariance as well as dispersal events and directions. These results were then compared and evaluated by evidence from the tectonic history of Earth and availability of land bridges.

2.5. Estimating divergence times

Information on divergence time is useful, in combination with information on tectonic history, to determine if a dispersal event inferred by DIVA is long distance. To estimate divergence times of lineages, we used the Bayesian method MULTIDIVTIME (Thorne et al., 1998; Kishino et al., 2001, available at http://statgen.ncsu.edu/thorne/multidivtime.html). At first, the BASEML program in PAML (Yang, 1997) was used to analyze the total molecular sequence data using the F84 (Felsenstein, 1984; Kishino and Hasegawa, 1989) model of sequence evolution, with rates allowed to vary among sites according to a discrete gamma (γ) distribution with four rate categories (Yang, 1994). ESTBRANCHES was then run to estimate branch lengths and their associated variance–covariance matrix using the output from previous analysis. The outgroups were pruned from the tree after that. The output of ESTBRANCHES (branch length and variance–covariance matrix) and the reported ages of fossils which were applied as the constraints of minimum age (lower bound) of the stem lineages (Tables 3 and 4 and Fig. 5) were used to estimate divergence time with MULTIDIVTIIME. The default "big time" of the program 250 mya was applied as the maximum age (upper bound) for the node splitting *Cornus* and *Alangium*. This maximum age is conservative, older than the estimated age of Eudicots (Soltis et al., 2005). MULTIDIVTIME was run following the methods described in Xiang et al. (2005) for determining prior and posterior ages, their standard deviations, and the 95% credibility intervals by using the Markov Chain Monte Carlo (MCMC) method. The Markov chain was run for 1,000,000 generations and sampled every 100 generations after an initial burn-in period of 10,000 cycles. To check for convergence of the MCMC, analyses were run from two different starting points.

3. Results

3.1. Sequence data

The aligned ITS matrix consisted of 762 base pairs (bp) and 31 samples, while the aligned matrix of *trn*L-F contained 449 bp and 13 samples. The *matK* matrix consisted of 1796 bp and 13 samples. For the ITS dataset, 398 sites (52.2%) were variable and 238 sites (31.2%) were parsimony-informative characters. They were easily aligned within sections but quite divergent among sections. The combined *trn*L-F and *matK* dataset had 340 (15.1%) variable sites and 197 (8.78%) parsimony-informative sites. Seven indels were found in *matK*, and three in *trn*L-F. The combined ITS, *trn*L-F and *matK* sequence dataset included 23 samples, each with at least one of the two genome sequences and 3007 bp, among which, 706 (23.5%) were variable characters, and 415 (13.8%) were parsimony-informative characters.

3.2. Phylogenetic analyses: molecular and morphological data

The parsimony analyses of ITS sequence data revealed 153 shortest trees with 685 steps, a consistency index (CI) of 0.71 excluding uninformative characters, and a retention index (RI) of 0.86. The strict consensus tree showed completely resolved relationships of species with most nodes well supported by bootstrap analyses (Fig. 2). Sections *Alangium* and *Rhytidandra* of Bloembergen (1939) were monophyletic. Sect. *Marlea* was polyphyletic, because one member, *Alangium griffithii*, grouped with sect. *Alangium*, but with low bootstrap (<50%; Fig. 2) and Bayesian (64%; Fig. 2) support values. Sections *Alangium* and *Marlea* were sister clades together forming a clade sister to sect. *Rhytidandra*. Sect. *Conostigma* (represented by a single species here) was sister to all of these clades. *Alangium grisolleoides* from Madagascar grouped with sect. *Rhytidandra*. Within sect. *Marlea*, three component clades were resolved: *A. chinense–A. platanifolium*, *A. faberi–A. barbatum* and *A. kurzii–A. kwangsiense* (Fig. 2). In the first clade, the widespread species *A. chinense* forms two clades, one represented by an accession from Africa and the other consisting of samples from China, Japan, and India. The multiple accessions of *A. faberi* form a clade that was sister to *A. barbatum*. The Bayesian analysis of the ITS data recovered a tree topologically similar to the strict consensus tree in the parsimony analysis.

The parsimony analyses of cpDNA data (*matK* and *trn*L-F) recovered 6 shortest tree (Fig. 3) of 400 steps, with a CI, with uninformative characters excluded, of 0.86 and RI of 0.91. The ITS tree and the cpDNA trees differed only in the relationship among sects. *Alangium*, *Marlea* and *Rhytidandra*. In the ITS tree, sects. *Alangium* and *Marlea* were sisters, while in the cpDNA tree, sects. *Marlea* and *Rhytidandra* were sisters. Neither relationship, however, was strongly supported (54% and <50% bootstrap values, respectively; 85% and 54% Bayesian support values, respectively). *Alangium griffithii* again grouped with sect. *Alangium* with high bootstrap (94%) and Bayesian (100%) support. Bayesian analysis of the cpDNA data also recovered relationships identical to the MP tree of cpDNA.

Parsimony analyses and Bayesian analyses for the combined ITS-*matK*–*trn*L-F data yielded completely resolved and well supported relationships of sections congruent with those in the ITS phylogeny (Fig. 4), but with higher support at some nodes.

The parsimony analysis of morphological data of living species (Supplementary material 1) resulted in 864 shortest trees, each 292 steps in length with a CI of 0.71, excluding uninformative characters, and an RI of 0.86. The strict consensus tree indicated that three of the four sections were monophyletic, but the monophyly of sect. *Rhytidandra* was not well supported. In this analysis, sect. *Marlea* was sister to the remainder of the genus (Fig. 5), a relationship differing from the one recovered using molecular data, but only weakly supported (69% bootstrap value and no Bayesian support; Fig. 5). Relationships among sects. *Conostigma*, *Rhytidandra* and *Alangium* were unresolved. The four pollen types of *A. chinense* (Reitsma, 1970) formed a clade with no resolution among them (Fig. 5).

The strict consensus trees from the first three analyses for determining the phylogenetic affinities of fossils did not resolve the placements of any of the fossils and also resulted in reduced resolution of the relationships among extant species (trees not shown). The fourth analysis (parsimony analysis of morphological data using the total molecular tree as a backbone constraint) allowed us to place most of the fossils on the total molecular tree (Fig. 6), although with relatively low bootstrap support. Three fossils, *Palaeoalangiumpollis zoniporatus* (Phadtare and Kulkarni, 1983), *Alangium platanifolium* (Miki, 1956) and *Alangiopollis* sp.1

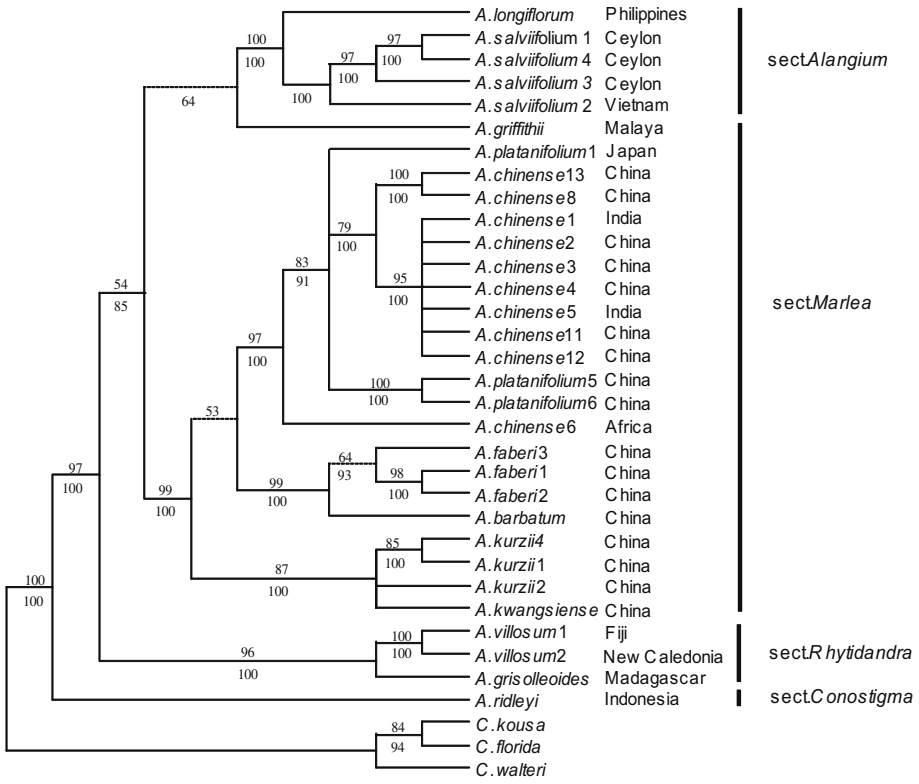


Fig. 2. One of the most parsimonious trees of the 153 shortest trees (each is 685 steps) resulting from parsimony analysis of ITS sequences which was similar to the Bayesian tree. Dashed line showed the missing branch in the strict consensus tree. Numbers above the branches are bootstrap values; numbers below the branches are Bayesian posterior probabilities from the corresponding Bayesian analysis. Geographic location of each sample is shown after the taxon name. *A.* = *Alangium*, *C.* = *Cornus*.

(Morley, 1982), could not be placed on the total molecular tree using this method. They did not group with any of the ingroup clades, but resolved as parts of a polychotomous clade within the ingroup. These three fossils were therefore mapped onto the total molecular tree according to our subjective assessment of their affinities based on their pollen or fruit morphology as reported in the literature (Tables 3 and 4).

3.3. Results of DIVA and MULTIDIVTIME analysis

Ancestral area reconstruction using DIVA under different constraints of maximum areas (2, 3, 4, 5) or without constraint required 19 (maxareas = 2 or 3) or 18 (maxareas = 4, 5 or unconstrained) dispersal events. Results from analysis with constraints of maximum areas of 4, 5 and unconstrained were identical at all nodes. The ancestral areas at most nodes were also the same as those optimized under constraints of maximum areas of 2 and 3. Differences were found at seven nodes, including the three deepest nodes (nodes 0–2 in Fig. 6) and four upper nodes (nodes 5, 22, 23, 24; Fig. 6). For most nodes, various analyses resolved a single distribution in one or two areas. The node times and their 95% credibility intervals estimated from divergence time analysis using MULTIDIVTIME are shown in Table 5 and Fig. 6. For nodes with multiple optimal distributions suggested by DIVA, the node time was used to evaluate alternative distributions. Most of the wide distributions at these nodes involved non-connected areas (either areas not adjacent to each other or areas that had been separated at the time of the node age). These distributions required either prior extinction or long-distance dispersal for achieving the wide distribution at the node and were considered to be less optimal in comparison to the alternative distributions that do not require such events. For example, the wide distributions including Africa, Australia, and Asia at the deep nodes (i.e., node 0 connecting *Cornus* and *Alangium* and node 1 common ancestor of *Alangium* crown group in Fig. 6) were less optimal than the narrower distributions in Europe and Asia, or Asia alone, because Africa and Australia were well separated from each other and from Asia before 90 mya (Scotese, 2004), a time older than the age of nodes 0 and 1 (Fig. 6 and Table 5), while Asia and Europe were connected. The ancestral distributions marked on the tree in Fig. 6 excluded the disjunct distributions for these seven nodes (Fig. 6) and represented the optimal biogeographic history of the genus as inferred from DIVA. These results were largely the same as those from the analysis with the constraint of maximum area equaling 2. This biogeographic pathway required 17 dispersal events to explain the distribution of the species of *Alangium* (both extant species and fossils) (Fig. 6), and suggested an Asian origin for the genus (N1). *Alangium* diverged early into two lineages in Asia in the Late Cretaceous (69–87 mya). One continued to diversify within S.E. Asia and sur-

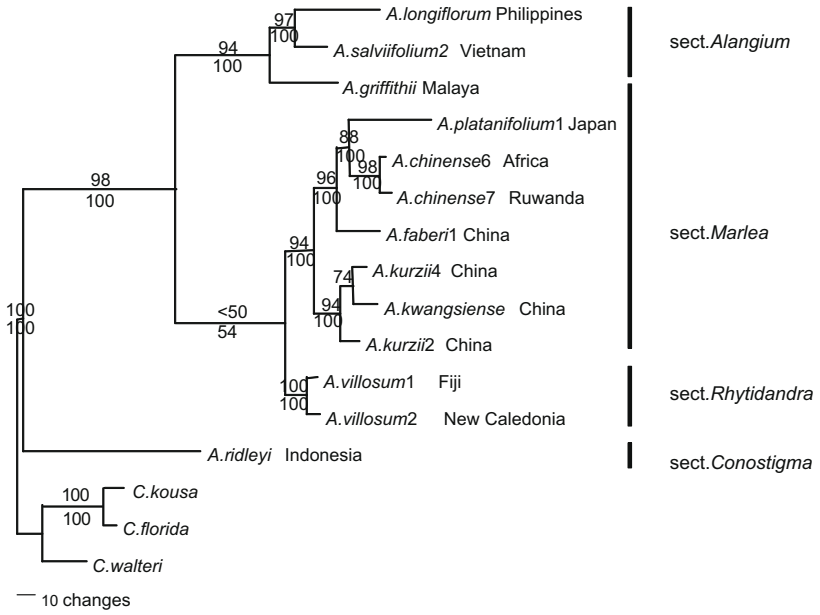


Fig. 3. One of the six most parsimonious trees with 400 steps resulting from analysis of *matK–trn*L-F sequence data. Geographic locations of analyzed samples are given after the taxon name. Numbers above branches are bootstrap values, and numbers below the branches are Bayesian posterior probabilities. *A.* = *Alangium*, *C.* = *Cornus*.

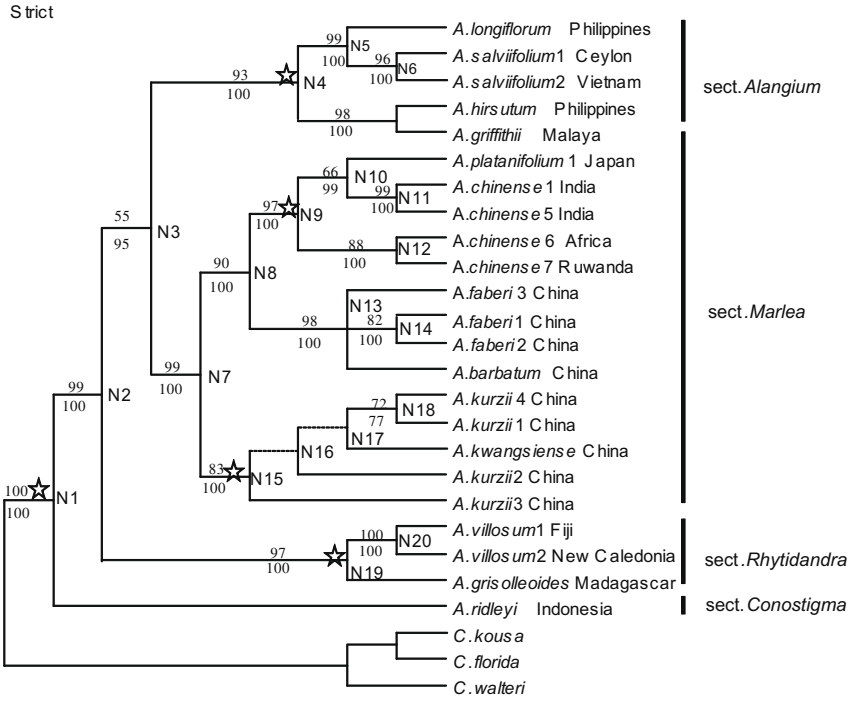


Fig. 4. The strict consensus tree of the 8 shortest trees (each is 1044 steps) resulting from the combined ITS–*matK–trn*L-F data. Geographic locations of the samples are given after the taxon name. Numbers above the branches are bootstrap values, numbers below the branches are Bayesian posterior probabilities. Node numbers N1–N20 correspond to those in Fig. 6. Stars mark the nodes where fossil ages were used as constraints for divergence time analysis using MULTIDIVTIME.

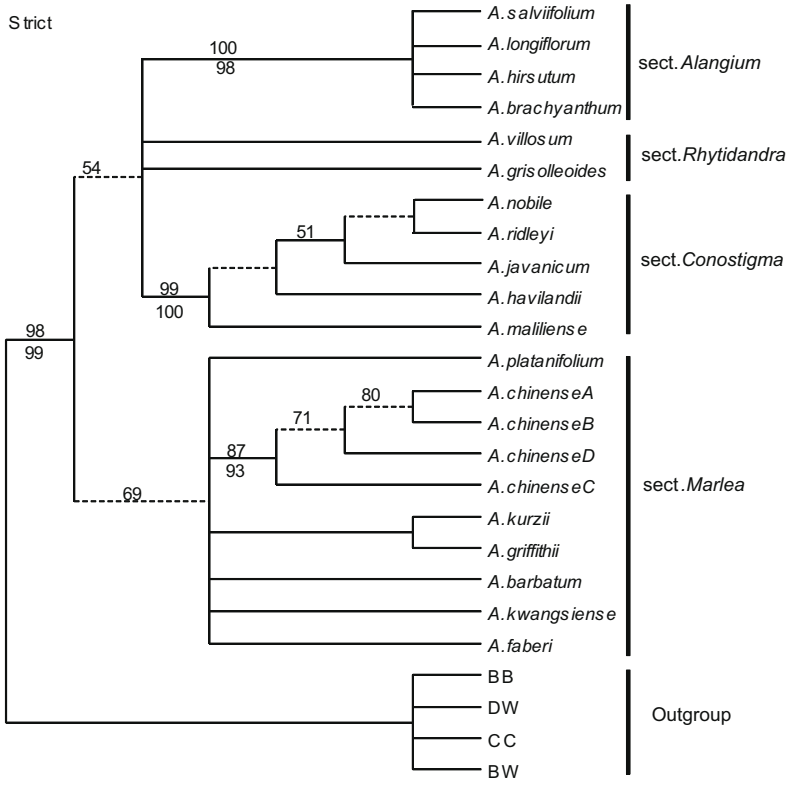


Fig. 5. The strict consensus tree of the 864 shortest trees (each is 292 steps) resulting from morphological data of extant species. Dashed line was the one missed in Bayesian tree. Numbers above branches are bootstrap values, numbers below the branches are Bayesian posterior probabilities. BB = Big-Bracted dogwoods, DW = Dwarf dogwoods, CC = Cornelian Cherries, and BW = Blue- or White-fruited dogwoods. They are the four major lineages in outgroup *Cornus*.

vived as sect. *Conostigma* and dispersed to India during Eocene. The other soon dispersed into Africa (Madagascar, represented by *A. grisolleoides*) (node 2; 66–79 mya; Fig. 6). Isolation of the lineage in Africa (Madagascar) and S.E. Asia led to the divergence and origin of sect. *Rhytidandra*. The African lineage (sect. *Rhytidandra*) later dispersed from Madagascar into Australia near the transition from the Cretaceous to the Tertiary (N19, Fig. 6) and diverged into two species in Australia in the Late Paleocene (node 21). One of the species (*A. villosum*) later spread to S.E. Asia. The remainder of the Asian taxa soon diverged into two lineages (node 3) in the Late Cretaceous; each further diversified and survived as sects. *Alangium* and *Marlea*. Sect. *Alangium* (node 4) subsequently spread into Australia (node 22) during the Oligocene and from Australia into India (node 23) during the Early Miocene, where it diverged into two species (node 23). One of them, *A. salviifolium*, later spread from India to Africa, S.E. Asia, and back to Australia. Multiple movements from S.E. Asia into Europe and North America occurred in sect. *Marlea* (node 7) during the Tertiary (twice into Europe, node 24 in the Paleocene and *A. barghoornianus* before the Oligocene), three times into North America, node 29 in the Early Paleocene, node 25 in the Early Eocene, and *A. barghoornianus* before the Oligocene). All became extinct and were represented only by fossils. In the Early Eocene, sect. *Marlea* also migrated into Africa (node 9; Fig. 6). *Alangium barbatum* (sect. *Marlea*) also moved from S.E. Asia into India sometime after the Early Eocene (node 13). Seven potential vicariance events were inferred from DIVA (nodes 0, 2, 19, 24, 25, 9, and 29; Fig. 6), but two of these (nodes 2 and 19) were likely false. The divergence time at these two nodes was younger than the separation of the distributional areas (Table 5; Scotese, 2004).

4. Discussion

4.1. Phylogenetic relationships and morphology

Phylogenetic analyses of both individual and combined data sets were consistent in supporting the monophyly of the four sections of Bloembergen (1939) with the exception of sect. *Marlea*. *Alangium griffithii* of sect. *Marlea* was resolved as a member of sect. *Alangium* by the molecular data. The finding of sect. *Conostigma* as the earliest diverging lineage agrees with the view of Eyde (1968). Eyde (1968) considered a number of morphological characters in the section, such as the pluricarpellate gynoecia, clavate style, and separate bundles in floral vasculature, to be primitive. Eyde (1968) also proposed that sect. *Alangium* was an advanced group derived from sect. *Conostigma* based on characters he considered to be advanced, such as the single carpel, united stigma, papery endocarp, and increased number of pores in the pollen. Our phylogenetic analyses indeed supported the evolution of sect. *Alangium* after sect. *Conostigma*.

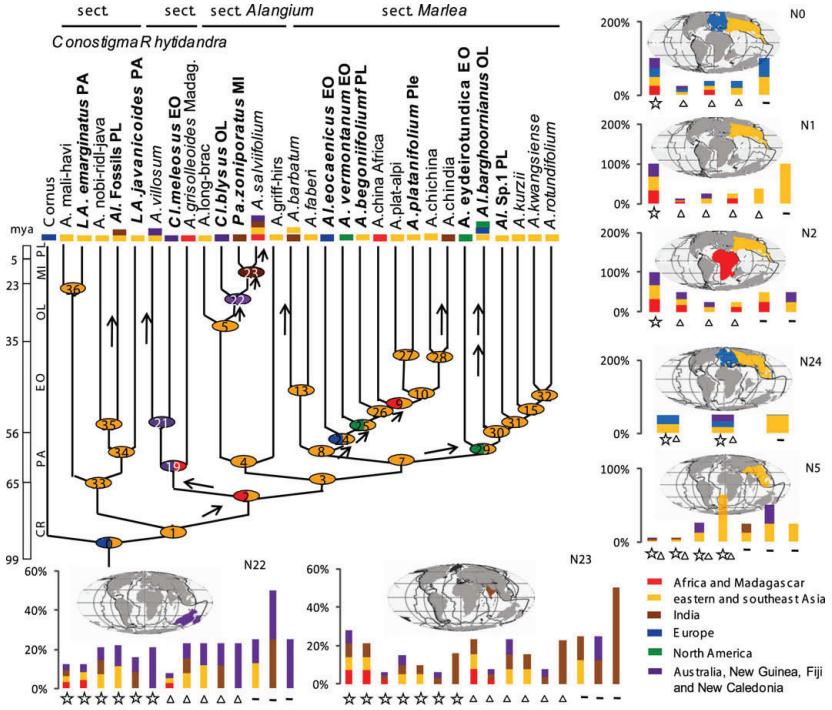


Fig. 6. Result of ancestral area reconstruction using DIVA. Nodes numbered 1–19 correspond to those in Fig. 4 and Table 5. Taxa in boldface italic are fossils. The ages of fossil representatives are indicated as CR = Cretaceous; PA = Paleocene; EO = Eocene; OL = Oligocene; MI = Miocene; PL = Pliocene; Ple = Pleistocene. Madag. = Madagascar, and Afr. = mainland Africa. The age of node 0 was set from divergence time analysis of Cornales (Xiang et al., unpublished). Optimal reconstruction(s) of ancestral areas by DIVA at each node are marked on the phylogenetic tree. For nodes where more than one ancestral distribution was inferred, the frequencies of the alternative solutions are shown in the column chart. Columns marked by an asterisk are results from analyses under no maximum area constraint and under constraints of maximum areas of 4 and 5 (38 equally parsimonious optimizations, results from these three analyses were the same). Columns marked by a triangle are result from analyses under constraint of maximum area set to 3 (208 equally parsimonious optimizations), while columns marked by a short dash are results from analysis under constraint of maximum area equaling to 2 (8 equally parsimonious optimizations). The most likely distribution at each of the variable nodes is also marked on the Paleogeographic map of the geological time corresponding to the age of the node. Some clades in Fig. 4 were reduced to a single branch in this figure when they contain taxa from the same geographic area. A. mali-havi = *A. maliliense* + *A. havilandii*; A.nobi-ridl-java = *A. nobile* + *A. ridleyi* + *A. javanicum*; A.long-brac = *A. longiflorum* + *A. brachyanthum*; A.griff-hirs = *A. griffithii* + *A. hirsutum*; A.plat-alpi = *A. platanifolium* + *A. alpinum*; A.chia africa = *A. chinense* from Africa; A.chindia = *A. chinense* from India, and A. chichina = *A. chinensis* from China. A. eydeirotundica = *A.eydei* + *A. rotundicarpum. Lanagiopollis* fossils include several fossils of *Alangium* found in the same location with similar ages. *A* = *Alangium*; *Al* = *Alangiopollis*; *Cl* = *Clavastephanocolporites*; *La* = *Lanagiopolli*; and *Pa* = *Palaeoalangiumpollis.* The branch of A. eydeirotundica EO also includes the wood fossil *Alangium oregonensis* Scott and Wheeler (Wheeler and Manchester, 2002) from the same locality and Stratigraphic level as the fruit fossil A. eydeirotundica.

Eyde (1968) further proposed that sects. *Marlea* and *Rhytidandra* shared a recent common ancestor based on the presence of features he considered primitive (e.g., bilocular ovaries in sect. *Marlea* and the two long-branched stigmas in sect. *Rhytidandra*). Although the cpDNA data suggested a sister group relationship for sects. *Marlea* and *Rhytidandra* (Fig. 3), the support values were relatively low (54% posterior possibility from Bayesian analysis and <50% bootstrap value). Analyses of the total molecular data did not support this relationship. The estimation of the divergence time, however, did suggest that these two sections evolved at almost the same time, suggesting that the divergence of sects. *Alangium* (node 4), *Marlea* (node 7), and *Rhytidandra* (node19) occurred relatively rapidly and the two sections shared a relatively recent common ancestor, although not the most recent (Fig. 6).

It is noteworthy that the exercise of character mapping on the molecular phylogeny using MacClade, as well as examining the autoapomorphies on the morphological tree resulting from analysis using total molecular tree as the constraint, largely agreed with Eyde's hypothesis on the evolution of morphological characters, e.g., increase in stamen number, fusion of stigmas, clavate (thickened toward the apex) to cylindrical style, hard to papery endocarp (data not shown).

The sectional relationships supported by the molecular data did not agree with Morley (1982), who considered sects. *Conostigma* and *Marlea* to be of equal age. The molecular phylogeny suggested that sect. *Conostigma* was older than sect. *Marlea* (Fig. 6). Our divergence time estimates, however, did support Morley's (1982) hypothesis regarding the antiquity of these sections. The divergence times indicate that the origins of these sections date back 66–87 mya (nodes 1 and 3; Fig. 6).

Alangium griffithii, once placed in sect. *Marlea* (Bloembergen, 1939), was suggested to be an early diverging species in sect. *Alangium* by the molecular data (Figs. 2–4). Analysis of the morphological data, however, suggested that the species was a member of sect. *Marlea*, as traditionally placed, and sister to *A. kurzii*. Although *A. griffithii* shared many morphological characters with sect. *Marlea*, including papery leaves, equal number of stamens and petals, four-lobed capitate stigma, short radicles, thin walled pith paren-

Table 5
Divergence times (mya—million years ago) estimated using Bayesian method of Thorne et al. (1998) and Thorne and Kishino (2002).

Node	Estimated divergence time (million years ago)	95% credibility intervals	Estimated ancestral area
N1	77.66085 (⩾65–55)	(69.97691, 87.55942)	Southeast Asia
N2	73.44130	(67.65111, 80.98023)	Africa and Southeast Asia
N3	72.06908	(66.36394, 79.46110)	Southeast Asia
N4	64.59074 (⩾33.5)	(57.50565, 72.48408)	Southeast Asia
N5	28.27841	(16.67519, 38.16155)	Southeast Asia
N6	22.49517	(11.77103, 32.26607)	Southeast Asia
N7	64.77656	(60.65069, 70.34095)	Southeast Asia
N8	61.72840	(57.86410, 66.74154)	Southeast Asia
N9	57.06686 (⩾55)	(55.07235, 61.50241)	Africa and Southeast Asia
N10	47.47355	(33.20003, 57.22837)	Southeast Asia
N11	16.29615	(1.25891, 36.89924)	India
N12	38.20134	(16.41354, 51.45844)	Africa
N13	47.44033	(34.84551, 56.53252)	Southeast Asia
N14	29.40123	(11.94391, 45.10801)	Southeast Asia
N15	57.50675 (⩾55)	(55.08772, 62.55266)	Southeast Asia and North America
N16	51.84312	(40.96877, 59.41986)	Southeast Asia
N17	47.65872	(33.99518, 56.96408)	Southeast Asia
N18	35.28246	(13.58891, 51.00474)	Southeast Asia
N19	67.72115 (⩾55)	(62.32220, 74.91118)	Africa and Australia
N20	57.57064	(55.07613, 62.96427)	Australia

Node numbers correspond to those on Figs. 5 and 6. Ages in parentheses are fossil ages used as constraints.

chyma, smooth endosperm, thick endocarp wall, and presence of pits on the endocarp, as emphasized by some authors, most of these feature were plesiomorphies also found in other sections. Only the papery leaf and four-lobed capitate stigma were uniquely shared by sect. *Marlea* and *A. griffithii*. Based on the molecular phylogeny, these two similarities must evolve independently.

The placement of the Madagascan *Alangium grisolleoides* in sect. *Rhytidandra* by Govindarajalu (1979) was strongly supported by our molecular analyses (Figs. 2 and 4), and further supported by several morphological synapmorphies, including the two palisade layers, unique pigment in the leaves, fruit with cuneate base and acute apex, narrow and faint costae of endoapertures, and sexine/nexine ratio (Supplementary material 1). Capuron (1962) and Eyde (1968) suggested *A. grisolleoides* to be a member of sect. *Conostigma* based on similarities in fruit structure (i.e., massive unilocular endocarp with many fine longitudinal ridges on the surface). However, these characters were also shared with *A. villosum* in sect. *Rhytidandra*. The molecular based phylogeny obtained in our study did not support the placement of *A. grisolleoides* in sect. *Conostigma* (Figs. 2 and 4).

The multiple samples of *Alangium chinense* and *A. platanifolium* from different areas were united in a single clade, while those of *A. chinense* were grouped into two subclades that diverged in the Early Eocene (55.1–61.5 mya; Figs 2, 4 and 6; Table 5, node 9 in Fig. 6). One subclade included samples from Africa and the other included samples from China and India. Our findings supported the recognition of at least two species, one in Africa and the other in S.E. Asia and India, within the traditional widely distributed *A. chinensis*. The China–India clade of *A. chinense* was closer to *A. platanifolium* than it was to the African clade (node 10, Figs. 4 and 6). The two subclades diverged in the Mid Eocene (55.1 to 61.5 mya; node 9, Fig. 6). This finding that *A. chinense* was not monophyletic was not surprising, given the wide geographic distribution of the species (Table 1) and the four distinct pollen types found within the species, with each endemic to a separate geographic area (Reitsma, 1970) while no other species in the genus had more than one type of pollen (Reitsma, 1970). The phylogenetic relationships of *A. chinense* outside of China, India, and Africa could not be evaluated due to lack of samples.

4.2. Biogeographic and evolutionary history- the importance of long-distance dispersal

The optimal biogeographic pathway of *Alangium* derived from DIVA implied 17 dispersal events (Fig. 6). Based on the divergence time and paleontological evidence, many of these dispersal events occurred between distant areas separated by oceans, thus, had to be long distance. For example, by the Late Cretaceous, India had been separated from all other landmasses and Australia was far from Laurasia and western Gondwanaland (Scotese, 2004). Movements involving these areas (e.g., events between nodes 2 and 19, 21 and *A. villosum*, 5 and 22, and 22 and 23 occurred in sections *Rhytidandra* and *Alangium*) therefore must have involved long-distance dispersal. The early dispersal from S. E. Asia to Africa/Madagascar in the Late Cretaceous (from node 1 to node 2; Fig. 6) could have been achieved via island hopping along the margin of the Tethys Seaway (Scotese, 2004; Tiffney, 1985) or long-distance dispersal from S.E. Asia into Madagascar directly. While the Tethys Seaway route required extinction of the lineage in mainland Africa (the African occurrence of the lineage is only known from Madagascar where *A. grisolleoides* is distributed), the long-distance dispersal scenario seemed more likely. Although the more recent dispersals of *A. salviifolium* from India into China and Africa in the later Tertiary were easy to explain because India had united with Asia by the Early Eocene (Scotese, 2004) and Africa was connected to Asia via Arabia in the Mid Miocene (~17 mya) (Raven and Axelrod, 1974), the dispersal of the species from India to Australia in the Miocene (Fig. 6) had to involve long-distance dispersal. Species of *Alangium* produce fleshy, drupaceous fruits that are dispersed by birds and mammals (Corlett, 1998; Au et al., 2006). The stony fruit endocarp containing copious oily endosperm also permits dispersal of seeds by water.

The majority of dispersal events in sect. *Marlea*, however, did not have to invoke long distance. For example, dispersal from S.E. Asia into Europe in the Paleocene or Eocene (nodes 8 to 24, node 30 to *A. barghoornianus* (Oligocene fossil; Fig. 6)) could be achieved by gradual migration along the island chain of the Tethys seaway (Tiffney, 1985; Heaney, 2004). During the Late Cretaceous and Paleocene, the Turgai Strait isolated Asia and Europe, but Africa, southern Europe, southwestern and southeastern Asia were linked

via a land bridge along the closing Tethys seaway (a circumequatorial current separating the ancient Laurasian and Gondwana supercontinents) that was also proposed for dispersal of mammals into Africa in the Paleocene (Sigé et al., 1990; Hooker et al., 1999). Dispersals from S.E. Asia into North America (nodes 24 to 25, node 7 to 29, and node 30 to *A. barghoornianus* (Oligocene fossil Fig. 6) in the early Tertiary (Paleocene to Eocene) could have been achieved by migrations across the Bering land bridge (BLB) or the North Atlantic land bridge (NALB). Migration from S.E. Asia to North America via the NALB required extinction in Europe, but the high latitudes and long winters of the BLB in the early Tertiary could have prevented the crossing of *Alangium*. Section *Marlea* is the only section in the genus with cold-tolerant species (e.g., *A. platanifolium* with a distribution to Siberia; Fedina et al., 2002). Thus, the high latitude of the BLB might not have been a problem for members of the *Marlea* lineage to cross. Similarly, movements of *A. barbatum* and *A. chinensis* from S.E. Asia into India after the Early Eocene (node 13 to *A. barbatum* and node 28 to *A. chinensis*; Fig. 6) did not have to involve long-distance dispersal because India had connected to Asia by the Early Eocene. The spread of sect. *Marlea* into Africa from S.E. Asia (node 26 to node 9) in the Early Eocene, however, may have involved long-distance dispersal or island hopping because the two areas were distantly separated at the time (Scotese, 2004).

Our results suggested that long-distance dispersal played a very important role in the biogeographic history, evolution, and speciation of *Alangium*. The many long-distance dispersal events across ocean barriers resulted in immediate isolation, allowing lineages to diverge. We did not detect vicariance events at any nodes connecting the living taxa (i.e., node time approximates the time of separation of areas of node distribution, as predicted by vicariance). All nodes connecting disjunct extant lineages with more than one area in distribution had ages younger than the isolation of the distributional areas, according to earth's plate tectonic history (Heaney, 2004; Sanmaritin and Ronquist, 2004; Scotese, 2004; Lomolino et al., 2006). For nodes connecting disjunct living and fossil lineages with distributions in connected areas (e.g., Asia and Europe, Asia and North America in the early Tertiary; nodes 24, 25, and 29; Fig. 6) in sect. *Marlea*, the divergence of the species now known only from fossils and their sisters could be the results of vicariance events due to extinction.

Evidence from fossils also suggested that considerable extinction occurred during the evolutionary history of *Alangium* and the current distribution of the genus was reduced from its once wider distribution in the past. Sect. *Conostigma* is now confined to southern regions of S.E. Asia, but fossil pollen indicates that the lineage was present in the Paleocene and Eocene of China (*La. javanicoides*) (Table 3) and an Oligocene wood fossil from India suggested its possible migration to that area (sectional affinity of the wood fossil was uncertain, possibly sect. *Conostigma* or sect. *Marlea* based on perforations and length of vessel elements; Awasthi, 1968; Wheeler and Manchester, 2002). During the Eocene–Oligocene transition, cooling of the global climatic caused extinction of many tropical elements and favored the spread of deciduous plants (Wolfe, 1975; Tiffney, 1985; Miller et al., 1987; Miller, 1992; Tiffney and Manchester (2001); Zachos et al., 2001). It is possible that the extinction of sect. *Conostigma* in China and India was due to a failure to adapt to the colder climate during that period. All species of the section are evergreen tropical trees, suggesting that the lineage was not adapted to the cool climates during the long evolutionary history. Section *Marlea* is now absent from Europe and North America, but fossil evidence indicated its presence in those areas in the early Tertiary (Figs. 1 and 6). The extinction of this lineage in Europe and North America was likely the consequence of climatic deterioration at the Eocene/Oligocene transition and later climatic cooling during the late Tertiary (Prothero and Berggren, 1992). Consistent with this hypothesis is the fact that most fossils found in Europe and North America were in the Eocene to Miocene (Tables 3 and 4). Unlike sect. *Rhytidandra*, however, sect. *Marlea* evolved a few deciduous species, such as *A. faberi*, *A. chinensis*, *A. kurzii*, and *A. platanifolium*, that were adapted to the cooler climate in the Early Oligocene (Fig. 6).

Our analysis largely supports Morley's (1982) hypothesis that the genus evolved in S.E. Asia and diverged relatively early into two separate evolutionary clades, one tropical represented by sect. *Conostigma*, and the other subtropical to temperate, represented by sect. *Marlea*. Our work does not support the North American origin hypothesis of Krutzsch (1989). Reitsma (1970)'s hypothesis of an Indian origin of *A. chinensis* (Fig. 1) cannot be evaluated in this study due to limited sampling of the species, although our results suggested that *A. chinensis*, as traditionally circumscribed, is polyphyletic. Our analyses did not support the hypothesis of Martin et al. (1996; Fig. 1) that sect. *Rhytidandra* evolved in Australia and subsequently spread into S.E. Asia. Instead, it suggested an African/Madagascan origin of the section in the Late Cretaceous and later dispersal to Australia and S.E. Asia (nodes 2 and 19; Fig. 6).

The spread of thermophilic taxa in the northern hemisphere has been proposed to involve primarily migrations across the North Atlantic land bridge (Tiffney, 1985) in the early Tertiary. Phylogenetic evidence from other taxa including Malpighiaceae (Davis et al., 2002), Lauraceae (Chanderbali et al., 2001), Melastomataceae (Renner and Meyer, 2001), Leguminosae (Schrire and Lavin, 2005), and *Cornus* (Xiang et al., 2005), mostly agree with this hypothesis, but migration of such taxa across the Bering land bridge in the middle Tertiary was also suggested for some taxa, e.g., Melastomataceae (Renner and Meyer, 2001). The terrestrial route on the margins of the Tethys seaway was also proposed as a possible migration route for the spread of boreotropical flora (Wolfe, 1975) around the tropical regions in the early Tertiary (Tiffney, 1985; Xiang et al., 2006). The inferred biogeographic history of *Alangium* was unable to exclude the NALB or BLB for the early Tertiary migration between Eurasia and North America, but supports the role of the Tethys Seaway island chains (connecting S.E. Asia, southern Europe, and Africa) for plant exchange between the two super-continents in the tropics in the early Tertiary. Furthermore, our study indicated long-distance dispersal has played a major role in the diversification and range expansion of the genus, suggesting that long-distance dispersal is likely important in the evolution of plants in the tropics.

Acknowledgments

The authors thank the curators of herbaria at Missouri Botanical Garden, Leiden, Harvard University, Royal Botanical Garden, Kew, and the Smithsonian Institution for loaning their specimens. Special thanks to C.X. Fu, X. Chen, Y. Liu, D.E. Boufford, C.Y. Li, and P. Raven for assistance in obtaining plant materials, E.A. Wheeler for discussion on wood anatomy characters, D.E. Boufford for reading the mansucript. We thank the Deep Time Research Coordination Network supported by the National Science Foundation (NSF) funded to D.E. Soltis (DEB-0090283) and the Phytogeography of the Northern Hemisphere working group supported by NSF via NESCent for travel support to their workshops and symposia. This study benefited from discussion with colleagues in the deep time network and the phytogeography working group. We also thank the reviewers for valuable comments to improve this manuscript. This research was also supported by the National Science Foundation (DEB-0444125) funded to Xiang.

Appendix A

Voucher information and GenBank accession numbers for taxa used in this study. A dash indicates the region was not sampled.

Voucher specimens are deposited in the following herbaria: MO = Missouri Botanical Garden; L = Leiden.

Taxon—GenBank accessions: *ITS*, *trn*L-F, *matK*; Source; Voucher specimen (DNA samples from herbarium specimens except for those marked by underline which were from silica-dried leaves).

A.1. Section Alangium

Alangium hirsutum Bloembergen, — —, FJ610034, —; L: K. Sidlyase, 1196, Philippines *Alangium longiflorum* Merrill — FJ610000, FJ610031, FJ644638; MO: Blas F. Hernaez, 635, Philippines *Alangium salviifolium* (L.) Wangerin, Acc#1 — FJ610001, FJ610032, –; MO: B.&K. Bremer, 1805, Ceylon *Alangium salviifolium* (L.) Wangerin, Acc#2 — FJ610002, FJ610033, FJ644639; MO: Mai Van Sinh, MVX100, Vietnam *Alangium salviifolium* (L.) Wangerin, Acc #3 — FJ610003, —; —; MO: S.H. Sohmer, 8647, Ceylon *Alangium salviifolium* (L.) Wangerin, Acc#4 — FJ610004, — ,—; MO: Suzanne Ripley, 154, Ceylon.

A.2. Section Marlea

Alangium barbatum (C.B. Clarke) Harms, Acc#1 — FJ610025, FJ610048, —; Xiang, 02256, China *Alangium chinense* (Loureiro) Harms, Acc#1 — FJ610008, FJ610036, —; MO: Walter Koelz, 4731, India *Alangium chinense* (Loureiro) Harms, Acc#2 — FJ610009, —; Xiang, 04c72, China *Alangium chinense* (Loureiro) Harms, Acc#3 — FJ610010, –; Xiang, 02257, China *Alangium chinense* (Loureiro) Harms, Acc#4 — FJ610011, –; Xiang 04c27, China *Alangium chinense* (Loureiro) Harms, Acc#5 — FJ610012, FJ610037, —; MO: S.K. Jain and R.C. Bharadwaja, India *Alangium chinense* (Loureiro) Harms, Acc#6 — FJ610013, FJ610038, FJ644641; MO: M.G. Bashonga, 567, Africa *Alangium chinense* (Loureiro) Harms, Acc#7 — —, FJ610039, FJ644642; MO: G.Bouxin, Ruwanda *Alangium chinense* (Loureiro) Harms, Acc#8 — FJ610014, –, –; Xiang, c4044, China *Alangium chinense* (Loureiro) Harms, Acc#11 FJ610015, –, –; Xiang, c4c24, China *Alangium chinense* (Loureiro) Harms, Acc#12 — FJ610016, –, –; Xiang, 04c31, China *Alangium chinense* (Loureiro) Harms, Acc#13 — FJ610017, –, –; MO: D.E. Boufford and B. Bartholomew, 23970, China.

Alangium faberi Oliv. Acc#1 — FJ610023, FJ610046, FJ644645; MO: Wen He Qun, W068, China.

Alangium faberi Oliv. Acc#2 — FJ610024, FJ610047, –; Xiang, 02111, China.

Alangium faberi Oliv. Acc#3 — FJ610022, FJ610045, –; Xiang, 02-82, China.

Alangium griffithii (Clarke) Harms — FJ610026, FJ610049, FJ644646; L: M. Jacobs, 8300, Malaya.

Alangium kurzii Craib, Acc#1 — FJ610018, FJ610040, –; Xiang,0272,China.

Alangium kurzii Craib, Acc#2 — –, FJ610041, FJ644643; MO: J.B.Zuo, 0543, China.

Alangium kurzii Craib, Acc#3 — FJ610019, FJ610042, –; Xiang, 02242, China.

Alangium kurzii Craib, Acc#4 — FJ610020, FJ610043, FJ644650; MO: Shukuai Lai, Hanrong Shan, 3613, China *Alangium kwangsiense* Melch. In Notizbl— FJ610021, FJ610044, FJ644644; MO: Zhijian Feng, 83705, China *Alangium platanifolium* (Siebold and Zuccarini) Harms, Acc#1 — FJ610005, FJ610035, FJ644640; MO: S. Tsugaru and T. Takahashi, 16393, Japan *Alangium platanifolium* (Siebold and Zuccarini) Harms, Acc#5 — FJ610006, –, –; Xiang, 02107, China *Alangium platanifolium* (Siebold and Zuccarini) Harms, Acc#6 — FJ610007, –, –; Xiang, 02118, China.

A.3. Section Rhytidandra

Alangium grisolleoides Capuron — FJ610029, FJ610052, –; MO: L.M. Randrianjanaka, 596, Madagascar *Alangium villosum* (Blume) Wangerin, Acc#1 FJ610027, FJ610050, FJ644647; MO: Bloemb; Dtto. Degener, 15321, Fiji *Alangium villosum* (Blume) Wangerin, Acc#2 — FJ610028, FJ610051, FJ644648; MO: Gordon Mc pherson, 6231, New Caledonia.

A.4. Section Conostigma

Alangium ridleyi King — FJ610030, FJ610053; FJ644649; MO: Ambriansyah and Z. Arigin, 359, Indonesia.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.01.017.

References

Au, A.Y.Y., Corlett, R.T., Hau, B.C.H., 2006. Seed rain into upland plant communities in Hong Kong, China. Plant Ecology 186 (1), 13–22.

Awasthi, N., 1968. A new fossil wood belonging to the family Alangiaceae from the Tertiary of South India. The palaeobotanist 17, 322–325.

Baumann, P., Oesterle, H., Suminta Wibisono, 1972. Stratigraphic correlation in the Tertiary of Java and Sumatra. In: Proceedings of the Indonesia Petroleum. Association 1st Annual Convention, pp. 31–24.

Bloembergen, S., 1939. A revision of the genus *Alangium*. Archipel Drukkerij, Buitenzorg.

Capuron, R., 1962. Contributions. A letude de la flore forestiere de Madagascar. VII. Presence a Madagascar du genre Alangium et description dune espece nouvelle. Adansonia 2, 282–284.

Chanderbali, A., van der Werff, S.H., Renner, S.S., 2001. Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. Annals of the Missouri Botanical Garden 88, 104–134.

Chase, M.W., Soltis, D.E., Olmstead, G., et al., 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene rbcL. Annals of the Missouri Botanical Garden 80, 528–580.

Cronquist, A., 1981. An Integrated System of Classification of Flowering Plants. Columbia University Press, New York. pp. 664–665.

Corlett, R.T., 1998. Frugivory and seed dispersal by vertebrates in the Oriental (Indomalayan) Region. Biological Reviews 73 (4), 413–448.

Davis, C.C., Bell, C.D., Mathews, S., Donoghue, M.J., 2002. Laurasian migration explains Gondwanan disjuncts: evidence from Malpighiaceae. Proceedings of the National Academy of Sciences, USA 99, 6833–6837.

Eyde, R.H., 1968. Flowers, fruits, and phylogeny of Alangiaceae. Journal of the Arnold Arboretum 49 (2), 168–192.

Eyde, R.H., Bartlett, A., Barghoorn, E.S., 1969. Fossil record of *Alangium*. Bulletin of the Torrey Botanical Club 96 (3), 288–314.

Fan, C., Xiang, Q.Y., 2001. Phylogenetic relationships within Cornus L. (Cornaceae) based on 26S rDNA sequences. American Journal of Botany 88, 1131–1138.

Fan, C., Xiang, Q.Y., 2003. Phylogenetic analyses of Cornales based on 26S rRNA and combined 26S rDNA *matK*–rbcL sequence data. American Journal of Botany 90, 1357–1372.

Farabee, M.J., Canright, J.E., 1986. Stratigraphic palynology of the lower part of the Lance Formation (Maestrichtian) of Wyoming. Palaeontographica Abt. B 199 (1–3), 1–89.

Fedina, L.A., Pavlova, N.S., Kudryavtseva, E.P., Kovalev, V.A., 2002. *Alangium platanifolium*, a species of Alangiaceae, the family new for the flora of Russia. Botanicheskii zhurnal 87 (12), 126–129.

Felsenstein, J., 1984. Distance methods for inferring phylogenies: A justification. Evolution (38), 16–24.

Fui, H.K., 1978. Stratigraphic framework for oil exploration in Sarawak. Geological Society of Malaysia Bulletin 10, 1–14.

Govindarajalu, E., 1979. The comparative morphology of the Alangiaceae: 6. On the foliar anatomy of two new species of *Alangium*. Proceedings of the Indian Academy of Sciences 88 (4), 283–291 (Illus. 4).

Harms, H., 1897. Die Gattungen der Cornaceae. Ber. Deutsche Bot. Ges. 15, 21–29.

Hartono, H.M.S., 1969. Globigerina marls and their planktonic foraminifera from the Eocene of Nanggulan, Central Java. Cushman Found. Foraminiferal Res. Contrib. 20, 152–159.

He, M., Bai, J., He, X., 1996. A new species of Alangiaceae from Sichuan. Journal of Sichuan University 33 (6), 739–740 (illus.).

Heaney, L.R., 2004. Conservation biogeography in oceanic archipelagoes. In: Lomolino, M.V., Heaney, L.R. (Eds.), Frontiers of Biogeography: New Directions in the Geography of Nature. Cambridge University Press, Cambridge. pp. 345–360.

Hooker, J., Russell, D.E., Phélizon, A., 1999. A new family of Plesiadapiformes (Mammalia) from the Old World Lower Paleogene. Palaeontology 42 (3), 377–407.

Huelsenbeck, J.P., Ronquist, F., 2003. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Jain, A., Katewa, S.S., Galav, P.K., Sharma, P., 2005. Medicinal plant diversity of Sitamata wildlife sanctuary, Rajasthan, India. Journal of Ethnopharmacology 102, 143–157.

Keast, A., 1996. Pacific biogeography: Patterns and Process In The origin and evolution of Pacific island biotas. In: Keast, Allen, Miller, ScottE. (Eds.), New Guinea to Eastern Polynesia: Patterns and Processes. SPB Academic Publishing bv.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29, 170–179.

Kishino, H., Thorne, J.L., Bruno, W.J., 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Molecular Biology and Evolution 18, 352–361.

Kroenke, L.M., 1996. Plate tectonic development of the western and southwestern Pacific Mesozoic to the present. In: Keast, A., Scott, E., Miller, S.P.B. (Eds.), The Origin and Evolution of Pacific Island Biotas, New Guinea to Eastern Polynesia: Patterns and Processes. Academic Publishing bv.

Krutzsch, W., 1962. Stratigraphisch bzw. Botanisch wichtige neue Sporen und Pollen formen aus dem deutschen Tertiär. Geologie 11, 265–307.

Krutzsch, W., 1989. Paleogeography and historical phytogeography (paleochorology) in the Neophyticum. Plant Systematics and Evolution 162, 5–61.

Leichti, P., Rue, F.W., Haile, N.S., 1960. The Geology of Sarawak, Brunei and the western part of North Borneo. Bulletin of the British Borneo Geological Survey Borneo. 3.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic Biology 50 (6), 913–925.

Liu, G., 1986. A Late Tertiary palynological assemblage from the Yaoshan Formation of Shanwang, Linju County, Shandong. Acta Palaeobotanica et Palynologica 1 (1), 65–84.

Lomolino, M., Riddle, B.R., Brown, J.H., 2006. Biogeography, third ed. Sinauer Associates. pp. 227–274.

Maddison, D.R., Maddison, W.P., 2000. MacClade 4: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA.

Mai, D.H., 1976. Fossile Früchte und Samen aus dem Mitteleozän des Geiseltales. Abhandlungen des Zentralen Geologischen Insitutes Paläontologishe Abhandlungen. pp. 93–149.

Manchester, S.R., 1994. Fruits and seeds of the Middle Eocene Nut Beds flora, Clarno Formation, Oregon. Palaeontographica Americana 58, 1–205.

Manchester, S.R., 1999. Biogeographical relationships of North American Tertiary floras. Annals of the Missouri Botanical Garden 86, 472–522.

Martin, A.H., Macphail, K.M., Partridge, D.A., 1996. Tertiary *Alangium* (Alangiaceae) in eastern Australia: evidence from pollen. Review of Palaeobotany and Palynology 94, 111–122.

Metcalfe, I., 1999. Gondwana Dispersion and Asian Accretion. A.A. Balkema, Rotterdam, The Netherlands.

Miki, S., 1956. Endocarp remains of Alangiaceae, Cornaceae and Nyssaceae in Japan. J. Inst. Polytech. Osaka City Univ. 7, 275–297.

Miller, K.G., 1992. Middle Eocene to Oligocene stable isotopes, climate, and deep-water history: the terminal Eocene event? In: Prothero, D.R., Berggren, W.A. (Eds.), Eocene–Oligocene Climatic and Biotic Evolution. Princeton University Press, Princeton, NJ, pp. 160–177.

Miller, K.G., Fairbanks, R.G., Mountain, G.S., 1987. Tertiary oxygen isotope synthesis, sea level history, and continental margin erosion. Paleoceanography 2, 1–19.

Morley, R.J., 1982. Fossil pollen attributable to *Alangium* Lamarck (Alangiaceae) from the tertiary of Malaysia. Review of Palaeobotany and Palynology 36 (1/2), 65–94 (illus. maps).

Muller, J., 1964. Palynological contributions to the history of Tertiary vegetation in NW, Borneo. In: Abstr. X Interantional Botanical Congress, Edinbergh, p. 271.

Phadtare, N.R., Kulkarni, A.R. 1983. Palynological assemblage of lignitic exposure of Ratnagiri Distt. In: Proc. XthInd. Coll. Micropaleontol Stratigr, pp. 515–532.

Phadtare, N.R., Thakur, B., 1990. Fossil pollen of *Alangium* from the Eocene lignite of Gujarat, India, with comments on its stratigraphic antiquity. Review of Palaeobotany and Palynology 63, 281–297.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Prothero, D.R., Berggren, W.A., 1992. Eocene–Oligocene Climatic and Biotic Evolution. Princeton University Press, Princeton, NJ.

Rambaut, A., Drummond, A., 2004. TRACER v.1.1 MCMC Trace Analysis Tool. University of Oxford.

Raven, P.H., Axelrod, D.I., 1974. Angiosperm biogeography and past continental movements. Annals of the Missouri Botanical Garden 61, 539–673.

Ronquist, F., 1996. DIVA version 1.1. Uppsala Univeristy, Uppsala (ftp.uu.se or ftp.systbot.uu.se).

Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Systematic Biology 46, 195–203.

Reid, E.M., Chandler, M.E.J., 1933. The London Clay Flora. British Museum (Natural History), London.

Reitsma, T.J., 1970. Pollen morphology of the Alangiaceae. Review of Palaeobotany and Palynology 10, 249–332.

Renner, A., Meyer, K., 2001. Melastomeae come full circle: biogeographic reconstruction and molecular clock dating. Evolution 55, 1315–1324.

Sang, T., Crawford, D.J., Stuessy, T.F., 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). American Journal of Botany 84, 1120–1136.

Sanmaritin, I., Ronquist, F., 2004. Southern Hemisphere biogeography inferred by event-based models: Plants vs Animal patterns. Systematic Biology 53, 216–243.

Scotese, C.R., 2004. Cenozoic and Mesozoic paleogeography: changing terrestrial biogeographic pathways. In: Lomolino, M.V., Heaney, L.R. (Eds.), Frontiers of Biogeography—New Directions in the Geography of Nature. Sinauer Associates, Inc., Sunderland, MA, pp. 9–26.

Sigé, B., Jaeger, J.J., Sudre, J., Vianey-Liaud, M., 1990. Altiatlasius koulchii n. Gen. et sp., primate omomyidé du Paléocène supérieur du Maroc, et les origines des euprimates. Palaeontographica Abt. A 214, 31–56.

Srivastava, S.K., 1972. Some spores and pollen from the Paleocene Oak Hill member of the Naheola Formation, Alabama (USA). Review of Palaeobotany and Palynology 14, 217–285.

Stanley, S.M., Steven, M., 2005. Earth System Biology. W.H. Freeman, New York.

Stone, B.C., Kochummen, K.M., 1975. A new *Alangium* (Alangiaceae) from Sarawak. Blumea 22 (2), 219–222.

Soltis, D.E., Soltis, P.S., Endress, P.K., Chase, M.W. 2005. Phylogeny and evolution of angiosperms Sinauer Associates, INC. Publishers Sunderland, Massachusetts.

Song, Z., Wang, W., Huang, F., 2004. Fossil pollen records of extant angiosperm in China. Botanical Review 70 (4), 425–458.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods), version 4.0 b10. Sinauer Associates, Sunderland, MA.

Takhtajan, A., 1980. Outline of the classification of flowering plants (Magnoliophyta). Botanical Review 46, 225–359.

Takhtajan, A., 1997. Diversity and Classification of Flowering Plants. Columbia University Press, New York.

Thorne, J.L., Kishino, H.I., Painter, S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Molecular Biology and Evolution 15, 1647–1657.

Thorne, J.L., Kishino, H.I., 2002. Divergence time and evolutionary rate estimation with multilocus data. Systematic Biology 51, 689–702.

Thorne, R.F., 1992a. Classification and geography of the flowering plants. Botanical Review 58, 225–348.

Thorne, R.F., 1992b. An updated phylogenetic classification of the flowering plants. Aliso 13, 365–389.

Tiffney, H., 1985. The Eocene North Atlantic Land Bridge: its importance in tertiary and modern phytogeography of the Northern hemisphere. Journal of the Arnold Arboretum 66 (2), 243–273.

Tiffney, B.H., Manchester, R.S., 2001. The use of Geological and Paleontoligical evidence in evaluating plant phylogeographic hypotheses in the northern hemisphere tertiary. International Journal of Plant Science 162, S3–17.

Traverse, A., 1955. Pollen analysis of the Brandon Lignite of Vermont. U.S. Bur. Mines, Rep. Invest. 1551, 107.

Traverse, A., 1994. Palynofloral geochronology of the Brandon Lignite of Vermont, USA. Review of Palaeobotany and Palynology 82, 265–297.

Wangerin, W., 1910. Alangiaceae. In: Wngler, A. (Ed.), Das Pflanzenreich, IV, vol. 220b, pp. 1–224.

Wheeler, E.A., Manchester, S.R., 2002. Woods of the Eocene Nut Beds flora, Clarno Formation, Oregon, USA. International Association of Wood Anatomists Journal, Supplement Pages 3, 188.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M., Gelfand, D., Sninsky, J., White, T. (Eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press, San Diego, pp. 315–322.

Whitmore, T.C., 1973. Plate tectonics and some aspects of Pacific plant geography. New Phytologist 72, 1185–1190.

Wolfe, J.A., 1975. Some aspects of plant geography of the Northern Hemisphere during the late Cretaceous and Tertiary. Annals of the Missouri Botanical Garden 62, 264–279.

Xiang, Q.Y., Soltis, D.E., Morgan, D.R., Soltis, P.S., 1993. Phylogenetic relationships of Cornus L. Sensu lato and putative relatives inferred from *rbcL* sequence data. Annals of the Missouri Botanical Garden 80, 723–734.

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

Xiang, Q.Y., Moody, M.L., Soltis, D.E., Fan, C.Z., 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., Manchester, S.R., Thomas, D., Zhang, W.H., Fan, C.Z., 2005. Phylogeny, biogeography, and molecular dating of cornelian cherries (*Cornus*, Cornaceae)—tracking Tertiary plant migration. Evolution 59, 1685–1700.

Xiang, Q.Y., Thomas, D.T., Zhang, W.H., Manchester, S.R., Murrell, Z., 2006. Species level phylogeny of the dogwood genus Cornus (Cornaceae) based on molecular and morphological evidence: implication in taxonomy and Tertiary intercontinental migration. Taxon 55 (1), 9–30.

Yang, Z., 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. 39, 306–314.

Yang, Z., 1997. PAML: a program for package for phylogenetic analysis by maximum likelihood. CABIOS 15, 555–556.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms and aberrations in global climate 65 Ma to present. Science 292, 686–693.

Chapter II

Shoot regeneration of dwarf dogwood (*Cornus canadensis* L.) and morphological characterization of the regenerated plants

Chun-Miao Feng[1], Rongda Qu[2], Li-Li Zhou[1], De-Yu Xie[1], Qiu-Yun (Jenny) Xiang[1]

[1]Department of Plant Biology, North Carolina State University, Raleigh, NC 27695, USA

[2]Department of Crop Science, North Carolina State University, Raleigh, NC 27695, USA

Published in Plant Cell Tissue Organ Culture, 2009, 97: 27-27

Chunmiao did all of the work reported in this paper, Dr. Li-Li Zhou provided technical assistance on tissue culture, Dr. Rongda Qu, Dr. De-Yu Xie and Dr. Jenny Xiang provided scientific advice and guidance.

Plant Cell Tiss Organ Cult (2009) 97:27–37
DOI 10.1007/s11240-009-9495-0

ORIGINAL PAPER

Shoot regeneration of dwarf dogwood (*Cornus canadensis* L.) and morphological characterization of the regenerated plants

Chun-Miao Feng · Rongda Qu · Li-Li Zhou · De-Yu Xie · Qiu-Yun (Jenny) Xiang

Received: 17 October 2008 / Accepted: 16 December 2008 / Published online: 29 January 2009
© Springer Science+Business Media B.V. 2009

Abstract Dwarf dogwoods (or the bunchberries) are the only suffrutex in *Cornaceae*. They are attractive ground cover ornamentals with clusters of small flowers surrounded by petaloid bracts. Little has been reported on plant regeneration of dogwoods. As a step toward unraveling the molecular basis of inflorescence evolution in *Cornus*, we report an efficient regeneration system for a dwarf dogwood species *C. canadensis* through organogenesis from rejuvenated leaves, and characterize the development of the plantlets. We used the nodal stem segments of vegetative branches as explants. Micropropogated shoots were quickly induced from axillary buds of nodes on an induction medium consisting of basal MS medium supplemented with 4.44 μM BAP and 0.54 μM NAA. The new leaves of adventitious shoots were used as explants to induce calli on the same induction medium. Nearly 65% of leaf explants produced calli, 80% of which formed adventitious buds. Gibberellic acid (1.45 μM) added to the same induction medium efficiently promoted quick elongation of most adventitious buds, and 0.49 μM IBA added to the basal MS medium promoted root formation from nearly 50% of the elongated shoots. The growth of plantlets in pot soil was characterized by the development of functional woody rhizomes, which continuously developed new aboveground vegetative branches, but not flowering branches, within the past 12 months. Potential reasons causing the delay of flowering of the regenerated plants are discussed. The establishment of this regeneration system facilitates developing a genetic transformation system to test candidate genes involved in the developmental divergence of inflorescences in *Cornus*.

Keywords *Cornus canadensis* · Organogenesis · Regeneration · Rhizomes · Tissue culture

Abbreviations

BAP	6-Benzylaminopurine
GA	Gibberellic acid, $C_{19}H_{22}O_6$
IBA	Indole-3-butyric acid
MS	Murashige and Skoog
NAA	1-Naphthaleneacetic acid
PVPP	Polyvinylpolypyrrolidone
FAA	Formaldehyde

Introduction

The dogwood genus *Cornus* L. (*Cornaceae*) consists of species belonging to four major clades that are mostly trees and shrubs. These clades are the big-bracted dogwoods (e.g., *C. florida* L.), the dwarf dogwoods (or bunchberries, e.g., *C. canadensis* L.), the cornelian cherries (e.g., *C. mas* L.), and the blue- or white-fruited dogwoods (e.g., *C. alternifolia* L.f.). The four lineages exhibit striking differences in

Electronic supplementary material The online version of this article (doi:10.1007/s11240-009-9495-0) contains supplementary material, which is available to authorized users.

C.-M. Feng · L.-L. Zhou · D.-Y. Xie (✉) · Q.-Y. (Jenny) Xiang (✉)
Department of Plant Biology, North Carolina State University, Raleigh, NC, USA
e-mail: deyu_xie@ncsu.edu

Q.-Y. (Jenny) Xiang
e-mail: jenny_xiang@ncsu.edu

R. Qu
Department of Crop Science, North Carolina State University, Raleigh, NC, USA

inflorescence architectures (S-Fig. 1, available online). For example, the big-bracted dogwoods produce capitulate cymes subtended by four large petaloid bracts; while the dwarf dogwoods, the only herbaceous perennials with woody rhizomes in the genus, bear minute compound cymes subtended by four petaloid bracts. The cornelian cherries produce umbellate cymes subtended by four non-petaloid, scale-like bracts. Finally, the blue- or white-fruited dogwoods have large corymbose or paniculate cymes, with their branches associated with early deciduous, non-modified bracts (Xiang and Thomas 2008; Zhang et al. 2008). With a robust phylogeny (S-Fig. 1, available online); reconstructed from molecular and morphological data and abundant fossils (Xiang et al. 2006; Eyde 1988), the genus *Cornus* has served as a good model to examine questions concerning speciation and morphological evolution (including inflorescence architectures) and their relationship to molecular evolution (Xiang and Thomas 2008; Xiang et al. 2008). Inflorescence structures play important roles in angiosperm evolution and crop production, but our knowledge of the genetic basis of inflorescence development and evolution is largely limited to a few model plants (Benlloch et al. 2007). Comparative studies among closely related species with distinct inflorescence architectures will be essential to improve our understanding of the molecular basis of inflorescence evolution in flowering plants (Cronk et al. 2002; Kellogg 2004; Kellogg 2006; Malcomber et al. 2006; Singer et al. 1999). The presence of divergent inflorescence types among the closely related dogwood species provides an ideal system for such a study to uncover molecular changes that have altered inflorescence architectures during flowering plant evolution.

Identifying molecular mechanisms of inflorescence development will ultimately require analyses of candidate gene functions. At present, in vivo functional analyses of genes in plants have commonly employed the model plant *Arabidopsis thaliana*, which limits the implication of results, because the phenotypic effects of genes observed in *Arabidopsis* may be restricted to *Arabidopsis* and its close relatives. Obviously, transformation systems established for the particular species under study are ideal for in vivo analyses of gene functions in the relevant plants, e.g., dogwoods. In this regard, an efficient regeneration and genetic transformation system in dogwoods is an essential step toward the goal of unraveling the causal mechanisms of inflorescence divergence in *Cornus*.

In addition, many dogwood species are of great ornamental values due to their showy inflorescences, fruits, and foliage. For example, *C. florida* L. is the state flower of North Carolina and Virginia, USA. *C. kousa* Hance, *C. mas* L. *C. canadensis* L., and *C. sericea* L. are also popular garden plants. Some species are medicinally important (e.g., *C. officinalis* Seib. & Zucc.; Zao-Bi in Chinese medicine) (Do et al. 2004). Over the past 20 years, numerous efforts have been made to develop rapid micropropagation for dogwoods through plant tissue culture. As a result, micropropagation techniques have been established for several species, e.g., in the cornelian cherry (or Japanese yellow dogwood) *C. officinalis* Sieb. & Zucc. from immature embryos (Lu 1984lib; Lu 1985) and stems (Xue et al. 2003), in *C. florida* from shoot proliferation (Declerck and Korban 1994; Trigiano et al. 1992) and axillary buds of seedlings (Kaveriappa et al. 1997), in the rare pacific dogwood *C. nuttallii* Audubon from shoots (Edson et al. 1994), and in *C. canadensis* (Pennell 1983). In addition, plant regeneration of *C. florida* through somatic embryogenesis was reported by Trigiano et al. (1989). Regenerated *C. florida* plantlets grown in pot soil could develop the first pair of true leaves (Trigiano et al. 1989), although the growth of plantlets stopped prior to the formation of the second pair of true leaves (Trigiano et al. 1992). This demonstrated a potential to use plant tissue culture technology to breed value-added varieties of *C. florida* in the future. To our knowledge, little has been reported for plant regeneration of any dogwood species from callus.

In order to understand the genetic and molecular mechanisms of inflorescence evolution in dogwoods, we choose *Cornus canadensis*, a species from the dwarf dogwood group, as our model plant. *Cornus canadensis* (also known as bunchberry, bunchberry dogwood, and Canadian dwarf cornel) is a perennial diminutive suffrutescent species (Murrell 1994). The plant develops annual aboveground stems with usually 2 or 3 nodes from the perennial, woody rhizomes every spring. Only the upper node of the stem bears a "whorl" of six green leaves (two opposite leaves at the node, each bearing two axil leaves). The annual stem develops a terminal pedunculate compound cyme that is minute and subtended by four petaloid bracts. The flowers bloom from May through early July. It is native to boreal regions, extending south to the Rocky Mountains in Colorado and New Mexico, the southern Appalachians in Western Virginia and Maryland, and mountains in Japan. It thrives in forests and forest edges with moist, well-drained soils. The species is commonly grown as ground cover plants in the landscapes of the northern states in North America. We have grown this species for 4 years in a controlled environment. In growth chambers at the North Carolina State University Phytotron, the plants continuously develop new woody rhizomes and normal flowers and fruits as in nature. In this study, we report plant regeneration of *C. canadensis* and characterize morphological development of the regenerated plants.

Materials and methods

Chemicals, culture media and growth conditions

All chemicals used to make different culture media in this study were purchased from PlantMedia Company (Dublin, OH, USA).

The basal MS medium (Murashige and Skoog 1962) was used in all culture media in this study. All working media were supplemented with 3% (w/v) of sucrose, solidified by 0.8% phytoagar, adjusted pH value to 5.7 with 0.2 M NaOH, and then autoclaved at 121°C for 20 min. Plant growth regulators used in culture media were sterilized by filtering through a 0.2-μm membrane and added to the media after autoclaving when the temperature had cooled to 50–60°C. All cultures were incubated in a lighted growth chamber at 23 ± 1°C with a 16 h light/8 h dark photoperiod. The light intensity was 33 μmol m^{-2} s^{-1} from cool white fluorescent tubes.

Plant materials

The source plants of *C. canadensis* for tissue culture were originally collected in June 2004 from Spruce Knob, West Virginia. These plants (Fig. 1a–c) have been grown in the Phytotron at NCSU under the following conditions: 9 h high intensity of light (450–550 μmoles m^{-2} s^{-1}) at 18°C from 7:30 AM to 4:30 PM, followed by 15 h dark period at 14°C, during which a 3-h interruption of dark from 11:00 PM to 2:00 AM-was applied by illumination of an incandescent light (40 μmoles m^{-2} s^{-1}). The plants were watered daily.

Induction of adventitious shoots from nodal stem segments

Aboveground young vegetative shoots (approximately 1 month old or less) (Fig. 1b) were collected and surface sterilized with 70% ethanol for 1–2 min, followed by rinsing five times with autoclaved double deionized water (ddH_2O). They were further sterilized with 10% (v/v) bleach solution (sodium hypochlorite 6%; Clorox, Oakland, CA, USA) for 10 min, followed by rinsing five times with ddH_2O. The sterilized shoots were blot-dried with sterile paper towels. Leaves were then removed from the shoots, and the stems were cut into 8 mm long segments, each containing a node. Ninety nodal stem segments (Fig. 1d) were horizontally cultured on the basal MS medium (Murashige and Skoog 1962) supplemented with 4.44 μM of 6-benzylaminopurine (BAP) and 0.54 μM of 1-naphthaleneacetic acid (NAA), which were contained in 1 × 9 cm (height × diameter) plastic petri dishes. This medium was employed for micropropagation of other *Cornus* species (Declerck and Korban 1994; Edson et al.

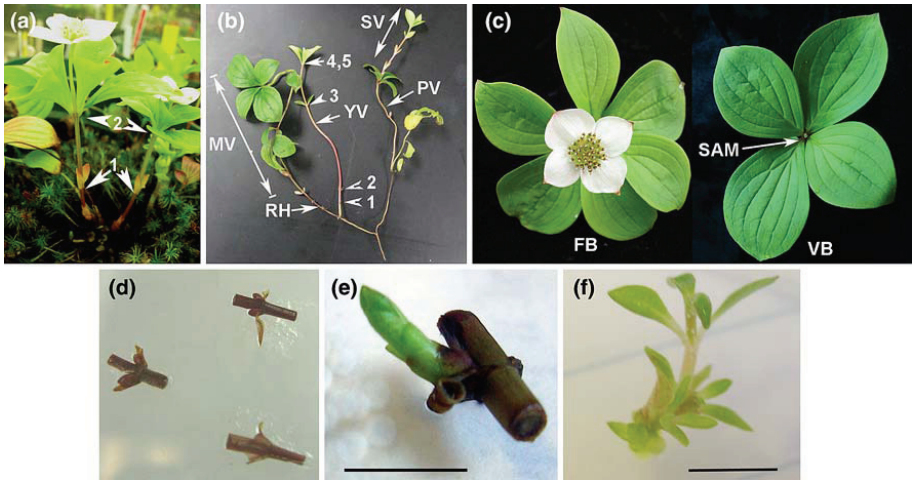


Fig. 1 Morphological identification of aboveground flowering and vegetative branches and micropropagation from the nodal segment of vegetative branches. **a** Two aboveground flowering branches; *1*, *2* the two nodes. **b** Rhizomes and aboveground vegetative branches; *1–5* the five nodes, *MV* mature aboveground vegetative branches, *PV* primary vegetative branch, *RH* rhizome, *SV* secondary vegetative branch, *YV* young aboveground vegetative branch. **c** Comparison of the top of an aboveground flowering branch and vegetative branch; *FB* flowering branch, *SAM* shoot apex meristem, *VB* vegetative branch. **d** Three nodal stem segments used as explants. **e** An example of nodal stem segment producing an adventitious shoot on the induction medium containing 4.44 μM BAP and 0.54 μM NAA. **f** Multiple adventitious shoots and leaves used as explants to induce calli

1994; Kaveriappa et al. 1997; Trigiano et al. 1992). The experiment was replicated three times.

Callus induction

Leaves from the micropropagated adventitious shoots and young leaves from aboveground vegetative stems of plants growing in a growth chamber were used as explants for callus induction.

Five media composed of the basal MS medium supplemented with 0.54 μM NAA and various concentrations (0.44, 2.22, 4.44, 6.67, or 8.89 μM) of BAP were tested to obtain optimal conditions of callus induction, and 30 ml of solid medium was contained in 1 × 9 cm plastic petri dishes. At least 60 leaves were tested for each treatment, with the leaves cut into approximately 0.3 × 0.5 cm pieces, 20 of which were inoculated into each petri dish with the abaxial side toward the medium. Callus induction was examined under a dissection microscope and imaged with a digital camera. Number of leaves that produced callus was recorded after 10 days of culture, and the rates of explants forming callus from each treatment were determined. This experiment was repeated three times.

Adventitious bud induction and elongation

Calli started to form adventitious buds after they were cultured on the callus induction medium containing 2.22, 4.44, or 6.67 μM BAP and 0.54 μM NAA for 1 week. Number of calli producing adventitious buds was recorded and percentage of shoot induction was calculated. These adventitious buds were subcultured on the same medium for additional 3 weeks before they were transferred onto a solid elongation medium.

Based on our previous work (Xie and Hong 2001), 1.45 μM gibberellic acid (GA_3) was included in the callus induction medium above to form an elongation medium, which was used to enhance elongation of multiple adventitious buds. Nearly 40-day-old multiple adventitious buds were transferred onto the elongation medium and cultured for a further 3 weeks.

Root induction and plant growth in soil

Seven media were tested to induce root formation. These root induction media were composed of the basal MS medium supplemented with indole-3-butyric acid (IBA) (0.49, 2.46, 4.93, or 7.39 μM) or NAA (0.54, 2.69, or 5.38 μM). Tests were conducted in 125-ml Erlenmeyer (E)-flasks containing forty ml of solid root induction medium.

Approximately 3- to 5-cm-long adventitious shoots with 3–5 internodes were excised from the base of the elongated clusters. Four to seven adventitious shoots were cultured in each E-flask. Between 12 and 21 adventitious shoots were tested for each medium. This experiment was repeated three times. After 2 weeks, adventitious shoots started to form adventitious roots.

After the plantlets formed at least 5 lateral roots, each longer than 1.5 cm, they were transplanted into pots containing soil mixed with vermiculite based on the original "Cornell Mix" (Boodley and Sheldrake 1972), with pH value of 5.7. Plantlets were grown in a growth chamber under fluorescent light (450–550 μmoles m^{-2} s^{-1}) with a photoperiod consisting of 9 h of day time at 18°C, followed by 15 h dark time at 14°C, during which a 3-h lighting interruption from 11:00 PM to 2:00 AM was provided by an incandescent light (40 μmoles m^{-2} s^{-1}). The plantlets were covered with plastic wraps (Anchor Packaging, Fenton, MO, USA) for the first week and watered with our nutrition solution (http://www.ncsu.edu/phytotron) daily. After 2 months, the percentage of surviving plantlets was calculated.

Histological analysis

Five- to seven-day-old calli induced from the medium containing BAP (4.44 μM) and NAA (0.54 μM) were used for histological analysis to examine the formation of adventitious buds from calli. Calli were fixed in 3.7% formaldehyde (FAA) solution for 2 h, then dehydrated through a gradual ethanol concentration series from 50 through 60, 70, 85, to 95% sequentially, 30 min each, and finally dehydrated in 100% ethanol for 30 min. They were further permeated with an analytical grade xylene series (25, 50, 75 and 100%) sequentially, 30 min each, and infiltrated by chips of paraplast®Plus (Fisherbrand, Houston, TX, USA) and embedded in Paraplast®Plus. Toluidine blue powder (O type from Sigma) of 0.025 g was dissolved in 100 ml of pH 5.5 sodium phosphate-citric acid buffer (containing 0.2 M disodium hydrogen phosphate and 0.1 M sodium citrate) to stain tissues. The embedded tissues were sectioned at a thickness of 8 μm and stained in toluidine blue (0.025%), mounted with Cytoseal-XYL (Apogent, Kalamazoo, MI, USA), and dried in an exhaust hood overnight (Ruzin 1999). Slides were examined under a Zeiss Axioscop2 microscope and images captured with a micropublisher 5.0 RTV digital camera and Q capture software (Q Imaging, Surrey, BC, Canada).

Fresh rhizomes were harvested from original ("mother") plants and from the regenerated plantlets. Hand sectioning was carried out to examine anatomical structures of the rhizomes. Freshly hand-sectioned tissues were stained as above and observed under the same microscope.

Statistical analysis

Analysis of variance (one-way ANOVA, Motulsky 1995) was performed to test the significance of differences between means obtained among the treatments in each experiment at the 5% level of significance ($P < 0.05$).

Results

Identification of vegetative and flowering branches for micropropagation

In the experiments, we observed that the underground woody rhizomes of *Cornus canadensis* growing in the NCSU Phytotron developed two different erect aboveground branches, flowering stems and vegetative stems, which are functionally and morphologically different (Fig. 1a, b). This observation was critical for us to develop a protocol of plant regeneration. A mature flowering branch is characterized by two nodes, the lower node with a pair of opposite scales and the upper one with one pair of opposite green leaves. From the axil of each leaf, one pair of decussate green leaves develop, giving the appearance of a whorl of six leaves at the upper node of the flowering stem (Fig. 1a). In addition, the growth of flowering branches is terminated by a small, compound, cymose inflorescence subtended by four petaloid bracts (Fig. 1a and c).

A vegetative branch (which does not develop a terminal inflorescence) is typically characterized by 3–5 nodes (Fig. 1b). In a stem with 5 typical nodes, the 2 lower nodes (defined as 1st and 2nd) develop a pair of opposite scales; the middle node (the 3rd node) has two small, opposite green leaves and is well separated from the 2nd and the two upper nodes (the 4th and 5th nodes). The 2 upper nodes are very close to each other without an obvious internode and each bears a pair of opposite, normal green leaves (Fig. 1b). In the stem with 3 or 4 nodes, 2 or 1 of the 3 lower nodes are missing compared to the stem with 5 nodes. All vegetative branches have an obvious shoot apex (Fig. 1c). Under phytotron conditions, some shoot apical meristems develop a second vegetative branch (referred to hereafter as the secondary branch) on top of the first branch (referred to hereafter as the primary branch) that directly develops from the rhizome (Fig. 1b). These morphological observations provide a basis for characterizing the regenerated plants in vitro.

Micropropagation of adventitious shoots

After 10 days of culture, adventitious shoots were induced from approximately 70% of the nodal segments (Fig. 1d) cultured on the basal MS medium (Murashige and Skoog 1962) supplemented with 4.44 μM of 6-benzylaminopurine (BAP) and 0.54 μM of 1-naphthaleneacetic acid (NAA) (Fig. 1d–f). Most nodal segments cultured produced 1–3 shoots from axillary buds (Fig. 1e, f). Two weeks later, leaves from these shoots (Fig. 1f) were used as explants for callus induction in this study. This experiment was repeated three times and similar results were obtained. In addition, we also tested nodal segments from flowering branches, but failed to get adventitious shoots on the same medium in this experiment.

Induction of calli and adventitious buds

Leaf explants from micro-propagated adventitious shoots started to form light green calli at the wound sites in 10 days (Fig. 2a). Among the five combinations of BAP (0.44, 2.22, 4.44, 6.67, and 8.89 μM) and NAA (0.54 μM) tested, the combination of 0.54 μM NAA and 4.44 μM BAP promoted nearly 62% of explants to form calli, slightly higher than the combination of 6.67 μM BAP and 0.54 μM NAA (Fig. 3a). The two combinations were significantly more efficient for callus induction than the rest of concentrations tested (Fig. 3a). Calli were subcultured on the same medium 2 weeks later.

During the 2 weeks of continuous growth, the initiation of adventitious buds started from 5- to 7-day-old calli that were induced from explants on the medium containing 2.22, 4.44, or 6.67 μM BAP and 0.54 μM NAA. The average percentage (78%) of adventitious bud formation from medium with 4.44 μM BAP and 0.54 μM NAA was higher than that formed on the other two media (data not shown). Transverse sections examined under a microscope supported the formation of adventitious buds from calli. The early adventitious bud development included an obvious apical meristem-like structure and one pair of leaf primordia (Fig. 2c). An axillary meristem-like structure was also observed at the bottom of leaf primordia (Fig. 2c). Further growth of these structures developed into obvious adventitious buds (Fig. 2d).

In contrast, young leaf explants directly collected from aboveground vegetative branches growing in the growth chamber could not form calli when cultured on the five media used above. This most likely resulted from the quick browning of the explants, which secreted brownish compounds. In an attempt to reduce the level of browning of explants, activated charcoal or polyvinylpolypyrrolidone (PVPP) in the concentrations of 0.1–6.0 g/l were tested in the experiments to inhibit the occurrence of explant browning (Bharadwaj and Ramawat 1993; Madhusudhanan and Rahiman 2000; Trigiano and Gray 2000). In another attempt, the explants were rinsed with 2.0 g/l ascorbic acid before they were cultured. However, the brown exudates persisted and no callus was induced from these explants.

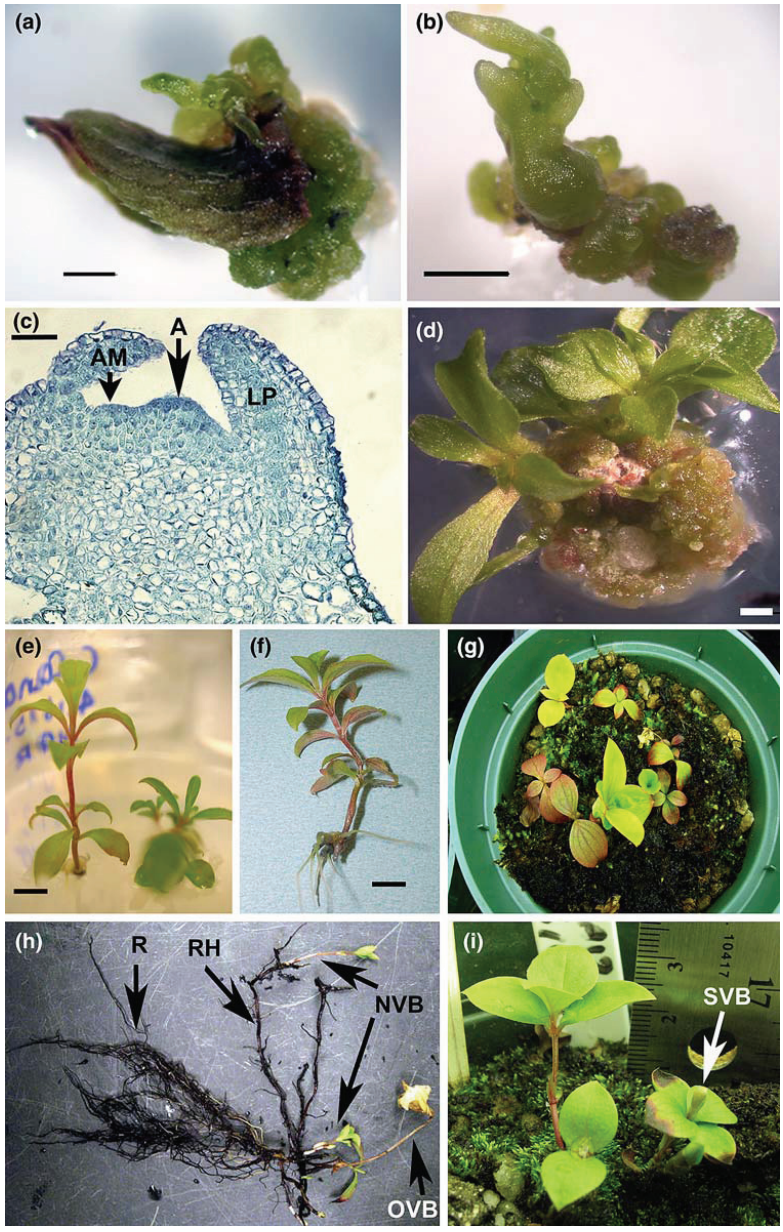


Fig. 2 Plant regeneration of dwarf dogwood through organogenesis. **a** Callus formation. **b** The early stage of an adventitious bud. **c** A microscopic image of transverse section shows the formation of an adventitious bud; *A* apical meristem, *AM* axillary meristem, *LP* leaf primordia. **d** Multiple adventitious buds. **e** Root induction of elongated adventitious shoots. **f** A plantlet with well developed adventitious roots. **g** Regenerated plants growing in pot soil. **h** Functional rhizomes developed from a regenerated plantlet; *NVB* new vegetative branches, *OVB* an old vegetative branch senesced, *R* roots, *RH* rhizomes. **i** Aboveground vegetative branches developed from regenerated plant rhizomes; *SVB* a secondary vegetative branch

The elongation of the adventitious buds

GA_3 is widely used to promote elongation of adventitious shoots in plant regeneration, e.g., for in vitro shoot proliferation of cassava (Bhagwat et al. 1996) and regeneration of *Acacia mangium* via organogenesis (Xie and Hong 2001). In this study, the elongation of the adventitious buds was very slow on the medium containing 4.44 μM BAP and 0.54 μM NAA. During 2 months of subculture on this medium, no adventitious bud could grow longer than

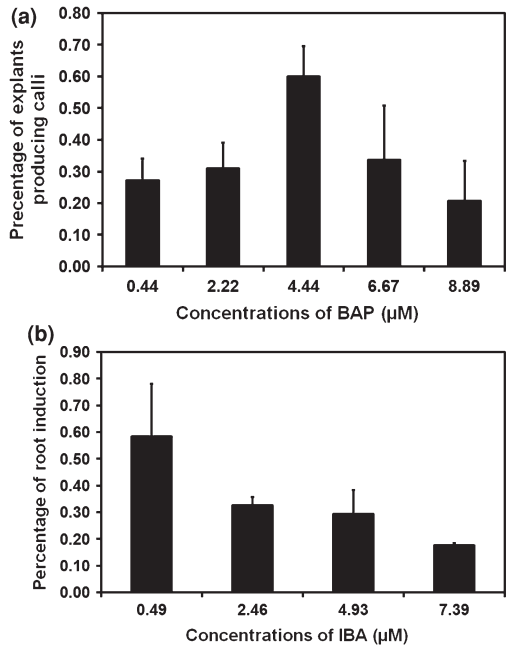


Fig. 3 Effect of BAP concentrations on callus formation (**a**) and effect of IBA concentrations on root induction (**b**). Statistical analysis is significant ($n = 12–21$, $P < 0.05$)

1.0 cm. Based on our previous work (Xie and Hong 2001), we added 1.45 μM of GA_3 into this medium. All adventitious buds (0.5–1.0 cm in length) tested in this study grew up to 2–4 cm long within just 3 weeks. Most of the adventitious shoots showed 4–5 nodes, each with one pair of opposite green leafy leaves (Fig. 2e). This study showed that GA_3 at 1.45 μM was effective to enhance elongation of the adventitious buds of *C. canadensis*.

Root induction of adventitious shoots

We evaluated the effects of IBA and NAA on root induction from elongated adventitious shoots. Four concentrations of IBA, 0.49, 2.46, 4.93, and 7.39 μM, were tested in this experiment. Adventitious shoots, 2–3 cm in length, were transferred onto rooting media for root induction. After 1 week, shoots started to form the first adventitious root on all media containing different concentrations of IBA tested and continued to produce more roots during the next 3 weeks of culture (Fig. 2f). In a 30-day period of root induction, 0.49 μM of IBA led nearly 48% of adventitious shoots to form adventitious roots, significantly higher than other concentrations tested (Fig. 3b). In contrast, the three NAA concentrations (0.54, 2.69, and 5.38 μM) tested could not induce root formation.

Growth of plantlets on agar medium and morphological properties

The morphology of plantlets was compared with both flowering and vegetative branches from original ("mother") plants. The morphology of 94 plantlets grown on solid media (Fig. 2f) is similar to that of a young aboveground vegetative branch rather than a flowering branch developed from rhizomes. After root formation, plantlets continued to grow new internodes and nodes from the shoot apex. However, the formation of the 6-clustered leaves as well as development of rhizomes was not observed in any plantlets on the rooting medium.

Developmental properties of plantlets in pot soil

In this experiment, 94 plantlets with 4–5 adventitious roots were planted in pot soil and grown in a lighted growth chamber. Fertilization and watering was applied in the same way as to the "mother" plants. Finally, 67 of them survived and were healthy after 2 months. The robust plantlets produced no or 1–2 new nodes after they were transferred to soil. More importantly, we observed that the growth of plantlets was mainly characterized by the development of rhizomes in the soil, from which new aboveground shoots started to form (Fig. 2g–i). Histological analysis by a cross-sectioning and a microscopic examination showed that the newly regenerated woody rhizomes were composed of epidermis, cortex, phloem, xylem, and pith (Fig. 4b). Because the regenerated plants were only 12 months old, the number of annual ring in the rhizomes was only one, fewer than a sample of rhizome taken from their original ("mother") plants (Fig. 4a). The dead plantlets did not show formation of the woody rhizomes. All plantlets that survived had developed rhizomes.

Numerous new aboveground branches were continuously produced from regenerated rhizomes after they were grown in the growth chamber for 12 months. However, no flowering branches were observed. All the new shoots developed from the regenerated plants are vegetative stems, as characterized at the beginning of "Results" (Fig. 2i). Similar to the mother plants, some mature vegetative branches developed a new secondary vegetative branch from the shoot apex of the primary branch (Fig. 2i). These observations indicate that the regenerated plants likely inherit the capability to develop normal aboveground vegetative stems of the mother plants.

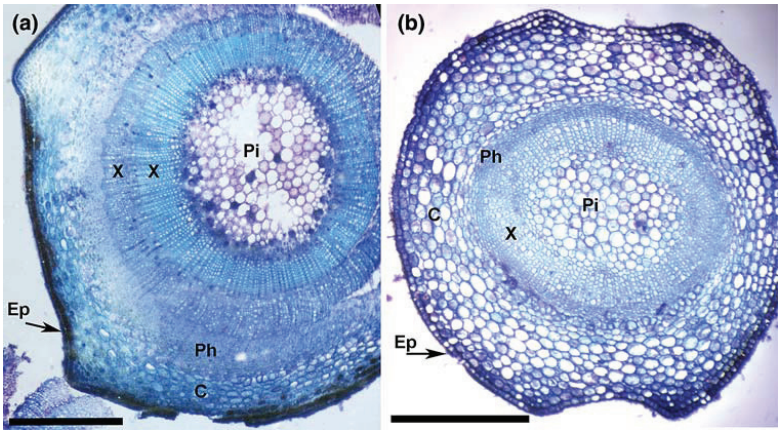


Fig. 4 Microscopic images showing anatomic structure of rhizomes. **a** A cross-section of a rhizome from an original "mother" plant. **b** A cross-section of a rhizome developed from a regenerated plant. *C* Cortex, *Ep* epidermis, *Ph* phloem, *Pi* pith, *X* xylem

Discussion

Many studies have shown that leaves from tissue-cultured adventitious shoots are useful to induce calli for plant regeneration. Examples include studies on plant regeneration of *Acacia mangium, Vaccinium macrocarpon*, *Vitis idaea*, and *V. corymbosum* (Xie and Hong 2002; Marcotrigiano et al. 1996; Meiners et al. 2007; Qu et al. 2000). We similarly found that juvenile leaves from micropropagated adventitious shoots efficiently produced calli in *Cornus canadensis*, which subsequently formed adventitious buds. This is in contrast to the young leaves from new shoots of the mother plant growing in pots, of which the explants did not grow calli, but showed a high content of brownish pigments. A high content of phenolics was reported to inhibit regeneration capacity of explants and callus (Abdelwahd et al. 2008; Lozovaya et al. 1996). It is clear that the failure of callus development from the explants derived from non-tissue cultured young leaves was due to inhibition by the brownish pigments (considered to be products of oxidation of phenolic compounds). It also suggests that the leaves from micropropagation through tissue culture do not produce brownish pigment on the medium after excision.

In this study, internodal stem segments from micropropagated adventitious shoots were also used to induce calli. The average percentage of callus induction from these stem segments was nearly 36% on the medium containing 4.44 μM BAP and 0.54 μM NAA which was used for leaf calli induction. However, unlike leaf-derived calli, these calli could not form adventitious buds on the same medium tested (data not shown), indicating differential responses of stem-derived calli and leaf-derived calli to the same medium. This observation differs from our previous report on *Acacia mangium* tissue culture, in which stem-derived calli could form adventitious bud-like structures just like the leaf-derived calli (Xie and Hong 2001).

Some shrub or subshrub species have been studied for regeneration due to their significant economic values, e.g., cranberry (*Vaccimum* spp.), lavender (*Lavandula angustifolia*), thorny burnet (*Sarcopoterium spinosum*),, periwinkle (*Catharanthus roseus*), brittlebrush (*Encelia farinose*), white burrobush (*Hymenoclea salsola*), thyme (*Thymus vulgaris*), and deerweed (*Lotus scoparius*) (Christodoulakis et al. 1990; Comstock and Mencuccini 1998; Marcotrigiano et al. 1996; Montalvo and Ellstrand 2000; Motomu et al. 2000; Nobel 1997; Nobel et al. 1998). Plant regeneration for cranberry was well established through organogenesis and micropropagation, and the regenerated cranberry plantlets were successfully grown in pot soil (Debnath and McRae 2001; Marcotrigiano et al. 1996; Qu et al. 2000). Regenerated periwinkle plantlets via organogenesis and somatic embryogenesis were shown to develop flowers after being transferred to pot soil (Junaid et al. 2007; Kim et al. 1994; Silva et al. 2003). However, little was reported about the development and propagation of woody rhizomes for cranberry and periwinkle plantlets. In our study, we found that the formation of woody rhizomes was probably critical for continuous survival. The new functional woody rhizomes were characterized by woody structures similar to that of the original mother plants (Fig. 4a, b) and continuously develop aboveground branches from nodes (Fig. 2i). This finding also suggests that formation of underground woody stems was probably the basis of continuous propagation of those regenerated plants of cranberry and periwinkle.

We observed multiple nodes on the vegetative stems and sometimes a secondary branch on top of the primary stem of *C. canadensis* growing in the Phytotron. This observation is surprising and a deviation from a typical natural

plant. In nature, all aboveground stems of the species normally have only 2 or 3 nodes and green leaves are produced only on the uppermost nodes. A secondary vegetative branch developed from the shoot apex of the primary branch has never been reported. However, an examination of specimens collected from the natural population of the source plant indicated that a flowering stem on the top of a vegetative stem sometimes occurs. Greenhouse conditions have probably interrupted the growth program of the species to some extent due to the lack of the winter period. The mother plants in the Phytotron continuously grow new shoots and develop flowers and fruits throughout the year. It takes about 2–3 months from a bud from the rhizome to grow mature fruits. The plants in the Phytotron have never gone through a dormant stage as in the natural environment. The lack of a dormant period may have altered the internal programming of vegetative growth, causing the stem to produce more nodes and more green leaves, an odd situation that rarely happens in nature. The regenerated plants replicated the deviated morphology of the vegetative stems (Figs. 1b, 2i).

At the present, the regenerated plants have not been able to produce flower stems, unlike the original mother plants which are continuously producing flowers and fruits year round. This result is different from those reports for regenerated periwinkle plantlets which developed flowers (Dhandapani et al. 2008; Junaid et al. 2007; Kim et al. 1994; Silva et al. 2003). A major difference between the dwarf dogwoods and the periwinkle plants growing in greenhouses is that the latter do not have obvious differentiation of aboveground vegetative and flowering branches. The result in dwarf dogwood is, however, consistent with the report of micropropagated plants from rhizomes and leaf sheath of mango ginger (*Curcuma amada*) (Prakash et al. 2004).

The development of inflorescences and flowers from regenerated dwarf dogwoods is essential to perform further studies on the molecular and genetic mechanisms of dogwood inflorescence evolution. The mechanism of why regenerated plantlets and new sprouts from regenerated rhizomes could not start to flower is still a mystery. It is possible that there is a generation time longer than 12 months required for this perennial species to mature to flower. The generation time for *C. canadensis* in the wild is unknown. Nursery information suggests that it may take 5 years for plants germinated from seeds to flower. It is also possible that the original vegetative branches used for explants may lack signals essential for the induction of inflorescence and flowers, e.g., expression of flowering genes *LEAFY* (Blazquez et al. 1997; Weigel et al. 1992) and *APETALA* (Irish and Sussex 1990; Jack et al. 1992; Mandel et al. 1992; Skipper 2002). If this is the case, the regenerated dwarf dogwood plants and rhizomes may have inherited the properties of vegetative growth, which results in no formation of flowers until the establishment of flowering signals de novo. This hypothesis can be tested by using aboveground flowering sprouts as explants to regenerate plants. Furthermore, the formation of flowering sprouts from regenerated rhizomes may require special treatments, e.g., treatments of growth hormones, low temperature, and special wavelength of light, etc., to stimulate flowering. We are currently testing the effects of temperature and GA_3 to gain some insights into this problem. The physiological condition of regenerated rhizomes may also be critical to the development of new flowering stems. It was reported that a relatively long storage duration of ginger rhizomes reduced the number of days required for flower initiation (Paz et al. 2005). The rhizomes of the regenerated dwarf dogwoods are approximately 1 year old, younger than their original "mother" plants (Fig. 4a, b). Thus, the rhizomes may still be too young to produce aboveground flowering stems. Finally, the regenerated plants could have been genetically changed due to mutations caused by tissue culture, which resulted in no development of flowers. We consider this to be less likely because mutations caused by tissue culture should occur randomly and are not expected to result in the same phenotypic effect of no flowering in all regenerated plantlets.

In conclusion, the accomplishments in this study provide a foundation for our subsequent work on testing candidate genes involved in the inflorescence development and evolution in the dogwood genus. Currently, by using the regeneration protocol established in this study, we have obtained transgenic calli and buds expressing a GUS reporter gene (Feng et al., in preparation). The target flowering genes that we are interested in include LEAFY, AP3 genes, etc. Our subsequent work will clone and functionally analyze gene expression and functions in the development of dwarf dogwood inflorescences. Furthermore, this study provides important information for regenerating other dogwood species and studying regeneration of other recalcitrant subshrub plants.

Acknowledgements The authors thank S. Azhakanandam for assistance in histological analysis, R. Franks, N. Robertson, W. Boss, and H. Saderoff for helpful discussion, R. Franks for facilities in histological analysis, and Norm Douglas for critical reading of the manuscript and English corrections. This study was supported by a Multidisciplinary Faculty Research Grant of North Carolina State University funded to Xiang, Xie, and Franks, and benefited from an USDA award (USDA 2006-35318-17431) and a NCSU university start-up fund to Xie. In addition, we thank NCSU Phytotron and staff for their help.

References

Abdelwahd R, Hakam N, Labhilili M, Udupa SM (2008) Use of an adsorbent and antioxidants to reduce the effects of leached

phenolics in in vitro plantlet regeneration of faba bean. Afr J Biotechnol 7:997–1002
Benlloch RA, Berbel A, Mislata S, Madueno F (2007) Floral Initiation and inflorescence architecture: a comparative view. Ann Bot (Lond) 100:659–676. doi:10.1093/aob/mcm146
Bhagwat B, Vieira LGE, Erickson LR (1996) Stimulation of in vitro shoot proliferation from nodal explants of cassava by thidiazuron, benzyladenine and gibberellic acid. Plant Cell Tissue Organ Cult 46:1–7. doi:10.1007/BF00039690
Bharadwaj L, Ramawat KG (1993) Effect of antioxidants and absorbents on tissue browning associated metabolism in *Coccu-lus pendulus* callus cultures. Indian J Exp Biol 31:715–718
Blazquez MA, Soowal LN, Lee I, Weigel D (1997) LEAFY expression and flower initiation in *Arabidopsis*. Development 124:3835–3844
Boodley JW, Sheldrake R Jr (1972) Cornell peat-lite mixes for commercial plant growing. Ext Info Bul 43 Cornell Univ
Christodoulakis NS, Tsimbani H, Fasseas C (1990) Leaf structural peculiarities in *Sarcopoterium spinosum*, a seasonally dimorphic subshrub. Ann Bot (Lond) 65:291–296
Comstock J, Mencuccini M (1998) Control of stomatal conductance by leaf water potential in *Hymenoclea salsola* (T. & G.), a desert subshrub. Plant Cell Environ 21:1029–1038. doi:10.1046/j.1365-3040.1998.00353.x
Cronk QCB, Bateman RM, Hawkins JA (2002) Developmental Genetics and Plant Evolution. Taylor & Francis, London and New York
Debnath S, McRae K (2001) An efficient in vitro shoot propagation of cranberry (*Vaccinium macrocarpon* ait.) by axillary bud proliferation. In Vitro Cell Dev Biol Plant 37:243. doi:10.1007/s11627-001-0043-9
Declerck V, Korban SS (1994) Effects of source of macro-nutrients and plant growth regulator concentrations on shoot proliferation of *Cornus florida*. Plant Cell Tiss Org Cult 38:57–60
Dhandapani M, Kim D, Hong S-B (2008) Efficient plant regeneration via somatic embryogenesis and organogenesis from the explants of *Catharanthus roseus*. In Vitro Cell Dev Biol Plant 44:18–25. doi:10.1007/s11627-007-9094-x
Do JR, Kang SN, Kim KJ, Jo JH, Lee SW (2004) Antimicrobial and antioxidant activities and phenolic contents in the water extract of medicinal plants. Food Sci Biotechnol 13:640–645
Edson JL, Wenny DL, Leege-Brusven A (1994) Micropropagation of Pacific dogwood. HortScience 29:1355–1356
Eyde RH (1988) Comprehending Cornus—puzzles and progress in the systematics of the dogwoods. Bot Rev 54:233–351. doi:10.1007/BF02868985
Irish VF, Sussex IM (1990) Function of the apetala-1 gene during Arabidopsis floral development. Plant Cell 2:741–753
Jack T, Brockman LL, Meyerowitz EM (1992) The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. Cell 68:683–697. doi:10.1016/0092-8674(92)90144-2
Junaid A, Mujib A, Bhat M, Sharma M, Šamaj J (2007) Somatic embryogenesis and plant regeneration in *Catharanthus roseus*. Biol Plant 51:641–646. doi:10.1007/s10535-007-0136-3
Kaveriappa KM, Phillips LM, Trigiano RN (1997) Micropropagation of flowering dogwood (*Cornus florida*) from seedlings. Plant Cell Rep 16:485–489
Kellogg EA (2004) Evolution of developmental traits. Curr Opin Plant Biol 7:92–98. doi:10.1016/j.pbi.2003.11.004
Kellogg EA (2006) Progress and challenges in studies of the evolution of development. J Exp Bot 57:3505–3516. doi:10.1093/jxb/erl132
Kim S, Song N, Jung K, Kwak S, Liu J (1994) High frequency plant regeneration from anther-derived cell suspension cultures via somatic embryogenesis in *Catharanthus roseus*. Plant Cell Rep 13:319–322
Lozovaya V, Gorshkova T, Yablokova E, Zabotina O, Ageeva M, Rumyantseva N, Kolesnichenko E, Waranyuwat A, Widholm J (1996) Callus cell wall phenolics and plant regeneration ability. J Plant Physiol 148:711–717
Lu WX (1984) Tissue culture of immature embryos of *Cornus officinalis*. Zhong Yao Tong Bao 9:7–8
Lu WX (1985) Rapid propagation of plants from the immature embryos of *Cornus officinalis*. Zhong Yao Tong Bao 10:9–10
Madhusudhanan K, Rahiman BA (2000) The effect of activated charcoal supplemented media to browing of in vitro cultures of Piper species. Biol Plant 43:297–299. doi:10.1023/A:1002729032527
Malcomber ST, Preston J, Reinheimer R, Kossuth J, Kellogg EA (2006) Developmental gene evolution and the origin of grass inflorescence diversity. Adv Bot Res 44:425–481
Mandel MA, Gustafson-Brown C, Savidge B, Yanofsky MF (1992) Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. Nature 360:273–277. doi:10.1038/360273a0
Marcotrigiano M, McGlew SP, Hackett G, Chawla B (1996) Shoot regeneration from tissue-cultured leaves of the American cranberry (*Vaccinium macrocarpon*). Plant Cell Tissue Organ Cult 44:195–199. doi:10.1007/BF00048524
Meiners J, Schwab M, Szankowski I (2007) Efficient in vitro regeneration systems for *Vaccinium* species. Plant Cell Tissue Organ Cult 89:169–176. doi:10.1007/s11240-007-9230-7
Montalvo AM, Ellstrand NC (2000) Transplantation of the subshrub *Lotus scoparius*: Testing the home-site advantage hypothesis. Conserv Biol 14:1034–1045. doi:10.1046/j.1523-1739.2000.99250.x
Motomu A, Yoshimoto O, Goki A (2000) Production of thymol by shoot culture of *Thymus vulgaris*. J Soc High Tech Agric 12:18–22
Motulsky H (1995) Intuitive biostatistics. Oxford University Press, New York, pp 299–300
Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol Plant 15:473–497. doi:10.1111/j.1399-3054.1962.tb08052.x
Murrell ZE (1994) Dwarf dogwoods: intermediacy and the morphological landscape. Syst Bot 19:539–556. doi:10.2307/2419776
Nobel PS (1997) Root distribution and seasonal production in the northwestern Sonoran Desert for a C3 subshrub, a C4 bunchgrass, and a CAM leaf succulent. Am J Bot 84:949–955. doi:10.2307/2446285
Nobel PS, Zhang H, Sharifi R, Castañeda M, Greenhouse B (1998) Leaf expansion, net CO_2 uptake, Rubisco activity, and efficiency of long-term biomass gain for the common desert subshrub *Encelia farinosa*. Photosynth Res 56:67–73. doi:10.1023/A:1006079631961
Paz MdP, Kuehny JS, McClure G, Graham C, Criley R (2005) Effect of rhizome storage duration and temperature on carbohydrate content, respiration, growth, and flowering of ornamental ginger. Acta Hortic 673:737–744
Pennell D (1983) The future of micropropagation in the United Kingdom. Proc Int Plant Prop Soc 33:249–253
Prakash S, Elangomathavan R, Seshadri S, Kathiravan K, Ignacimuthu S (2004) Efficient regeneration of *Curcuma amada* Roxb. Plantlets from rhizome and leaf sheath explants. Plant Cell Tissue Organ Cult 78:159–165. doi:10.1023/B:TICU.0000022553.83259.29
Qu L, James P, Nicholi V (2000) A highly efficient in vitro cranberry regeneration system using leaf explants. HortScience 35:948–952
Ruzin S (1999) Plant microtechnique and microscopy. Oxford University Press, New York

Silva RLD, Sousa CM, Santos RPD, Miranda RM (2003) In vitro regeneration of *Catharanthus roseus* nodal segment explants, under different auxin and cytokinin combinations. Agronomia 37:50–54

Singer S, Sollinger J, Maki S, Fishbach J, Short B, Reinke C, Fick J, Cox L (1999) Inflorescence architecture: a developmental genetics approach. Bot Rev 65:385–410. doi:10.1007/BF02857756

Skipper M (2002) Genes from the APETALA3 and PISTILLATA lineages are expressed in developing vascular bundles of the tuberous rhizome, flowering stem and flower primordia of *Eranthis hyemalis*. Ann Bot (Lond) 89:83–88. doi:10.1093/aob/mcf009

Trigiano RN, Gray DJ (2000) Plant tissue culture concepts and laboratory exercises. CRC Press, Boca Raton

Trigiano RN, Beaty RM, Dietrich JT (1989) Somatic embryogenesis and plantlet regeneration in *Cornus florida*. Plant Cell Rep 8:270–273. doi:10.1007/BF00274127

Trigiano RN, Beaty RM, Lowe KW (1992) Micropagation of dogwoods (*Cornus* spp.). In: Bajaj YPS (ed) Biotechnology in agriculture and forestry. Springer, Berlin, pp 81–90

Weigel D, Alvarez J, Smyth DR, Yanofsky MF, Meyerowitz EM (1992) LEAFY controls floral meristem identity in Arabidopsis. Cell 69:843–859. doi:10.1016/0092-8674(92)90295-N

Xiang Q-YJ, Thomas DT (2008) Tracking character evolution and biogeographic history through time in Cornaceae. Does choice of methods matter? J Syst Evol 46:349–374

Xiang Q-YJ, Thomas DT, Zhang WH, Manchester SR, Murrell Z (2006) Species level phylogeny of the Dogwood genus *Cornus* (*Cornaceae*) based on molecular and morphologicalevidence—implication in taxonomy and tertiary intercontinental migration. Taxon 55:9–30

Xiang Q-YJ, Thorne JL, Seo T-K, Zhang WH, Thomas DT, Ricklefs RE (2008) Rates of nucleotide substitution in Cornaceae (Cornales)—pattern of variation and underlying causal factors. Mol Phylogenet Evol 49:327–342. doi:10.1016/j.ympev.2008.07.010

Xie D, Hong Y (2001) In vitro regeneration of *Acacia mangium* via organogenesis. Plant Cell Tissue Organ Cult 66:167. doi:10.1023/A:1010632619342

Xie DY, Hong Y (2002) Agrobacterium-mediated genetic transformation of *Acacia mangium*. Plant Cell Rep 20:917. doi:10.1007/s00299-001-0397-9

Xue JP, Zhang AM, Wang YH, Sheng W (2003) Study on plant tissue culture of *Cornus officinalis*. Zhong Guo Zhong Yao Za Zhi 28:118–121

Zhang WH, Xiang Q-YJ, Thomas DT, Wiegmann BM, Soltis DE (2008) Molecular evolution of PISTILLATA-like genes in the dogwood genus *Cornus* (*Cornaceae*). Mol Phylogenet Evol 47:175–195. doi:10.1016/j.ympev.2007.12.022

Chapter III

Phylogeny-based developmental analyses illuminate evolution of inflorescence architectures in dogwoods (*Cornus* s. l., Cornaceae)

Chun-Miao Feng[1], Qiu-Yun (Jenny) Xiang[1] and Robert G. Franks[2]

[1]Department of Plant Biology, North Carolina State University, Raleigh, NC 27695,USA

[2]Department of Genetics, North Carolina State University, Raleigh, NC 27695, USA

Published in New Phytologist, 2011, accepted

Chunmiao did all of the work reported in this paper, Dr. Jenny Xiang and Dr. Robert Franks provided scientific advice and guidance.

New Phytologist

Research

Phylogeny-based developmental analyses illuminate evolution of inflorescence architectures in dogwoods (*Cornus* s. l., Cornaceae)

Chun-Miao Feng[1], Qiu-Yun (Jenny) Xiang[1] and Robert G. Franks[2]

[1]Department of Plant Biology, North Carolina State University, Raleigh, NC 27695, USA; [2]Department of Genetics, North Carolina State University, Raleigh, NC 27695, USA

Authors for correspondence:
Qiu-Yun (Jenny) Xiang
Tel: +1 919 515 2728
Email: jenny_xiang@ncsu.edu

Robert G. Franks
Tel: +1 919 513 7705
Email: rgfranks@ncsu.edu

Received: *13 January 2011*
Accepted: *23 February 2011*

New Phytologist (2011)
doi: 10.1111/j.1469-8137.2011.03716.x

Key words: ancestral character state reconstruction, *Cornus,* evolutionary development, inflorescence evolution, umbels and heads.

Summary

• Inflorescence architecture is important to angiosperm reproduction, but our knowledge of the developmental basis underlying the evolution of inflorescence architectures is limited. Using a phylogeny-based comparative analysis of developmental pathways, we tested the long-standing hypothesis that umbel evolved from elongated inflorescences by suppression of inflorescence branches, while head evolved from umbels by suppression of pedicels.

• The developmental pathways of six species of *Cornus* producing different inflorescence types were characterized by scanning electron microscopy (SEM) and histological analysis. Critical developmental events were traced over the molecular phylogeny to identify evolutionary changes leading to the formation of umbels and heads using methods accounting for evolutionary time and phylogenetic uncertainty.

• We defined 24 developmental events describing the developmental progression of the different inflorescence types. The evolutionary transition from paniculate cymes to umbels and heads required alterations of seven developmental events occurring at different evolutionary times.

• Our results indicate that heads and umbels evolved independently in *Cornus* from elongated forms via an umbellate dichasium ancestor and this process involved several independent changes. Our findings shed novel insights into head and umbel evolution concealed by outer morphology. Our work illustrates the importance of combining developmental and phylogenetic data to better define morphological evolutionary processes.

Introduction

Alterations in inflorescence architecture result in biological changes associated with pollination and reproduction that can drive new ecological adaptation and speciation (Wyatt, 1982; Fishbein & Venable, 1996; Friedman & Harder, 2005). During angiosperm evolution, innovation in inflorescence architecture has repeatedly occurred in different angiosperm lineages via alterations in the developmental pathways responsible for the ancestral forms (Tucker & Grimes, 1999; Soltis *et al.*, 2005). Despite the importance of understanding how developmental and genetic changes have shaped inflorescence architecture, little progress has been made, in part because existing model organisms exhibit little variation in these traits and there have been few comparative studies between species from a phylogenetic perspective (Cronk *et al.*, 2002; Kellogg, 2004, 2006).

The architecture of an inflorescence in part depends on its branching pattern and the relative position at which flowers are borne and is controlled by diverse developmental pathways (Benlloch *et al.*, 2007). In an indeterminate (racemose) inflorescence, the apices do not end in a terminal flower, and are able to grow for a long period, generating a continuous main axis that laterally produces floral meristems or branches. In a determinate (cymose) inflorescence, the apices of the main and lateral axes terminate in a flower. Much additional diversity of forms is determined by variation in the duration of the 'vegetative

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

state' of the inflorescence meristem (IM) during which the bud produces additional inflorescence branches instead of generating flowers (Prusinkiewicz *et al.*, 2007) as well as the variation in the length of inflorescence branches and pedicels. Among the great diversity of inflorescence architectures, raceme, spike, corymb, cyme, umbel, and head are the most common simple forms. These simple forms may assemble in a variety of ways to form more complex architectures. Raceme, spike, and corymb are elongated indeterminate inflorescences. The cyme is an elongated determinate form, while umbel and head are condensed forms that can be either determinate or indeterminate. At the present, the development of raceme and cymes are well understood in a few model plants such as *Arabidopsis*, *Antirrhinum*, rice, maize, petunia and tomatoes (Coen *et al.*, 1990; Alvarez *et al.*, 1992; Weigel *et al.*, 1992; Ingram *et al.*, 1995; Blazquez *et al.*, 1997; Bradley *et al.*, 1997; Lee *et al.*, 1997; Souer *et al.*, 1998; Samach *et al.*, 1999; Chuck *et al.*, 2002; Komatsu *et al.*, 2003; Satoh-Nagasawa *et al.*, 2006; Lippman *et al.*, 2008; Rebocho *et al.*, 2008; Souer *et al.*, 2008; Li *et al.*, 2009; Li *et al.*, 2010), but we know little regarding the developmental and genetic changes underlying the evolution of these inflorescence types, let alone the heads and umbels for which the developmental bases are poorly understood. Although an elegant model has been proposed to explain inflorescence diversity (Prusinkiewicz *et al.*, 2007), this model does not explain well the formation of umbels and heads. It has long been hypothesized that determinate umbels and heads evolved from branched inflorescences (e.g. panicles) by the suppression of inflorescence branches to form umbels and by suppression of pedicels in umbellate forms to produce heads (Parkin, 1914; Stebbins, 1974; Wyatt, 1982; Harris, 1999). Recently, Endress (2010) raised questions on this hypothesis based on the evidence that basal angiosperm taxa do not bear panicles, and suggested that detailed comparative inflorescence studies combined with phylogeny are needed to test this hypothesis and to better understand inflorescence evolution. To our knowledge, there have been no previous comparative studies combining the analysis of developmental morphologies and phylogeny to elucidate the evolutionary trend and the underlying developmental basis between elongated branched inflorescences and condensed forms such as umbels and heads in any angiosperm lineages.

In this paper, the genus *Cornus* (dogwood) was chosen as our model group to investigate the evolutionary direction and underlying developmental basis among elongated compound cymes (panicles, *sensu* Endress, 2010), umbels (i.e. sciadioid in Endress, 2010) and heads (i.e. cephalioid in Endress, 2010). *Cornus* contains *c.* 50 species that are divided into four major subgroups (Eyde, 1988; Murrell, 1993; Xiang *et al.*, 1998, 2006; Fan & Xiang, 2001). The phylogenetic relationships of these four groups are well resolved by phylogenetic studies based on multiple gene sequences and morphological data (Fan & Xiang, 2001, 2003; Xiang *et al.*, 2006, 2008; Fig. 1). These four groups are divergent in inflorescence architectures, exhibiting a range of branch condensation patterns as well as petaloid bracts in some species (Fig. 1). The first inflorescence type, found in the blue- or white-fruited (BW) lineage, is characterized by paniculate cymes with rudimentary bracts. The second inflorescence type found in the dwarf dogwood (DW) group is a condensed form of a dichasium with fewer and shorter branches, referred to as a 'minidichasium' hereafter. In the dwarf dogwood (DW) group the entire inflorescence is subtended by four large petaloid involucral bracts. The third inflorescence type is found in the cornelian cherry (CC) group bearing a determinate umbel (ciadioid) with four unexpanded, nonpetaloid, involucral bracts. The fourth inflorescence type found in *Cornus* is a completely condensed structure, a determinate head (cephaloid) with four or six large petaloid, involucral bracts that is found in the big-bracted (BB) group. In an earlier analysis of inflorescence evolution in *Cornus* based on outer morphology, umbels were inferred to be derived from paniculate cymes while heads were inferred to have evolved either directly from paniculate cymes or via umbels (Xiang & Thomas, 2008). These hypotheses could not be distinguished because of a lack of knowledge of the development morphologies.

In this paper, we conduct a comparative developmental morphological analysis of six *Cornus* species that represent the four major clades and cover the four major inflorescence types found in this genus. We characterized the developmental morphological progression for these six species via scanning electron microscopy (SEM) and histological analyses using materials collected over a 3-yr period. We

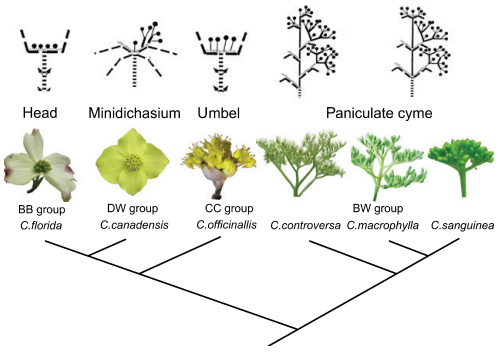


Fig. 1 Four major clades of *Cornus* L. showing their phylogenetic relationships from Xiang *et al.* (2008) and the images of inflorescence types. Schematic diagrams of inflorescence structures are reproduced from Murrell (1993) with permission from American Society of Plant Taxonomists. BB, big-bracted; DW, dwarf dogwood; CC, cornelian cherry; BW, blue- or white-fruited.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

describe a total of 24 developmental events, 13 correlate with alterations in inflorescence structure and size. We identify seven of these developmental events that are important to the development of umbels and heads. We trace these seven characters over the molecular phylogeny using methods that account for evolutionary time and phylogenetic uncertainty to uncover the evolutionary histories of umbels and heads. Our results suggest that heads and umbels in *Cornus* evolved independently from elongated forms via a small umbellate dichasium ancestor. Our results further reveal that the origins of umbels and heads each required multiple evolutionary changes that occurred independently in the CC and BB lineages in addition to the common changes that occurred in their common ancestor. These events include suppression of elongation of the rachis supporting the IM, suppression of inflorescence branch meristem (IBM) elongation as well as alterations in pedicel elongation. Additionally, changes in the pattern of primary IBM initiation and a reduction in the number of primary IBMs in early development support the formation of umbel and head forms, suggesting a much more complicated developmental evolutionary history than previously predicted from the analysis of the mature inflorescence morphology.

Materials and Methods

Sample selection and inflorescence bud collection

Six species spanning four groups of *Cornus* were chosen to cover the four different inflorescence forms in this genus: (*Cornus controversa* Hemsley, *C. sanguinea* L. and *C. macrophylla* Wall. from the BW group, producing paniculate cymes; *C. officinallis* Seib. & Zucc. from the CC group, producing umbels; *C. florida* L. from the BB group bearing heads (or cephaloid inflorescences); and *C. canadensis* L.f. from the DW group, bearing the minidichasium). Inflorescence bud samples of most of those species were collected every week after transiting to inflorescence growth (see the Results section) in 2007, 2008, 2009 and 2010 at J. C. Arboretum, Raleigh, NC, USA, except for *C. florida* and *C. canadensis*. The inflorescence buds of *C. florida* were collected from plants grown on the NC State University campus and those of *C. canadensis* were collected from plants grown in the phytotron of NC State University which were propagated from specimens collected in Cheshire, NH, USA. Two or more individuals were chosen for each species, if available, to cover potential variation within species. Usually 20–30 buds per individual were fixed in FAA (formaldehyde (37%) : acetic acid : 100% ethanol : ddH_2O, 2 : 1 : 10 : 7, v/v) for histological study and *c.* 10–15 buds per individual in 3.0% glutaraldehyde in 25 mM potassium phosphate buffer (pH 7.0) at 4°C for at least 24 h for SEM study.

Observation of inflorescence developmental stages in *Cornus*

For SEM observation, preserved inflorescence buds were first washed in 25 mM potassium phosphate buffer (pH 7.0) at 4°C (three times at 1 h each), then dehydrated in 30, 50, 70 and 95% ethanol at 4°C (1 h each), followed by dehydration in 100% ethanol for 24 h at 4°C, which was repeated with two different temperature regimes: 4°C followed by room temperature and only at room temperature. Samples were then critical point dried using liquid carbon dioxide for 15 min using a Tousimis SAMDRI-795 (Tousimis Research Corp., Rockville, MD, USA) located in the Center for Electron Microscopy (CEM), NC State University. Buds were then dissected under the dissecting microscope, mounted on aluminum stubs with Pelco tape™ (Ted Pella, Inc., Redding, CA, USA) and finally sputter-coated with gold/palladium using a Hummer 6.2 Sputter System (Anatech, Union City, CA, USA) in the CEM. Prepared inflorescence buds were examined using a JEOL JSM-5900LV SEM (Jeol, Peabody, MA, USA) at 10, 15 or 20 kV. Images were captured digitally and colorized using GIMP (http://www.gimp.org/).

The protocol for histological analysis followed Feng *et al.* (2009). For better penetration of dogwood inflorescence buds in FAA, hairy scales or hard bracts, which tightly protected the inflorescence buds, were removed and buds were vacuum-infiltrated for 30 min during fixation. After 24 h fixation, tissues were dehydrated and embedded in Paraplast® Plus (Paraplast Plus, Fisher Healthcare, Houston, Texas, USA). Sections were sectioned at 8 μm thickness using a rotary microtome and stained with toluidine blue (0.025%). Slides were observed under a Zeiss Axioscope2 microscope and images were captured using a Micropublisher 5.0 RTV digital camera and Q capture software (Q Imaging, Surrey, BC, Canada).

Analysis of sizes of IMs

Mature IMs were usually found to be dome-shaped or disk-like in *Cornus* species. The width and height of IMs were measured to compare the relative sizes of IMs in different species. The width of IM was defined as the distance between two attachment points of the two youngest leaves adjacent to the IM in the BW and DW groups in which bracts originated within the IM. In the CC and BB groups where bracts originated outside the IM, the width of the IM is the distance between the two attachment points of the second pair of the involucral bracts. The distance can be easily measured in the SEM images using ImageJ (http://rsbweb.nih.gov/ij/). For histological images, only the medium longitudinal section through the pair of youngest leaves or bracts could be used. Height of the IM was defined as the height of the group of meristematic cells in histological

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

New Phytologist (2011)
www.newphytologist.com

Table 1 Comparison of developmental events (DEs) that differ among species of *Cornus*

		BW group			CC group	BB group	DW group
DE	Characters	*C. san**	*C. mac**	*C. con**	*C. off**	*C. flo**	*C. can**
2–3	Bract initiation pattern	After maturation of IMs	After maturation of IMs	After maturation of IMs	Before maturation of IMs	Before maturation of IMs	After maturation of IMs
4	**Size of inflorescence meristems (Ch.A)**	Large	Large	Large	Small	Small	Small
7	**Elongation of rachis supporting apical inflorescence meristem (Ch.B)**	Visible	Visible	Visible	Invisible	Invisible	Invisible
5–9	**Initiation pattern of 1^{O} IBMs (Ch.C)**	Decussately, sequentially	Decussately, sequentially	Decussately, sequentially	Whorled, simultaneously	Decussately, sequentially	Decussately, sequentially
5–9	**Number of 1^{O} IBMs (Ch.D)**	8 in four pairs	8 in four pairs	5 or 6	10 in four pairs	6 in three pairs	4 in two pairs
10–12	**IBM elongation (Ch.E)**	Visible	Visible	Visible	Invisible	Invisible	Visible
10–11	Highest order of IBMs in distal pair of 1^{O} IBMs	2^{O} IBMs	3^{O} or 4^{O} IBMs	4^{O} IBMs	2^{O} IBMs	2^{O} IBMs	3^{O} or 4^{O} IBMs
12	Highest order of IBMs in inflorescence	4^{O} IBMs	5^{O} or 6^{O} IBMs	6^{O} IBMs	2^{O} IBMs	2^{O} or 3^{O} IBMs	3^{O} or 4^{O} IBMs
14	**Pedicel initiation during floral organogenesis (Ch.F)**	No initiation	No initiation	No initiation	Initiation	No initiation	Initiation
19	Awn development on petals	No awn	No awn	No awn	No awn	No awn	Awn developed
15–20	Timing pattern of floral organogenesis	Not all before winter	Not all before winter	Not all before winter	Before winter	Before winter	Before winter
21	**Pedicel initiation during anthesis (Ch.G)**	Initiation	Initiation	Initiation	No initiation	No initiation	No initiation
23	Bract petaloidy	Not petaloidy	Not petaloidy	Not petaloidy	Not petaloidy	Petaloidy	Petaloidy

Ch.A–G indicates characters A–G analyzed by character mapping, corresponding to those in Fig. 14 and Supporting Information Fig.S1.
*Abbreviation of species names: *C. san*, *C. sanguine*; *C. mac*, *C. macrophylla*; *C. con*, *C. controversa*; *C. off*, *C. officinallis*; *C. flo*, *C. florida*; *C. can*, *C. canadensis*. Characters were defined based on differences of DEs at five developmental stages. Characters in bold text were selected for analysis of ancestral character state reconstruction on a phylogenetic framework with character coding matrix shown in Table 2 and Fig. 14. Detailed information for stages and DEs are referred to in the legend of Fig. 2 and the Results section. IM, inflorescence meristem; IBM, inflorescence branch meristem; BB, big-bracted; DW, dwarf dogwood; CC, cornelian cherry; BW, blue- or white-fruited.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

sections in the inflorescence bud. The meristematic cells were identified based on their cellular morphology (smaller in size, more uniform in morphology and stain, typically rectangular with fewer and smaller vacuoles; see the Results section, Fig. 4). Pairwise comparison of the width and height of IMs among the four groups was performed using Student's *t*-test in R-2.11.1 (http://www.r-project.org) to determine whether there are significant differences in IMs among groups.

Tracing the evolutionary history of umbels and heads in *Cornus*

Seven developmental events (Table 1, characters in bold) were shown to differ between condensed and branched forms, suggesting that they are important to the formation of umbels and heads (Table 2). These characters were chosen for analysis of ancestral state reconstruction to identify the evolutionary developmental changes leading to heads and umbels. The characters were coded as binary data in the analysis. Character ancestral state reconstruction was performed using BayesTraits V1.0 (Pagel & Meade, 2006). The analysis used 1000 MrBayes trees from previous molecular phylogenetic study of *Cornus* (Xiang *et al.*, 2008) to account for phylogenetic uncertainty. The phylogeny contains 22 species of *Cornus* distributed among the four major clades and two outgroup taxa, the sister genus of *Cornus*, *Alangium* and *Diplopanax* from Cornales. For species present in the phylogeny but not included in the development study, their character states were coded based on their adult inflorescence morphology in comparison to the morphology of the studied species, to determine the presence or absence of a developmental event. When character state was uncertain for a taxon, it was coded as missing state. For character state coding of *Diplopanax*, the inflorescence morphology of *Curtisia* and *Grubbia*, recently found to be the sister clade of *Cornus-Alangium* (Xiang *et al.*, 2011), was considered. The structure of adult inflorescences in various species was determined from herbarium specimens or living collections in the J. C. Raulston Arboretum. Species of the genus that were not included in the phylogeny were all from the BW group, members of which have their inflorescence structure con-

Table 2 Character state matrix of seven characters important to the development and evolution of umbels and heads in *Cornus*

	Characters						
Species	A	B	C	D	E	F	G
C. nuttallii Audubon	1	1	0	2	1	0	1
C. florida L.	1	1	0	2	1	0	1
C. kousa Hance	1	1	0	2	1	0	1
C. disciflora Moc. & Sesse ex DC.	1	1	0	2	1	0	1
C. capitata Wall.	1	1	0	2	1	0	1
C. oligophlebia Merr.	0	0	0	0	0	0	0
C. hongkongensis Hemsley	1	1	0	2	1	0	1
C. alternifolia L.f.	0	0	0	1	0	0	0
C. controversa Hemsley	0	0	0	1	0	0	0
C. peruviana J.F.Macbr.	0	0	0	0, 2	0	0	0
C. walteri Wangerin	0	0	0	0	0	0	0
C. racemosa Lam.	0	0	0	0	0	0	0
C. oblonga Wall.	0	0	0	0	0	0	0
C. chinensis Wangerin	1	1	1	0	1	1	1
C. sessilis Torr. Ex Durand	1	1	1	0	1	1	1
C. eydeana QY Xiang & YM Shui	1	1	1	0	1	1	1
C. mas L.	1	1	1	0	1	1	1
C. officinallis Seib. & Zucc.	1	1	1	0	1	1	1
C. volkensii Harms	1	1	1	0	0,1	0,1	0,1
C. suecica L.	1	1	0	3	0	1	1
C. canadensis L.f.	1	1	0	3	0	1	1
C. unalaschkensis Ledeb.	1	1	0	3	0	1	1
Alangium	?	0	0	4	0	0	0

Characters and states (character number corresponding to Table 1):
A, size of inflorescence meristems: 0, large; 1, small.
B, elongation of rachis supporting apical inflorescence meristem: 0, visible; 1, invisible.
C, initiation pattern of primary (1^{O}) inflorescence branch meristems: 0, decussately and sequentially; 1, close to whorled and simultaneously.
D, number of 1^{O} inflorescence branch meristems: 0, 8 or 10 in four pairs; 1, 5 or 6; 2, 6 in three pairs; 3, 4 in two pairs; 4, fewer than four.
E, inflorescence branch meristem elongation: 0, visible; 1, invisible.
F, pedicel initiation during floral organogenesis: 0, no initiation during floral organogenesis; 1, initiation during floral organogenesis.
G, pedicel initiation during anthesis: 0, initiation during anthesis; 1, no initiation during anthesis.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

served as the elongated forms. A maximum likelihood (ML) analysis was first performed to obtain a sense of the average values of the rate parameters for the BayesTraits analysis. An interval from 0 to 15 was then set as the prior of the exponential distribution of the rate parameters. The values of 70 and 90 were chosen for the ratedev parameter to get an acceptance rate between 20 and 40% as suggested by the manual. A total of 10 000 000 iterations were run for the analysis of each character. The ancestral state with the highest value of the posterior possibility was reported for the node connecting the major clades. The density distributions of the posterior possibilities were depicted by R-2.11.1 (http://www.r-project.org/). For nodes where the posterior possibility of reported ancestral state was < 0.70, the fossil command was used to test whether there is positive support of the reported state over other states by calculating the Bayes factor as described in the manual (Pagel & Meade, 2006).

The results from BayesTraits were further compared with analyses using two alternative methods available in Mesquite 2.01 (Maddison & Maddison, 2007) using the best phylogenetic tree. Each character was traced over the phylogram of the best phylogenetic tree using the ML method with the mk1 model (one-parameter Markov k-state model, a generalization of the Jukes–Cantor model), following Xiang & Thomas (2008). For characters that are polymorphic for some species – for example, *Cornus peruviana* of the BW group and *C. volkensii* of the CC group are polymorphic for one and three characters, respectively (Table 2) – the parsimony method was used to trace the character evolution because the ML method does not accept polymorphism.

Results

Inflorescence development in *Cornus*

Inflorescence buds for the next year start to develop immediately after anthesis, or 1 or 2 months later depending on the species. We describe here a total of 24 developmental events (DEs) that characterize the inflorescence developmental pathway of the *Cornus* species we examined (Fig. 2). Some of the DEs are shared among all *Cornus* species we examined while others are lineage-specific (Fig. 2, Tables 1, 2). The developmental events are grouped here into five main developmental stages as described in the following.

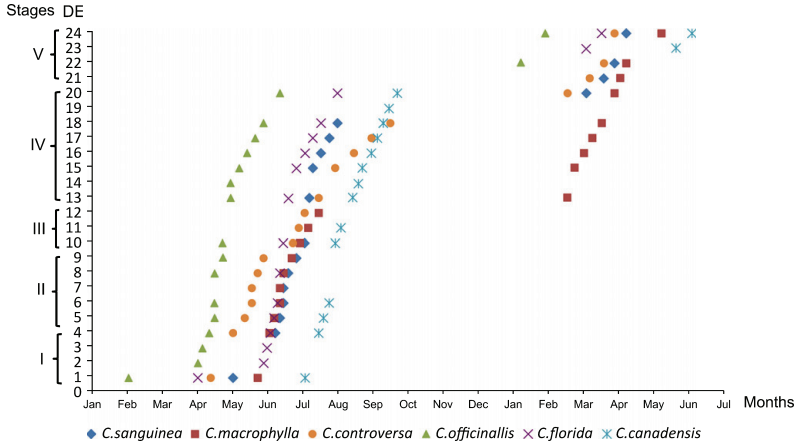


Fig. 2 Summary of inflorescence developmental pathways for the six species of *Cornus* examined in the study. Developmental events (DEs, *y*-axis) are plotted against the time when the events occur. Stage I – transition of vegetative meristem to inflorescence meristem (IM): DE1, vegetative meristem at the end of blooming season; DE2, initiation of the first pair of bract primordia in the peripheral zone of the IM; DE3, initiation of the second pair of bract primordia in the peripheral zone of the IM; DE4, mature IM ready to generate primary (1^{o}) inflorescence branch meristems (IBMs). Stage II – initiation of primary (1^{o}) IBMs: DE5, initiation of the first group of 1^{o} IBMs in the peripheral zone of the IM with or without bract subtending; DE6, initiation of the second group of 1^{o} IBMs in the peripheral zone of the IM with or without bract subtending; DE7, elongation of rachis supporting the apical IM; DE8, initiation of the third group of 1^{o} IBMs in the peripheral zone of the IM with or without bract subtending; DE9, initiation of the fourth group of 1^{o} IBMs in the peripheral zone of the IM with or without bract subtending. Stage III – initiation of higher-order IBMs: DE10, initiation of secondary (2^{o}) IBMs in the peripheral zone of the 1^{o} IBMs; DE11, initiation of tertiary (3^{o}) IBMs in the peripheral zone of the 2^{o} IBMs; DE12, initiation of other higher-order IBMs in the peripheral zone of the 3^{o} IBMs. Stage IV – floral organogenesis: DE13, initiation of floral meristems; DE14, initiation and elongation of pedicels during floral organogenesis; DE15, initiation of sepal primordia; DE16, initiation of petal primordia; DE17, petal primordia fully enclose floral bud; DE18, initiation of stamen primordia; DE19, development of awns on petals; DE20, initiation of gynoecium (carpel primordia). Stage V – inflorescence branch elongation and bract expansion: DE21, elongation of inflorescence branches; DE22, additional pedicel elongation during anthesis; DE23, expansion of bracts; DE24, floral bud open exposing mature stamen and release pollen.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

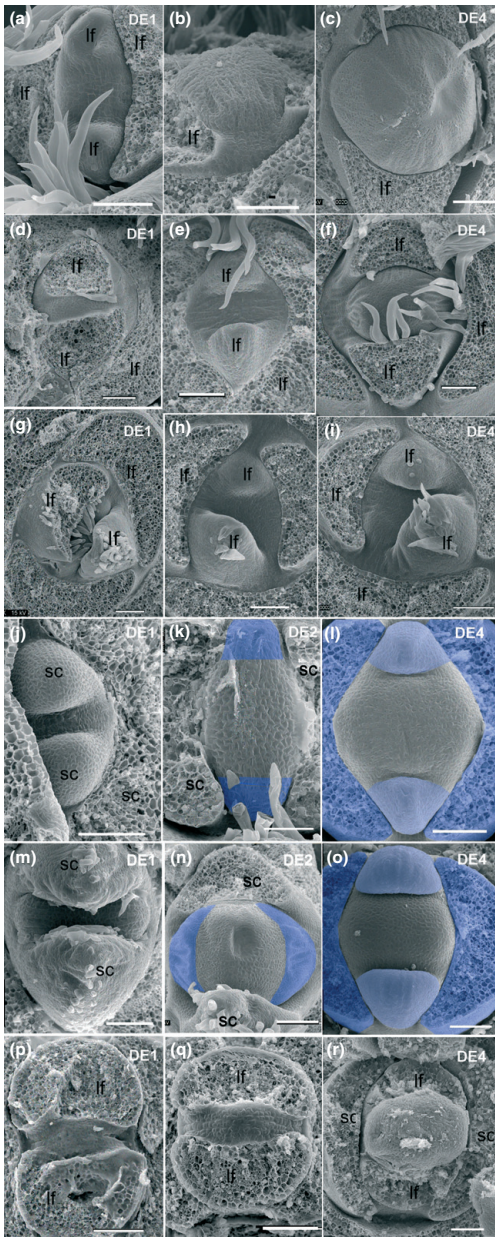


Fig. 3 Stage I of the inflorescence developmental pathway: vegetative meristem to inflorescence meristem transition (DE1 to DE4). (a–c) *Cornus sanguinea*; (d–f) *C. macrophylla*; (g–i) *C. controversa*; (j–l) *C. officinallis*; (m–o) *C. florida*; (p–r) *C. canadensis*. Blue colored area indicates bracts. lf, leaf; SC, scale. For a description of DE1–DE4, see Fig. 2. Bars, 100 μm.

Stage I: transition from vegetative meristem to IM – DE1 to DE4

The first stage in the inflorescence developmental pathway is the transition from a vegetative meristem (DE1) to an IM (DE4). During the vegetative to reproductive transition, the vegetative meristem enlarges and becomes more dome-shaped as it transitions into an IM that generates IBMs. In some, but not all, of the species we examined, the initiation of two pairs of bract primordia (DE2 and DE3) occurred during the vegetative to reproductive transition. The vegetative meristems of the *Cornus* species examined in this study are usually narrow and flat-topped (Figs 2, 3a,d,g,j,m,p, 4a,c,e,g). However, the morphology of the IMs varies among the different inflorescence types. They differ in size and in the organization of the meristematic cell layers (Figs 3–5).

In general, the IMs from species with paniculate cymes (*C. macrophylla*, *C. sanguinea* and *C. controversa*) are significantly larger than those in species from other three groups (Figs 3c,f,b, 5; $P < 0.01$; Tables 1, 2). In addition, meristematic cells in these species are more densely packed and the IMs have more corpus or mantle layers (Fig. 4b) than the condensed forms in *C. officinallis* (umbels, Fig. 4d) and *C. florida* (heads, Fig. 4f).

Another difference noted between the different species during the vegetative to reproductive transition is whether bract primordia initiate during the transition. In *C. florida* and *C. officinallis*, two pairs of bract primordia arise during the vegetative–reproductive transition, so they form outside of (abaxial to) the IM (Fig. 3k,l,n,o; Tables 1, 2). By contrast, in the branched forms (*C. canadensis* and the BW species), pairs of bracts initiate within the IMs subtending IBMs and initiate after the vegetative–reproductive transition (Fig. 6a,e,i,u; Tables 1, 2).

Stage II: initiation of primary IBMs from IMs – DE5 to DE9

After the transition from vegetative to reproductive development, the IMs begin to initiate pairs of primary IBMs (1^0 IBMs). In this paper, we define a 1^0 IBM as any lateral branch meristem that arises from the periphery of the apical IM. The 1^0 IBMs have the ability to generate additional lateral meristems: either higher-order IBMs (i.e. secondary, tertiary, etc.) or floral meristems. In species that display paniculate cymes (*C. sanguinea* and *C. macrophylla*), the apical IMs give rise to four pairs of 1^0 IBMs (Figs 2, 6; DE 5, 6, 8, and 9; each pair was color-coded as purple, salmon, orange, and yellow, respectively). These four pairs of 1^0 IBMs arise sequentially at the periphery of the meristematic dome in a decussate fashion. Concomitantly there is an apically directed elongation of the rachis supporting the IM (Figs 2, 6, 7; DE 7). This initiation pattern of 1^0 IBMs

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

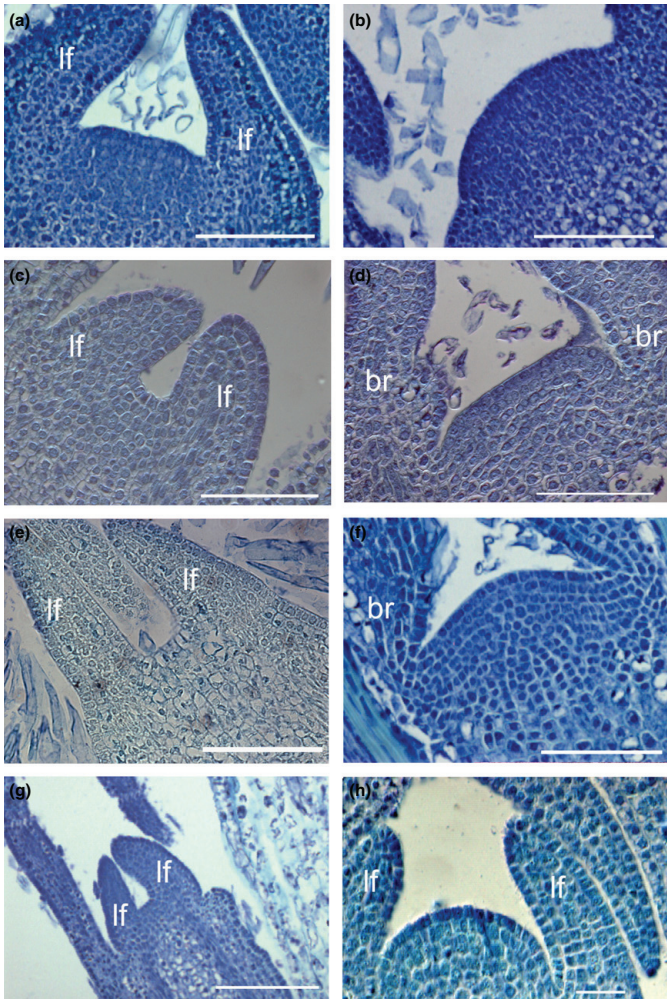


Fig. 4 Longitudinal sections of vegetative meristems and inflorescence meristems in four species of *Cornus*. (a) Vegetative meristem of *C. sanguinea*; (b) inflorescence meristem of *C. sanguinea*; (c) vegetative meristem of *C. officinallis*; (d) inflorescence meristem of *C. officinallis*; (e) vegetative meristem of *C. florida*; (f) inflorescence meristem of *C. florida*; (g) vegetative meristem of *C. canadensis*; (h) inflorescence meristem of *C. canadensis*. lf, leaf; br, bract; Bars: (a–f) 100 μm; (g) 200 μm; (h) 50 μm.

matches the decussate phyllotaxy of leaves that occurs almost exclusively in the genus (two sister species *C. controversa* Hemsley and *C. alternifolia* L.f. are exceptions). In *C. controversa*, the initiation pattern of 1^0 IBMs is spiral and matches the spiral leaf phyllotaxy observed in this species. In *C. controversa* the first three groups of 1^0 IBMs arise singularly and spirally from the periphery of the relatively broad apical IM (Fig. 6i,j,k). The rachis subtending the apical IM continues to elongate and the apical meristem generates another two branch meristems to give rise to a total of five 1^0 IBMs (Fig. 6l).

The apical IMs in *C. florida* and *C. canadensis* generate 1^0 IBMs with a similar pattern to those observed in *C. sanguinea* and *C. macrophylla* (decussate), but with a reduced number of 1^0 IBMs. In *C. florida*, three pairs were generated (Fig. 6s,t), and in *C. canadensis* only two pairs were produced (Fig. 6v,w,x). No elongation of the rachis supporting the IM was observed in these two species (Figs 6q–v, 7g–j).

The species that produces the umbel inflorescence (*C. officinallis*) shows a different initiation pattern of 1^0 IBMs. Unlike the sequential initiation pattern in other species, the first three groups of 1^0 IBMs (containing four, two and two IBMs per group, respectively) initiate almost simultaneously on the surface of the wide apical IM (Fig. 6m,n). Thus what may be a very compressed decussate initiation pattern of 1^0 IBM appears somewhat like a whorled pattern of initiation. Following the initiation of these eight 1^0

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

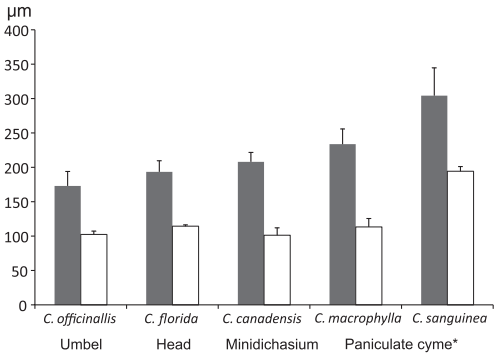


Fig. 5 Comparison of average width (closed bars) and height (open bars) of mature inflorescence among five species of *Cornus*. Significant difference: *, $P < 0.01$. Error bars are ±SD.

IBMs, another pair of 1^0 IBMs was formed from the center of apical IM (Fig. 6o,p) resulting in a total of 10 1^0 IBMs. Sometimes the second group 1^0 IBMs were observed to split into two IBMs and the fourth group of 1^0 IBMs could be aborted (Fig. 8).

Stage III: initiation of higher-order IBMs – DE10 to DE13

Each pair of 1^0 IBMs can give rise to a pair of secondary IBMs (2^0 IBMs, DE10) that initiate from peripheral portions of the primary IBMs. Secondary IBMs can lose IBM identity and differentiate into a floral meristem as observed in *C. officinallis* (Fig. 9j–l), for the third group of 1^0 IBMs of *C. florida* (Fig. 9n,o) and the most distal pair of 1^0 IBMs in *C. sanguinea* (Fig. 9a, c). Alternatively, the 2^0 IBMs can generate tertiary IBMs (3^0 IBMs, DE11). These 3^0 IBMs terminate as floral meristems in *C. canadensis* (Fig. 9q) and in most distal pair of primary branch meristems in *C. macrophylla* (Fig. 9e), or they can generate quaternary, quinary (fifth-order) or senary (sixth-order) IBMs (DE12) as in *C. macrophylla* and *C. controversa*. Senary IBMs are the highest order observed in the *Cornus* species we examined. All these high-order IBMs are arranged in dichasium structures. They initiate in a basipetal manner, starting from the most basal primary IBMs and continuing to more distal ones. The 1^0 IBMs in more distal positions produce fewer high-order IBMs. The IBMs in the BW species are elongated while only a slight elongation of IBMs was observed in *C. canadensis* (Fig. 7c,h), and no elongation of IBMs was observed in *C. officinallis* and *C. florida* (Fig. 7e,f,i,j).

Stage IV: floral organogenesis – DE14 to DE20

The central apical IM is the first to transition to a floral meristem in all species examined, as demonstrated by the initiation of sepal primordia (Fig. 9, DE 13). Flowers in the *Cornus* typically have tetramerous sepals, petals and stamens, two fused carpels and an inferior ovary (Fig. 10). Floral development is similar among different *Cornus* species. In all six species examined, four sepal primordia are the first floral organs to initiate at the margin of the floral meristem (DE15; Fig. 11a). Four petal primordia subsequently initiate adaxial to the sepal primordia to form the second floral whorl alternate with the sepal primordia (DE16; Fig. 11a). The petal primordia grow longer than sepals and grow to enclose the developing stamens and carpels (DE17; Fig. 11d). Stamen primordia initiate in the third floral whorl (DE18; Fig. 11e–g) followed by two carpel primordia from the remaining floral meristem apex (DE20; Fig. 11j).

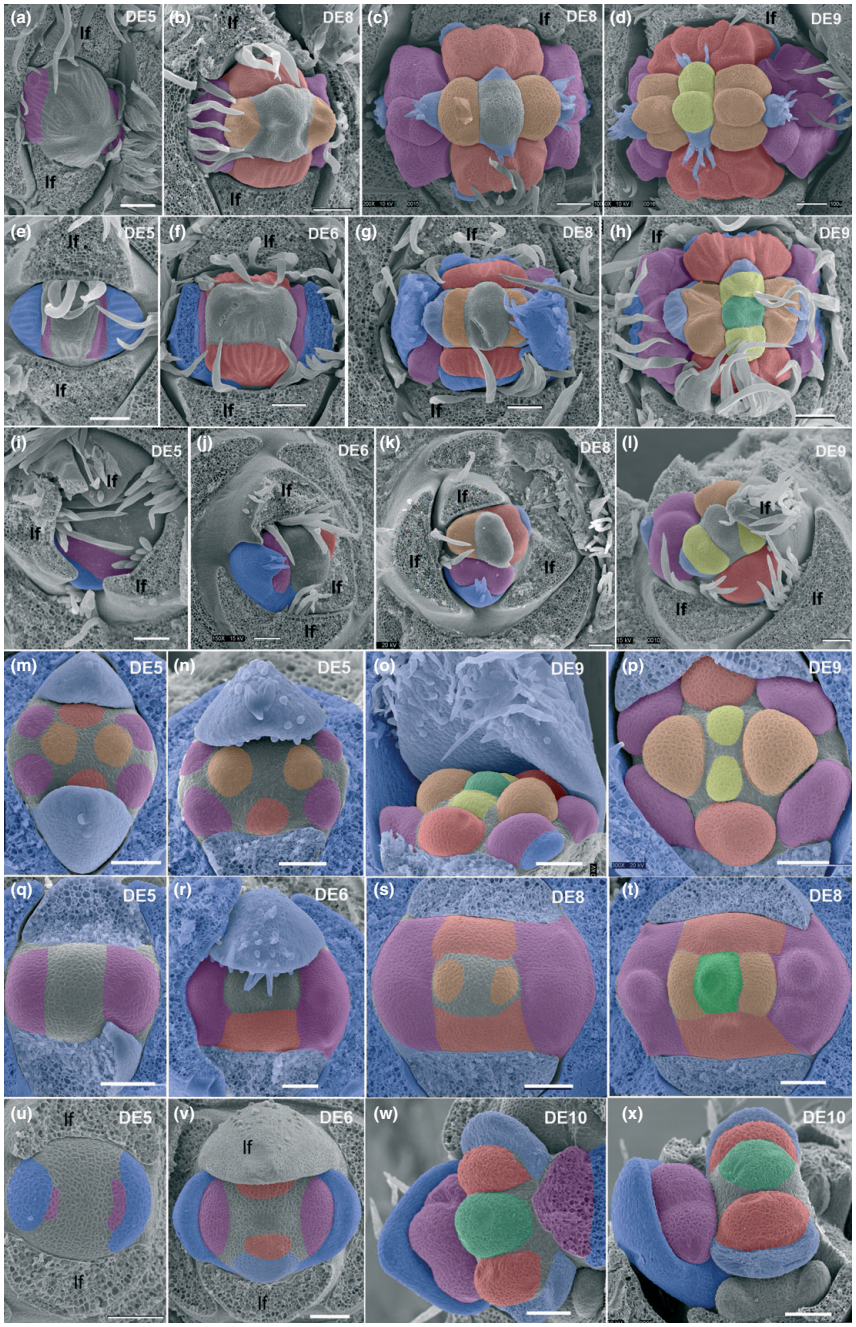
There are major differences of floral development among the *Cornus* species regarding the timing of floral organ initiation with respect to the calendar year (Fig. 2) and in the development of pedicels (DE14; Fig. 7d,f–h,k,l). In three species with condensed inflorescences, *C. florida*, *C. officinallis* and *C. canadensis*, floral organogenesis starts in the summer and all floral organs are initiated and well developed before winter. In two species with elongated inflorescences, *C. sanguinea* and *C. controversa*, floral organogenesis begins in the summer, but carpels do not develop until the following spring. In the third species with elongated inflorescences, *C. macrophylla*, which blooms later than *C. sanguinea* and *C. controversa*, floral organ organogenesis initiates in spring (Fig. 2).

Furthermore, pedicel development/elongation in *C. officinallis* (umbels) and *C. canadensis* (minidichasium) occurs during floral organogenesis (in the fall) (DE14; Figs 2, 7f–h). However, in species with elongated inflorescences, pedicel development/elongation occurs in the blooming season of the following spring, well after all floral organs are initiated (Fig. 12a). In *C. florida* (heads), flowers are sessile and pedicel development/elongation was not observed. In *C. canadensis*, awns (a unique feature of the DW group) develop on petals during this stage (DE19; Fig. 11h,i).

Stage V: brach elongation and bract expansion – DE21 to DE24

Species of the BW groups (*C. sanguinea*, *C. macrophylla* and *C. controversa* in this study) have a rapid increase in the length of branches and pedicels in the spring (DE21, 22; Figs 2, 12a), resulting in an expanded compound inflorescence. Longitudinal sections from developing branches at different stages indicated that the elongation of branches is supported by an increase of cell numbers in the pith and an increase in cell expansion in the cortex (Fig. 13). A minor elongation of the primary inflorescence branches and pedicels occurs in *C. canadensis*, but not in *C. florida*, in the

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust



© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

Fig. 6 Stage II of the inflorescence developmental pathway: initiation of primary inflorescence branch meristem (1^{O} IBM, DE5–DE9). (a–d) *Cornus sanguinea*; (e–h) *C. macrophylla*; (i–l) *C. controversa*; (m–p) *C. officinallis*; (q–t) *C. florida*; (u–x) *C. canadensis*; Blue, bracts; purple, first group of 1^{O} IBMs; salmon, second group of 1^{O} IBMs; orange, third group of 1^{O} IBMs; yellow, fourth group of 1^{O} IBMs; green, central floral meristem. lf, leaf; sc, scales; for DE5–DE9, see Fig. 2. Bars, 100 μm.

spring before flowering. In *C. officinallis*, only pedicels elongate; branches remain rudimentary (DE22; Fig. 12b). In *C. canadensis*, *C. florida* and *C. officinallis*, involucral bracts expanded, but became petaloid only in the first two species (Fig. 2, DE23; Fig. 12c,d).

Summary of inflorescence development pathways in *Cornus*

We identified 13 developmental events that differ among the species representing the four major inflorescence types we studied (Table 1; Fig. 2). The umbels and heads showed a combination of two unique developmental differences, respectively, while the minidichasium in *C. canadensis* display two unique differences and the elongated compound inflorescences in the BW groups showed four unique differences (Table 1; Fig. 2).

In addition, the relationship of the developmental events to the calendar year also varies among these species, especially for those in stages I, IV and V.

Reconstructing the developmental evolutionary history of umbels and heads in *Cornus*

The character matrix for the seven characters that are likely involved in the origin of umbels and heads is shown in

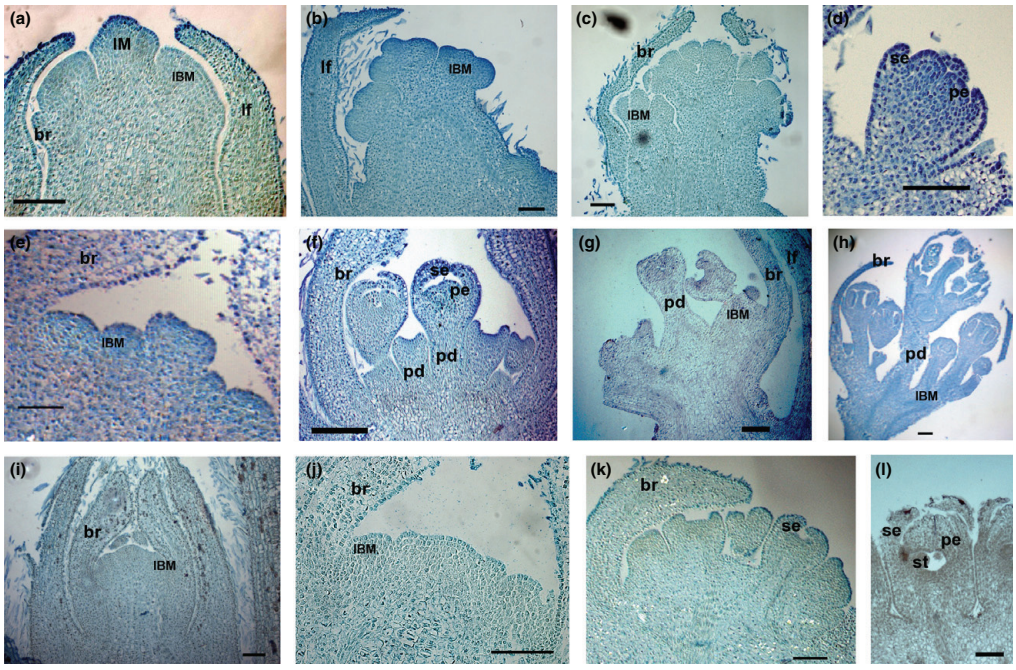


Fig. 7 Longitudinal sections of young *Cornus* inflorescences. (a–c) Longitudinal section of *C. macrophylla* inflorescences, showing vertical elongation of central rachis supporting the inflorescence meristem (IM) and branch elongation in branches supporting inflorescence branch meristems (IBMs); (d) longitudinal section of one developing flower of *C. macrophylla*, showing no pedicel development during floral organogenesis; (e, f) longitudinal section of *C. officinallis* young inflorescences, showing no elongation of rachis supporting the central IM or branches supporting IBMs; (e) pedicel development in *C. officinallis* during floral organogenesis (f); (g, h) longitudinal section of *C. canadensis* young inflorescences, showing slight elongation of branches supporting IBMs and pedicel development during floral organogenesis; (i–l) longitudinal section of *C. florida* young inflorescences, showing no elongation of central rachis supporting the IM and branches supporting the IBMs and no pedicel development during floral organogenesis. lf, leaf; br, bract; sc, scales; pd, pedicel; se, sepal; pe, petal; st, stamen. Bars: (a–k) 100 μm; (l) 50 μm.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

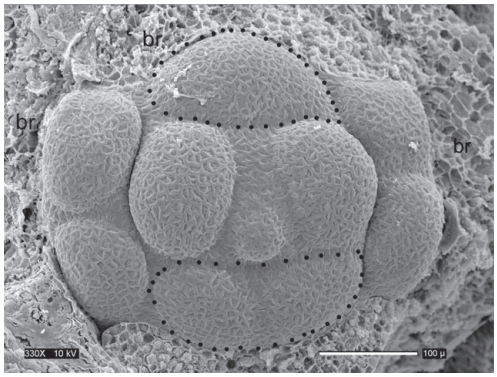


Fig. 8 Variable splitting of primary inflorescence branch meristems (1^{o} IBMs) in *Cornus officinallis*. Dotted line (lower) shows the splitting of one 1^{o} IBM into two 1^{o} IBMs in *C. officinallis*. br, bract. Bar, 100 µm.

Table 2. Results from ancestral character state reconstruction using BayesTraits were largely congruent with those inferred from ML and maximum parsimony (MP) methods (Fig. 14a–d; Supporting Information, Fig. S1), except for character A at node B (Fig. S1a). For this node, the BayesTraits suggested nearly equal possibility of state 0 and state 1, while ML prefers state 1. When comparing results between BayesTraits, ML, and MP methods, we found that ancestral character states with uncertainty were usually better resolved in BayesTraits (node C in Fig. S1a; nodes A, B, D in Fig. S1b; node A in Fig. 14a; nodes D, F in Fig. 14b; nodes D, F in Fig. 14c). Results from BayesTraits based on 1000 phylogenetic trees are described in the following.

Character A – size of IMs The ancestral state of the crown group of *Cornus* (Fig. S1a, node B) was not clear, but state 1 (small IM 1) is strongly supported for the most recent common ancestor of BB, DW and CC groups (node D, PP (posterior probability) = 0.99 ± 0.03), and state 0 (large IM) for the ancestor of BW group (node C, PP = 1.00 ± 0.01). The results supported divergence of the IM size between the ancestor of the BB, DW, and CC lineage and the ancestor of the BW lineage.

Character B – elongation of rachis supporting IMs Ancestral states at two deepest nodes (A and B in Fig. S1b) were resolved as state 0 (visible elongation) with high support (PP = 0.94 ± 0.10 and 0.81 ± 0.28, respectively). A switch to state 1 (invisible elongation) was highly supported at node D (common ancestor of BB-CC-DW groups, PP = 0.99 ± 0.04).

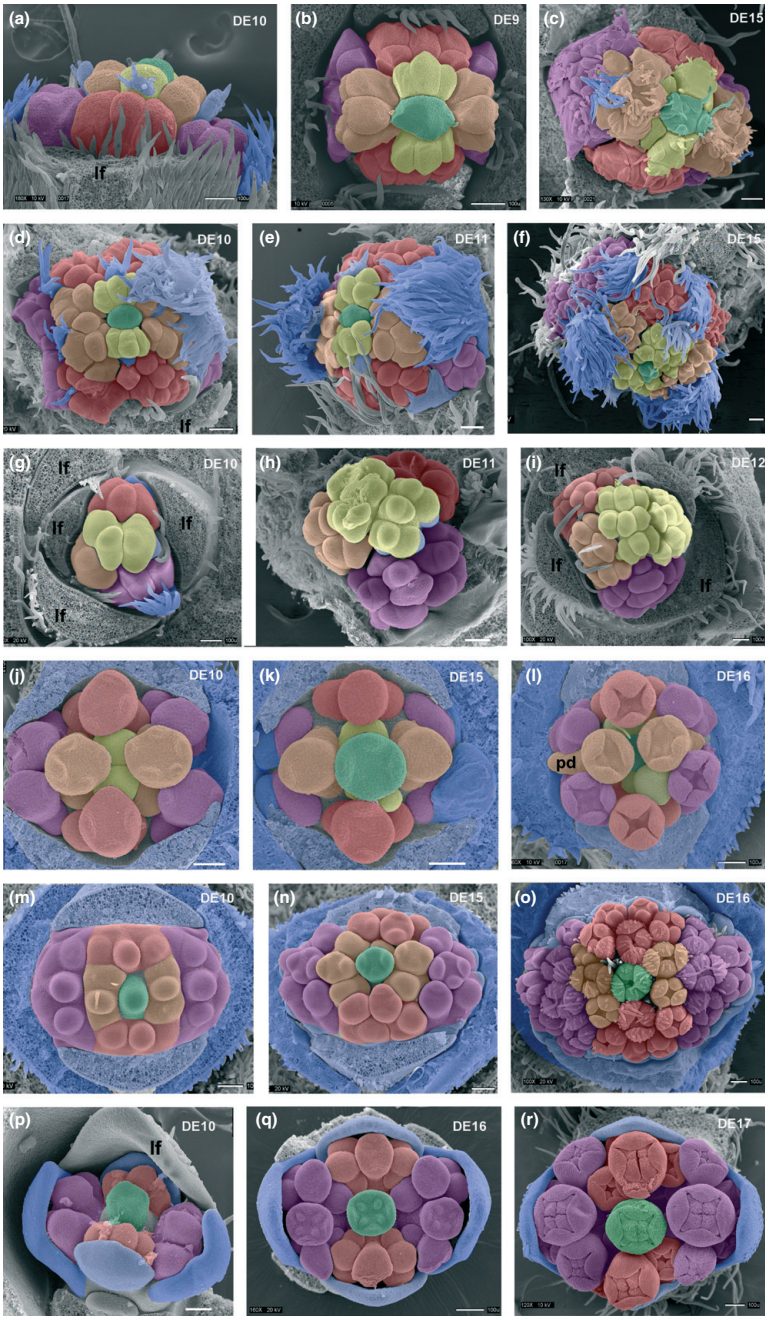
Character C – initiation pattern of 1^{0} IBMs The ancestral states of the root of *Cornus* (node B) and the most recent common ancestor of the BB-DW-CC clade (node D) were strongly supported to be 0 (PP = 0.99 ± 0.02 and 0.80 ± 0.23, respectively, Fig. S1c). There was a single evolutionary shift to state 1 (whorled and simultaneous) at the crown node of the CC lineage (node E, PP =1.00 ± 0.00).

Character D – number of 1^{0} IBMs State 0 (four pairs) was reconstructed for node B (crown of *Cornus*), node C (crown of the BW group) and node E (crown of the CC group) with strong support (PP = 0.95 ± 0.10, 0.93 ± 0.09 and 0.99 ± 0.01, respectively) and for node D (crown of BB-DW-BW clade) with relatively low, but significant support (PP = 0.60 ± 0.21) (Fig. 14a). There was a change in the number of 1^{o} IBMs at the node L (uniting *C. controversa* and *C. alternifolia*) within the BW group, from eight in four pairs to five or six 1^{0} IBMs (state 1). Evolutionary reduction in numbers occurred in the BB-DW clade. One scenario suggested a reduction from four pairs to three pairs (state 2) at the crown of the BB-DW clade (node F; PP: 0.48 ± 0.21), followed by a further reduction from three pairs to two pairs (state 3) at the crown node of DW (node G). The alternative scenario suggested a reduction from four pairs to two pairs at the node F followed by an increase from two pairs to three pairs at the crown node of the BB clade (node F; PP = 0.26 ± 0.18). The latter was supported by lower posterior probability, but the hypothesis test using Bayes Factor did not show a significant difference between these two scenarios.

Character E – IBM elongation The ancestral states at nodes A and B were strongly supported as state 0 with elongated IBMs (PP = 0.94 ± 0.08 and 0.85 ± 0.17, respectively, Fig. 14b). The ancestors of the BB-DW-CC and BB-DW clades were also estimated to have elongated IBMs although with relatively low (but significant) supports (PP = 0.55 ± 0.48* and 0.70 ± 0.32, respectively). A switch from elongated IBM (state 0) to nonelongated IBM (state 1) occurred independently in the ancestor of the CC (node E) and BB (node H) clades.

Fig. 9 Stage III of the inflorescence developmental pathway: initiation of higher-order inflorescence branch meristems (IBMs) (DE10–DE12). (a–c) *Cornus sanguinea*; (d–f) *C. macrophylla*; (g–i) *C. controversa*; (j–l) *C. officinallis*; (m–o) *C. florida*; (p–r) *C. canadensis*. Blue, bracts; purple, first group of primary (1^{o}) IBMs or flowers developed from the first group of 1^{o} IBMs; salmon, second group of 1^{o} IBMs or flowers developed from the second group of 1^{o} IBMs; orange, third group of 1^{o} IBMs or flowers developed from the third group of 1^{o} IBMs; yellow, fourth group of 1^{o} IBMs or flowers developed from the fourth group of 1^{o} IBMs; green, central floral meristem or central flower; lf, leaf; pd, pedicel. For a description of DE10–DE12, see Fig. 2. Bars, 100 µm.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust



© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

New Phytologist

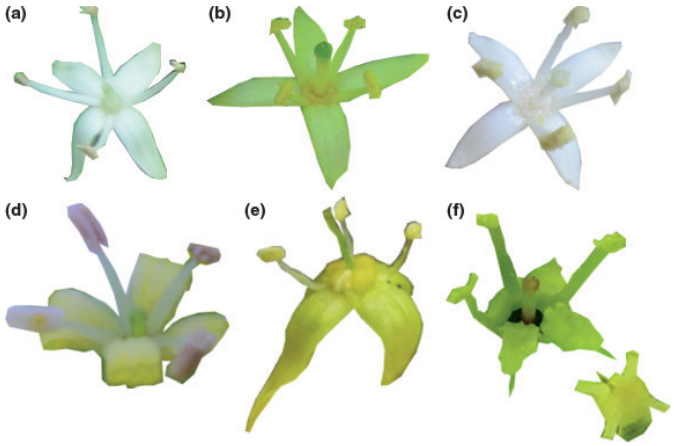


Fig. 10 Mature flowers in bloom from the six species of *Cornus* studied. (a) *C. macrophylla*, (b) *C. sanguinea*, (c) *C. controversa*, (d) *C. florida*, (e) *C. officinallis*, (f) *C. canadensis*.

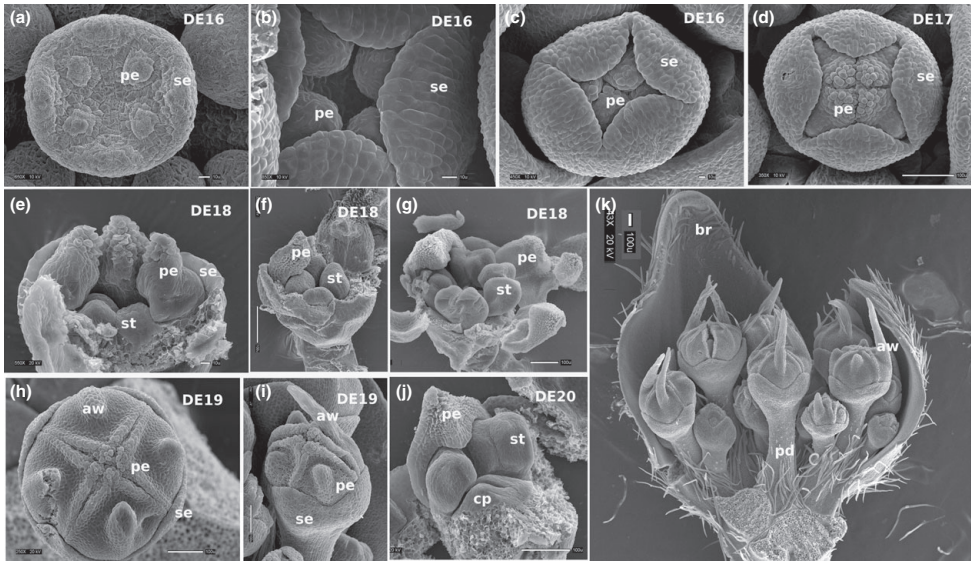


Fig. 11 Stage IV of the inflorescence developmental pathway: floral organogenesis (DE13–DE20), exemplified by *Cornus canadensis*. (a) Initiation of sepals and petals; (b, c) growth of sepals and petals; (d) petals covering the central meristem; (e) initiation of stamens; (f, g) growth of stamens; (h) initiation of awns on petals; (i) elongation of awns; (j) initiation of carpels; (k) mature inflorescence buds (two branches were taken off). For descriptions of DE13–DE20, see Fig. 2. br, bract; se, sepal; pe, petal; st, stamen; aw, awn; cp, carpel. Bars: (a–c, e) 10 μm; (d, f–k) 100 μm.

Character F – pedicel initiation during floral organogenesis The ancestral states were resolved as state 0 (no initiation during floral organogenesis) for all nodes with strong support, except for the CC and DW clades (Fig. 14c). The transition from state 0 to state 1 (initiation during floral organogenesis) was shown to have occurred independently in the CC and DW clades with high support (Fig. 14c, nodes E and G).

Character G – pedicel initiation during anthesis Pedicel initiation during anthesis (state 0) was resolved for the crown nodes of *Cornus* and the BW group with strong

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

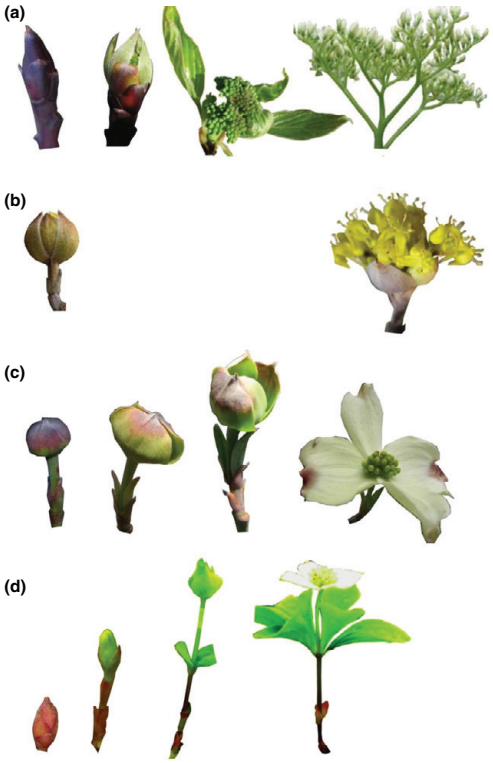


Fig. 12 Stage V of the inflorescence developmental pathway: branch elongation and bract expansion (DE 21–DE24). (a) Branch and pedicel elongation in *Cornus controversa*; (b) pedicel elongation in *C. officinallis*; (c) bract expansion and petaloidy in *C. florida*; (d) bract expansion and petaloidy in *C. canadensis*. For descriptions of DE21–DE24, see Fig. 2.

support (Fig. 14d, PP = 0.81 ± 0.28 and 1.00 ± 0.00, respectively). Loss of pedicel initiation during anthesis (state 1) occurred in the common ancestor of the BB-DW-CC clade (node D, PP = 0.99 ± 0.04).

These evolutionary changes are summarized in Fig. 15. Based on the evolutionary changes of the branches, the ancestral inflorescence type reconstructed for the BB-DW-CC clade was a small umbellate dichasium based on the following character states: small IM size (character A), loss of rachis elongation (character B), decussate and sequential initiation of 1^0 IBMs (character C), four pairs of 1^0 IBMs (character D), visible elongation of IBMs (character E), no initiation of pedicels during floral organogenesis (character F), and no pedicel initiation at the anthesis (character G).

Discussion

Temporal divergence of inflorescence developmental stages among *Cornus* species – new insights for BW and DW species

Our comparative developmental study revealed that all inflorescence types in *Cornus* initiate their development soon after the anthesis and complete the architecture formation in the autumn. This finding is new to the DW and BW lineages that produce elongated, branched inflorescences (minidichasium and paniculate cymes, respectively). The inflorescences in these two lineages had been, for a long time, believed to develop in the spring of the year of blooming, because their floral buds look similar to the leaf buds until they expand in the spring, after leaves have fully developed. Only the umbels and heads in the CC and BB lineages were considered to be preformed in the autumn of the previous year because they are clearly visible as head-like buds in the autumn, different from leaf bud morphology (Murrell, 1993; Xiang *et al.*, 2006). This temporal similarity in a broader scale among the inflorescence types indicates that the development of the branched inflorescence types in DW and BW clades is not as delayed as it appears to be based on the outer morphology.

We have also characterized at a finer scale the temporal divergence with respect to the calendar year of the developmental events among the species (Table 1; Fig. 2). Evolutionary shifts towards earlier inflorescence development and blooming occurred in the umbel (*C. officinallis*)

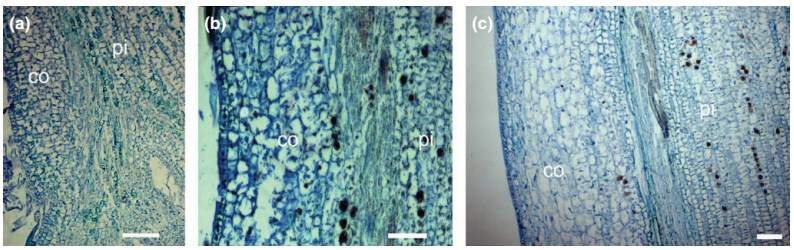


Fig. 13 Longitudinal section of inflorescence branches during branch elongation in *Cornus macrophylla*. (a) Inflorescence branches about to elongate; (b) elongating inflorescence branches; (c) elongated inflorescence branches. co, corpus; pi, pith; Bars, 100 μm.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

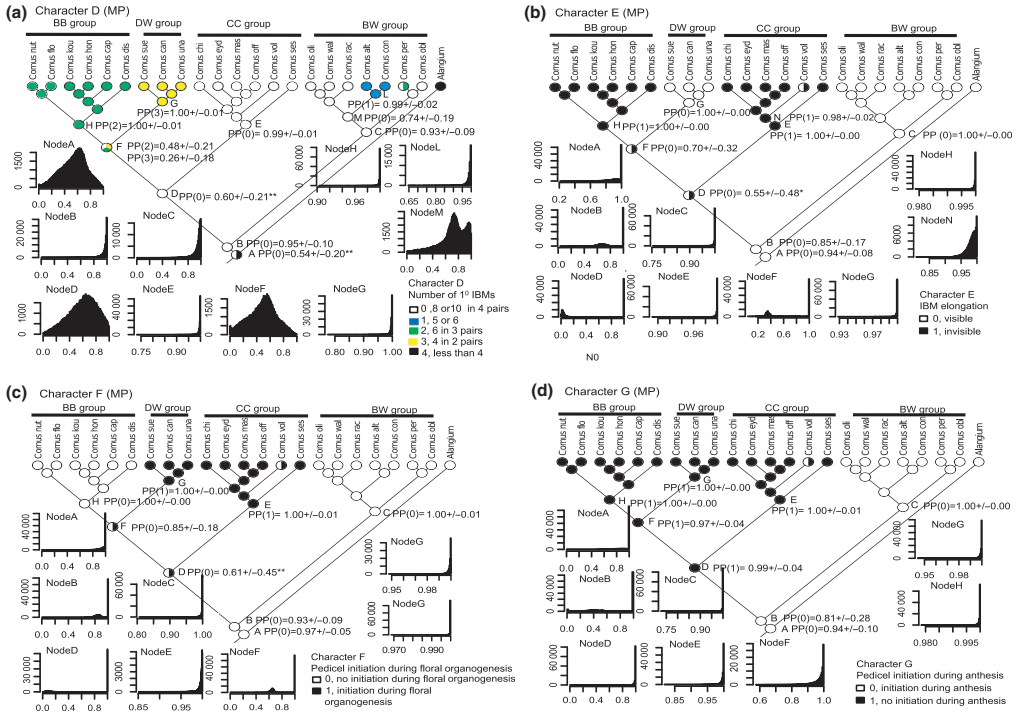


Fig. 14 Evolutionary trends of four of the seven developmental characters (characters D–G) important to the development of umbels and heads in *Cornus*. Results for the first three characters (A–C) are presented in Supporting Information, Fig. S1. Character states shown in ball patterns (or colors) were derived from maximum parsimony analyses implemented in Mesquite 2.01 using the best phylogenetic tree (*Diplopanax* was removed manually from the outgroup to save space). Maximum likelihood analyses of these characters were not possible because of the character state polymorphism in some taxa. Character state and its posterior probability derived from analysis using a Bayesian method implemented in BayesTraits 1.0 using 1000 phylogenetic trees were shown at nodes. The distribution densities of posterior probabilities of character states estimated for the nodes of interest by BayesTraits are shown in the panels. For ancestral states with posterior probabilities < 0.7, a hypothesis test using Bayes Factors was performed to determine whether there is support for one state over the other state. *, Bayes factor is > 2, positive support for one over the other state; **, Bayes factor is > 5, strongly support one character state over the other; (a) character D; (b) character E; (c) character F; (d) character G; details of characters and states are provided in Tables 1, 2. PP, posterior probability.

and head (*C. florida*) lineage (Figs 2, 16). The link between the temporal shift in flowering and the origins of umbels and heads is unclear. It was proposed that inflorescence architectures affect pollinator behaviors and influence pollination efficiency (Wyatt, 1982). Observation in nature and experimental studies on artificial inflorescence architectures have shown that the flat and closely packed structures of umbels and heads increased the number of flowers visited by insects or birds per inflorescence as a result of the decrease in foraging cost in these flattened or planar inflorescence structures (Hainsworth *et al.*, 1983; Cresswell, 1990; Jordan & Harder, 2006). By increasing floral visitation, umbels and heads could ensure a certain rate of pollination in early spring when the number of pollinators is limited. Therefore, selection may favor condensed inflorescence structures for early flowering plants or for plants occurring in places with scarce pollinators (e.g. in alpine habitat, as for the DW species). The evolution of condensed inflorescence structures in *Cornus* (e.g. umbels, heads, and minidichasium) could thus be the results of adaptation to early flowering or scarce pollinators. Inflorescence structures were also found to be correlated with types of pollinators in Arecaceae (Henderson, 2002). In this family, the condensed inflorescences were reported to be pollinated by beetles while the elongated inflorescences were found to be pollinated by bees, flies or wasps (Henderson, 2002). The dogwood species were reported to be pollinated by a wide range of insects (Eyde, 1988; Rhoades, 2010). Species producing condensed inflorescences (BB, DW, and CC groups) are pollinated by bees, flies, and beetles (Douglas, 1983; Barrett & Helenurm, 1987; Eyde, 1988; Rhoades, 2010), while species producing elongated,

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

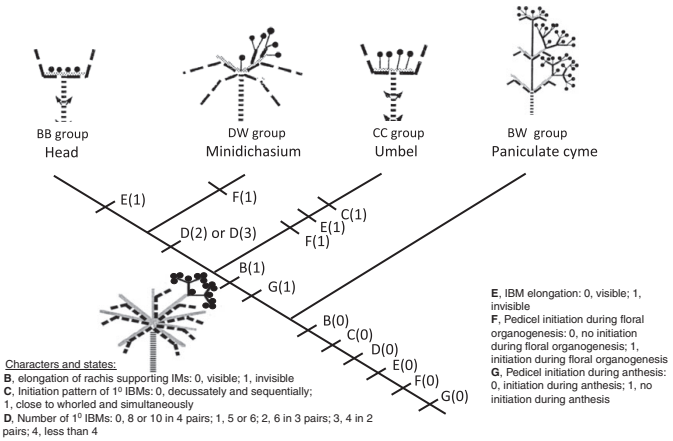


Fig. 15 Summary of inferred evolutionary developmental changes responsible for the origins of umbels and heads in *Cornus.* Ticks on branches indicate the evolutionary events (character A was not shown as the ancestral state, as the crown of *Cornus* was unclear). Numbers in parenthesis indicate character states. Schematic diagrams of inflorescence structures are reproduced from Murrell (1993, 1996) with permission from American Society of Plant Taxonomists. Detailed information about schematic diagrams is given in Fig. 1. BB, big-bracted; DW, dwarf dogwood; CC, cornelian cherry; BW, blue- or white-fruited.

branched inflorescences (the BW group) were reported to be pollinated by flies and beetles (Robertson, 1928; Parmenter, 1956; Waldbauer, 1983). The available evidence is unclear regarding whether the evolution of condensed inflorescence structures in *Cornus* is linked to pollinator preference.

Developmental evolutionary history of umbels and heads in *Cornus*

By tracking the evolutionary histories of developmental characters on the *Cornus* phylogeny, we have begun to illuminate the developmental basis of the origins of umbels and heads in the genus. Our results support the idea that umbels and heads were derived from the ancestral umbellate dichasium independently, each transition involving multiple developmental changes at different evolutionary times (Fig. 15), a much more complicated evolutionary history than can be deciphered from analysis of the mature external morphology (Xiang & Thomas, 2008). The initial developmental divergences that led to umbel and head forms occurred very early in the evolutionary history of the genus, before the BB-DW split from CC, dating back to the very early Tertiary based on molecular dating and fossil evidence (Xiang *et al.*, 2008). The origin of umbels required at least five evolutionary changes. Two occurred earlier in the common ancestor of the BB-DW-CC clade: suppression of rachis elongation and loss of pedicel initiation/elongation at anthesis. Three changes occurred later in the CC lineage after it diverged from the BB-DW clade: a change from a decussate to a 'whorled' pattern of primary IM initiation, loss of IBM elongation, and initiation and elongation of pedicels during floral organ initiation (Fig. 15). It should be noted here that the 'whorled' pattern of primary IBM initiation could also be interpreted as a severe compression of the rachis elongation and reduction of the time between the initiations of successive IBM pairs. The origin of heads involved four evolutionary changes from the common ancestor of *Cornus*. The first two changes were common to those for umbels and the latter two changes occurred independently, one in the ancestor of BB-DW (reduction in the number of primary IBMs) and one in the BB clade (loss of IBM elongation, Fig. 15). These data suggested that the ancestor of the umbel and head forms in *Cornus* was a small eight- or seven-branched umbellate dichasium. This type of inflorescence is still observed in some extant species of *Cornus*, such as the male inflorescence of *C. volkensii*, a member of the CC group from tropical Africa, and *C. peruviana* in the BW group (Murrell, 1996). Our data also indicate that evolution of heads and umbels in *Cornus* did not occur through simple suppression of branches and pedicels as previously believed, but involved other important evolutionary changes that would not be apparent without a comparative analysis of the progression of the developmental morphologies. These new insights deepen our understanding of inflorescence evolution in *Cornus*. The evolutionary pathways of umbels and heads revealed in the study differ from, and are clearer than, those inferred from mature external morphology alone in *Cornus* (Xiang & Thomas, 2008).

Our study on *Cornus* is consistent with the hypothesis that inflorescence architectures evolved from elongated to condensed forms in angiosperms more generally (Parkin,

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

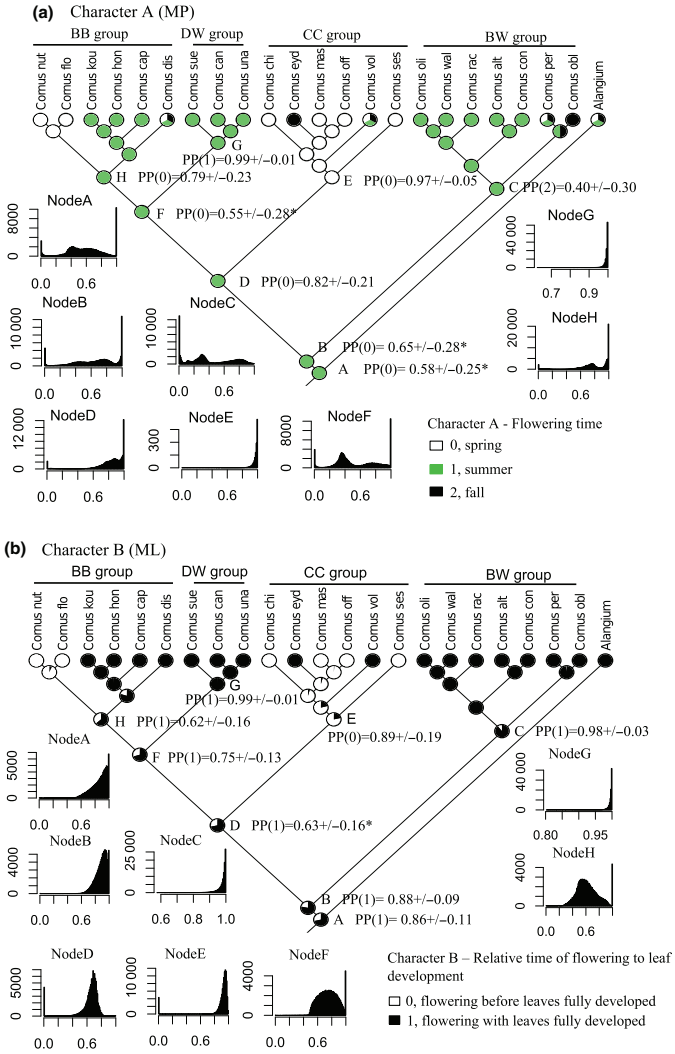


Fig. 16 Evolutionary trends of flowering time in *Cornus* inferred from analyses using BayesTraits 1.0 and Mesquite 2.01. Flowering time of species was determined based on herbarium specimen records, '*A California Flora*' (Munz & Keck, 1959), '*Manual of Vascular Plants of Northeastern United States and Adjacent Canada*' (Gleason & Cronquist, 1991), '*Floral of China*' (Xiang & Boufford, 2005), and several other articles (Murrell, 1996; Xiang *et al.*, 2003). Information on analyses and symbols are given in Fig. 14. PP, posterior probability.

1914; Stebbins, 1974; Wyatt, 1982; Harris, 1999). The data from *Cornus*, however, do not support the hypothesis that umbels evolved from elongated inflorescences via simple suppression of inflorescence branches and heads evolved from umbels. It is as yet unclear if the developmental pathways of umbels and heads found in *Cornus* are conserved in other plant lineages that also bear determinate umbels and heads. Comparative studies of developmental pathways of inflorescence architectures in other plant lineages would be helpful for a better understanding of the evolution of umbels and heads in a variety of Angiosperm species.

By contrast, developmental studies have been conducted for indeterminate umbels and heads in a few plants, for example, *Lotus* and *Coronilla* in Fabaceae (Dong *et al.*, 2005; Sokoloff *et al.*, 2007) and *Helianthus annuus* in Asteraceae (Marc & Palmer, 1981). These studies found

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

that, similar to the umbels and heads in *Cornus*, the IMs in these taxa also had a flat structure. The major developmental difference between these taxa and *Cornus*, however, is the lack of IBMs in these taxa. Initiation of IBMs was not observed in these taxa. Only floral meristem initiation was observed, which occurred in whorls directly from the expanded IM. This process was in contrast to the development of umbels and heads in *Cornus*, which formed IBMs first and subsequently suppressed their elongation. This suggests that the genetic controls in indeterminate heads and umbels are likely different from those in determinate heads and umbels. It further suggests that potentially fundamental differences exist in the formation of determinate and indeterminate umbels and heads.

Future work

Our phylogeny-based comparative developmental study clarifies the complex developmental pathways of the four inflorescence types in *Cornus* and identifies seven important characters differentiating the development of umbels and heads (Figs 14, 15, S1). These data suggest that the study of the genetic control of IM size, rachis elongation, IBM initiation and elongation, timing of pedicel formation and the transition of IM to floral meristem will be critical to understanding the evolution of umbels and head development in *Cornus*. We are currently working to establish a genetic transformation system for *Cornus canadensis* employing callus-regenerated plantlets (Feng *et al.*, 2009), which will allow us to test the function of candidate genes involved in these developmental processes and to clarify the molecular genetic basis for the evolution of these developmental differences.

In conclusion, our results revealed a more complex developmental basis underlying the evolution of umbels and heads in *Cornus* than is predicted from external morphology. Our findings shed novel insights into head and umbel evolution in the genus. Our work illustrates the importance of combining developmental and phylogenetic analyses to better understand morphological evolution.

Acknowledgements

The authors thank the J. C. Raulston Arboretum for providing living plants for experiments, the phytotron at North Carolina State University for caring cultured plants, D. E. Boufford for assistance in collecting *C. canadensis* from the field, M. Mackenzier Jr and V. Knowlton at NCSU Electron Microscopy Center for technical assistance with the SEM analysis, and the herbaria at NCSU and UNC, Chapel Hill for use of the *Cornus* collections. We acknowledge the travel awards from the Developmental & Structural Section of Botanical Society of America to C-M.F. The research was supported by a NCSU Multi-disciplinary Research grant and a NSF grant (IOS-1024629).

References

Alvarez J, Guli CL, Yu XH, Smyth DR. 1992. Terminal-flower – a gene affecting inflorescence development in *Arabidopsis thaliana*. *Plant Journal* **2**: 103–116.

Barrett SCH, Helenurm K. 1987. The reproductive biology of boreal forest herbs. I. Breeding systems and pollination. *Canadian Journal of Botany* **65**: 2036–2046.

Benlloch R, Berbel A, Serrano-Mislata A, Madueno F. 2007. Floral initiation and inflorescence architecture: a comparative view (vol 100, pg 659, 2007). *Annals of Botany* **100**: 659–676. 10.1093/aob/mcm293.

Blazquez MA, Soowal LN, Lee I, Weigel D. 1997. LEAFY expression and flower initiation in Arabidopsis. *Development* **124**: 3835–3844.

Bradley D, Ratcliffe O, Vincent C, Carpenter R, Coen E. 1997. Inflorescence commitment and architecture in Arabidopsis. *Science* **275**: 80–83.

Chuck G, Muszynski M, Kellogg E, Hake S, Schmidt RJ. 2002. The control of spikelet meristem identity by the branched silkless1 gene in maize. *Science* **298**: 1238–1241.

Coen ES, Romero JM, Doyle S, Elliott R, Murphy G, Carpenter R. 1990. Floricaula – a homeotic gene required for flower development in *Antirrhinum majus*. *Cell* **63**: 1311–1322.

Cresswell JE. 1990. How and why do nectar-foraging bumblebees initiate movements between inflorescences of Wild Bergamot *Monarda fistulosa* (Lamiaceae). *Oecologia* **82**: 450–460.

Cronk QCB, Bateman RM, Hawkins JA. 2002. *Developmental genetics and plant Evolution*. London, UK: New York, NY, USA: Taylor & Francis.

Dong ZC, Zhao Z, Liu CW, Luo JH, Yang J, Huang WH, Hu XH, Wang TL, Luo D. 2005. Floral patterning in *Lotus japonicus*. *Plant Physiology* **137**: 1272–1282.

Douglas S. 1983. Floral color patterns and pollinator attraction in a bog habitat. *Canadian Journal of Botany* **61**: 3494–3501.

Endress PK. 2010. Disentangling confusions in inflorescence morphology: patterns and diversity of reproductive shoot ramification in angiosperms. *Journal of Systematics and Evolution* **48**: 225–239.

Eyde RH. 1988. Comprehending *Cornus* – puzzles and progress in the systematics of the dogwoods. *Botanical Review* **54**: 233–351.

Fan CZ, Xiang QY. 2001. Phylogenetic relationships within *Cornus* (Cornaceae) based on 26S rDNA sequences. *American Journal of Botany* **88**: 1131–1138.

Fan CZ, Xiang QY. 2003. Phylogenetic analyses of Cornales based on 26S rRNA and combined 26S rDNA-matK-rbcL sequence data. *American Journal of Botany* **90**: 1357–1372.

Feng C, Qu R, Zhou L, Xie D, Xiang QY. 2009. Shoot regeneration of dwarf dogwood (*Cornus canadensis* L.) and morphological characterization of the regenerated plants. *Plant Cell Tissue and Organ Culture* **97**: 27–37.

Fishbein M, Venable DL. 1996. Evolution of inflorescence design: theory and data. *Evolution* **50**: 2165–2177.

Friedman J, Harder LD. 2005. Functional associations of floret and inflorescence traits among grass species. *American Journal of Botany* **92**: 1862–1870.

Gleason H, Cronquist A. 1991. *Manual of vascular plants of northeastern United States and adjacent Canada*. Bronx, NY, USA: New York Botanical Garden.

Hainsworth FR, Mercier T, Wolf LL. 1983. Floral arrangements and hummingbird feeding. *Oecologia* **58**: 225–229.

Harris EM. 1999. Capitula in the Asteridae: a widespread and varied phenomenon. *Botanical Review* **65**: 348–369.

Henderson A. 2002. *Evolution and ecology of palms*. Bronx, NY, USA: Botanical Garden Press.

Ingram GC, Goodrich J, Wilkinson MD, Simon R, Haughn GW, Coen ES. 1995. Parallels between unusual floral organs and fimbriata,

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

genes-controlling flower development in *Arabidopsis* and *Antirrhinum*. *Plant Cell* **7**: 1501–1510.

Jordan CY, Harder LD. 2006. Manipulation of bee behavior by inflorescence architecture and its consequences for plant mating. *The American Naturalist* **167**: 496–509.

Kellogg EA. 2004. Evolution of developmental traits. *Current Opinion in Plant Biology* **7**: 92–98.

Kellogg EA. 2006. Progress and challenges in studies of the evolution of development. *Journal of Experimental Botany*. **57**: 3505–3516.

Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyozuka J. 2003. FRIZZY PANICLE is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* **130**: 3841–3850.

Lee I, Wolfe DS, Nilsson O, Weigel D. 1997. A LEAFY co-regulator encoded by UNUSUAL FLORAL ORGANS. *Current Biology* **7**: 95–104.

Li F, Liu W, Tang J, Chen J, Tong H, Hu B, Li C, Fang J, Chen M, Chu C. 2010. Rice Dense and Erect Panicle 2 is essential for determining panicle outgrowth and elongation. *Cell Research* **20**: 838–849.

Li S, Qian Q, Fu Z, Zeng D, Meng X, Kyozuka J, Maekawa M, Zhu X, Zhang J, Li J *et al.* 2009. Short panicle1 encodes a putative PTR family transporter and determines rice panicle size. *Plant Journal* **58**: 592–605.

Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, Paran I, Eshed Y, Zamir D. 2008. The making of a compound inflorescence in tomato and related nightshades. *PLoS Biology* **6**: 2424–2435.

Maddison WP, Maddison DR. 2010. Mesquite: a modular system for evolutionary analysis [online]. Version 2.74. [WWW document]. URL http://mesquiteproject.org [accessed on 3 October 2010]

Marc J, Palmer JH. 1981. Photoperiodic sensitivity of inflorescence initiation and development in sunflower. *Field Crops Research* **4**: 155–164.

Munz PA, Keck DD. 1959. *A California flora*. Berkeley, CA, USA: University of California Press.

Murrell ZE. 1993. Phylogenetic relationships in *Cornus* (Cornaceae). *Systematic Botany* **18**: 469–495.

Murrell ZE. 1996. A new section of *Cornus* in South and Central America. *Systematic Botany* **21**: 273–288.

Pagel M, Meade A. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov Chain Monte Carlo. *The American Naturalist* **167**: 808–825.

Parkin J. 1914. The evolution of the inflorescence. *Botanical Journal of the Linnean Society* **42**: 511–563.

Parmenter L. 1956. Beetles visiting the flowers of dogwood. *Entomology Record of Journal Variation* **68**: 243–244.

Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E. 2007. Evolution and development of inflorescence architectures. *Science* **316**: 1452–1456.

Rebocho AB, Bliek M, Kusters E, Castel R, Procissi A, Roobeek I, Souer E, Koes R. 2008. Role of EVERGREEN in the development of the cymose petunia inflorescence. *Developmental Cell* **15**: 437–447.

Rhoades PR. 2010. *Four aspects of dogwood pollination: insect visitation, a novel approach to identity pollen, floral volatile emission and tracking parentage*. Master's thesis, University of Tennessee, Knoxville, TN, USA.

Robertson C. 1928. *Flowers and insects: lists of visitors of four hundred and fifty-three flowers*. Carlinville, IL, USA: Science Press Printing Company.

Samach A, Klenz JE, Kohalmi SE, Risseeuw E, Haughn GW, Crosby WL. 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.

Satoh-Nagasawa N, Nagasawa N, Malcomber S, Sakai H, Jackson D. 2006. A trehalose metabolic enzyme controls inflorescence architecture in maize. *Nature* **441**: 227–230.

Sokoloff DD, Degtjareva GV, Endress PK, Remizowa MV, Samigullin TH, Valiejo-Roman CM. 2007. Inflorescence and early flower development in loteae (leguminosae) in a phylogenetic and taxonomic context. *International Journal of Plant Sciences* **168**: 801–833.

Soltis DE, Soltis PS, Endress PK, Chase MW. 2005. *Phylogeny and evolution of Angiosperms*. Sunderland, MA, USA: Sinauer Associates, Incorporated.

Souer E, Rebocho AB, Bliek M, Kusters E, de Bruin RAM, Koes R. 2008. Patterning of inflorescences and flowers by the F-box protein double top and the leafy homolog aberrant leaf and flower of petunia. *Plant Cell* **20**: 2033–2048.

Souer E, van der Krol A, Kloos D, Spelt C, Bliek M, Mol J, Koes R. 1998. Genetic control of branching pattern and floral identity during Petunia inflorescence development. *Development* **125**: 733–742.

Stebbins GL. 1974. *Flowering plants: evolution above the species level*. Cambridge, MA, USA: Belknap Press.

Tucker SC, Grimes J. 1999. The inflorescence: introduction. *Botanical Review* **65**: 303–316.

Waldbauer GP. 1983. Flower associations of mimetic Syrphidae (Diptera) in northern Michigan. *Great Lakes Entomology* **16**: 79–85.

Weigel D, Alvarez J, Smyth DR, Yanofsky MF, Meyerowitz EM. 1992. Leafy controls floral meristem identity in Arabidopsis. *Cell* **69**: 843–859.

Wyatt R. 1982. Inflorescence architecture – how flower number, arrangement, and phenology affect pollination and fruit-set. *American Journal of Botany* **69**: 585–594.

Xiang QY, Boufford DE. 2005. Cornaceae. In: Wu ZY, Raven HR, Hong DY, eds. *Floral of China Volume 14 Apiaceae through Ericaceae*. Beijing, China: Science Press, St. Louis, MO, USA: Missouri Botanical Garden Press, 206–221.

Xiang QY, Shui YM, Murrell Z. 2003. *Cornus eydeana* (Cornaceae), a new cornelian cherry from China – notes on systematics and evolution. *Systematic Botany* **28**: 757–764.

Xiang QY, Soltis DE, Soltis PS. 1998. Phylogenetic relationships of cornaceae and close relatives inferred from matK and rbcL sequences. *American Journal of Botany* **85**: 285–297.

Xiang QY, Thomas DT. 2008. Tracking character evolution and biogeographic history through time in Cornaceae – does choice of methods matter? *Journal of Systematics and Evolution* **46**: 349–374.

Xiang QY, Thomas DT, Xiang QP. 2011. Resolving and dating the phylogeny of Cornales – effects of taxon sampling, data partitions, and fossil calibrations. *Molecular Phylogenetics and Evolution* **59**: 123–138.

Xiang QY, Thomas DT, Zhang WH, Manchester SR, Murrell Z. 2006. Species level phylogeny of the genus *Cornus* (Cornaceae) based on molecular and morphological evidence – implications for taxonomy and Tertiary intercontinental migration. *Taxon* **55**: 9–30.

Xiang QY, Thorne JL, Seo T, Zhang W, Thomas DT, Ricklefs RE. 2008. Rates of nucleotide substitution in Cornaceae (Cornales)-Pattern of variation and underlying causal factors. *Molecular Phylogenetics and Evolution* **49**: 327–342.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Evolutionary trends of the first three developmental characters (characters A–C; Tables 1, 2) important to the development of umbels and heads in *Cornus*.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

© 2011 The Authors
New Phytologist © 2011 New Phytologist Trust

Chapter IV

Alteration in selection, co-evolution, and expression of MADS-box B-class genes after gene duplication in dogwoods (*Cornus* s. l., Cornaceae) – Deciphering the genetic links to bract petaloidy evolution

Chun-Miao Feng[1], Xiang Liu[1], Yi Yu[1], Wenheng Zhang[2], Dr. Deyu Xie[1], Robert G. Franks[3] and Qiu-Yun (Jenny) Xiang[1]

[1]Department of Plant Biology, North Carolina State University, Raleigh, NC 27695,USA

[2]Department of Organism and Evolutionary Biology, Harvard University Herbaria, Cambridge, MA 02138, USA

[3]Department of Genetics, North Carolina State University, Raleigh, NC 27695, USA

Manuscript in preparation: follow the format of Molecular Biology and Evolution

Chunmiao did all the developmental study, examined gene expression patterns using *in situ* hybridization, and completed all the sequence analyses reported in this paper. Xiang Liu generated cDNA sequences of *PI* and *AP3* homologs of *Cornus* species. Xiang Liu and Yi Yu designed primers and performed RT-PCR for *PI* and *AP3* homologs of *Cornus* species. Wenheng Zhang generated partial genomic sequences of *AP3* homologs of *Cornus* species. Dr. Deyu Xie, Dr. Jenny Xiang and Dr. Robert Franks provided scientific advice and guidance.

Abstract

Transcription factor gene duplications followed by functional divergence have played an important role in shaping regulatory networks influencing key phenotypic characters. However, the mechanisms of evolutionary change after gene duplication leading to evolutionary novelty are still not well understood. B-class genes encode a group of MADS-box DNA-binding transcription factors controlling floral development in angiosperms. Extensive B-class gene duplication events have occurred during the evolution of angiosperms. Gene duplications followed by functional divergence have been hypothesized to be responsible for heterotopic petaloidy in organs beyond petals. Here we tested the hypothesis via investigation of B-class gene evolution and its relationship to bract petaloidy in the genus *Cornus*. We cloned the B-class genes in six species and reconstructed the evolutionary histories to identify gene duplication events, estimated the coevolution between the two B-class genes, and tested their roles in petaloid bract development in the genus via gene expression analyses. Our study reveals divergent developmental pathways of petaloid bracts and differential expression patterns of B-class genes in the two lineages with petaloid bracts. We report altered selection regimes on duplicated *PI* paralogs and demonstrate that coevolution between two B-class genes may affect the fate of duplicates. We find different partnerships between CorPI paralogs and CorAP3 in flower and petaloid bracts in *Cornus*. Our results suggested the importance of B-class protein C domains in functional divergence and provide new insights on B-class gene evolution and its potential link to the molecular mechanisms underlying the evolutionary novelty of bract petaloidy.

Keywords: *Cornus*, B-class gene, duplication, coevolution, bract petaloidy

Introduction:

Plant MADS-box genes encoding DNA-binding transcription factors play a crucial role in regulating flower development (Theissen et al. 1996, 2000; Munster et al. 2001; Stellari et al. 2004; Litt and Kramer 2010; Rijpkema et al. 2010; Sablowski 2010). Four functional classes of MADS-box homeotic genes controlling floral organ identity have been identified and their roles during organ identity specification have been summarized in the floral quartet model, where a heterotetramer including one A class protein, two B class proteins and one E class protein specify petal identity, a heterotetrameric complex composed of two B class proteins, one C class protein and one E class protein regulate stamen development, and a tetrameric protein complex composed of two C class proteins and two E class proteins is required for specification of carpels (Coen and Meyerowitz 1991; Pelaz et al. 2000; Honma and Goto 2001; Theissen and Saedler 2001). Gene duplication followed by functional divergence through shifts in gene expression and protein interaction specificity in these MADS-box genes has been considered to be the main factor in floral evolution in angiosperms (Lamb and Irish 2003; Litt and Irish 2003; Kramer et al. 2004; Stellari et al. 2004; Vandenbussche et al. 2004; Kim et al. 2005; Nei 2005; Nei and Rooney 2005: Zahn et al. 2005; Freeling and Thomas 2006; Irish 2006; Litt and Kramer 2010). In particular, *GLOBOSA*/*PISTILLATA* (*PI*) and *DEFICIENS*/*APETALA3* (*AP3*) orthologs, members of the B-class homeotic genes, appear more likely to be involved in angiosperm floral divergence. While other floral homeotic genes may experience high level of constraints due to their

pleiotropic roles in determining floral meristem identity, regulating apical meristems and specifying floral organ identity, B-class genes seem to have functions limited to petal and stamen development, and have been reported to evolve 20-40% faster than other plant MADS-box genes (Purugganan 1997). Moreover, enormous floral divergence is attributed to the morphological plasticity in petals and stamens, which was proposed to be the direct result of evolution of B-class genes (Bowman 1997; Kanno et al. 2003). Extensive gene duplication and subsequent functional divergence of duplicates of this gene lineage have been documented in angiosperms. For instance, *PI* and *AP3* paralogs were derived from an ancient duplication in the ancestor of angiosperms (Doyle 1994; Purugganan et al. 1995; Purugganan 1997; Theissen et al. 1996; Kim et al. 2004). In *Arabidopsis*, they form obligate heterodimers required for nuclear localization and DNA binding (McGonigle et al. 1996; Riechmann et al. 1996; Hill et al. 1998; Melzer and Theissen 2009). A later duplication event occurred at the base of the core eudicots for the *AP3* lineage, giving rise to the *TM6* and the *euAP3* lineages (Kramer et al. 1998). The *TM6* orthologs have either subsequently lost or specialized to specify stamen development while *euAP3* lineages are conserved for the development of petals and stamens (Kramer et al. 1998). Multiple smaller-scale duplication events of both *PI* and *AP3* lineages have been reported in several orders and families, e.g., in Ranunculales, the Asterids clade, and Orchidaceae (Kramer et al. 1998; Kramer et al. 2003; Mondragon-Palomino and Theissen 2008, 2009; Viaene et al. 2009). Differential expression patterns of the *AP3* duplicates were found to contribute to the ectopic petaloidy diversity in floral morphologies in Ranunculales and Orchidaceae (Kramer et al. 1998; Kramer et al. 2003;

Mondragon-Palomino and Theissen 2008, 2009). Multiple independent duplication events of *PI*-lineages were identified in the Asterids clade, including two ancient duplication events, one in the ancestor of Ericales and another in the ancestor of the core asterids (euasterids I and euasterids II) (Viaene et al. 2009). The expression patterns of *PI* duplicates in the clade indicated that the duplicated genes experienced differential selection forces and diverged in function for the novel floral morphologies in ericaceous flowers (Viaene et al. 2009).

These examples have illustrated the importance of B-class gene duplication and functional modification on floral divergences. However, the mechanisms of evolutionary changes after gene duplication that lead to novel morphologies are not well understood (Geuten and Irish 2010; Innan and Kondrashov 2010). In particular, it has been known that *PI*/*AP3* homologs usually form heterodimers and function in multi-subunit molecular complexes to promote expression of their downstream genes (Schwarz-Sommer et al. 1992; Trobner et al. 1992; Goto and Meyerowitz 1994; Riechmann et al. 1996). This interaction is maintained during evolution of B-class proteins in core eudicots. Studies on whether and how the protein interactions between PI- and AP3- like proteins affect the fate of duplicates will contribute to a better understanding of the molecular and evolutionary mechanisms shaping floral morphologies. Although functional divergence of B-class gene paralogs has been reported to account for diversity of floral morphology, their role in the evolution of petaloid bracts associated with inflorescences has not been well studied.

In this study, we reported results from analyses of molecular evolution and spatiotemporal expression patterns of *Cornus PI* orthologs (*CorPI*) and *Cornus AP3*

orthologs (*CorAP3*) in Cornaceae (the dogwood group). Cornaceae is an early-divergent lineage in the Asterids clade, consisting of four morphologically distinct groups. Two of the four groups (the big-bracted dogwood (BB) and the dwarf dogwood (DW)) contain species with petaloid bracts (Eyde 1987, 1988; Murrel 1993; Xiang et al. 2005, 2006; Fig. 1). In the previous study based on partial genomic DNA sequences, two copies of *CorPI* resulting from an ancient duplication event and a single *CorAP3* lineage were detected in *Cornus* (Zhang et al. 2008; Zhang 2008).

In this study, we examined the developmental pathways of petaloid bracts in two major groups of *Cornus* and compared the pathways to those of groups with non-petaloid bracts. We also isolated cDNA clones of *CorAP3* and *CorPI* from six species representing the four major groups in *Cornus*. We conducted phylogenetic analyses to characterize their evolutionary histories. Furthermore, we performed analyses to detect positive selection and molecular coevolution. Finally, we compared expression patterns of the *CorPI* paralogs at different flower and bract developmental stages using both RT-PCR and *in situ* hybridization. We detected divergence between paralogs of *PI* homologs in selection regimes, co-evolution with *CorAP3*, and expression patterns in *Cornus*. Results from developmental, evolutionary and genetic analyses suggest that the petaloid bracts in *C. florida* and *C. canadensis* evolved independently via divergent developmental and molecular pathways.

Materials and Methods:

Bract sample collection and developmental study

Bract materials were collected from four species representing the four major clades of *Cornus* for SEM analyses, *C. macrophylla* from the blue- or white- fruited group (BW; with minute, early deciduous, rudimentary, non-petaloid bractlets on inflorescence branches), *C. officinallis* from the cornelian cherry group (CC; with four modified, non-petaloid bracts at the base of the umbellate inflorescence), *C. florida* from the big-bracted group (BB; with four, large, petaloid bracts at the base of a capitate inflorescence; the morphologies of bracts were similar to those in CC at early stages, but later enlarged and become petaloid), and *C. canadensis* from the dwarf dogwood group (DW; with four large and petaloid bracts at the base of a minidichasium) (see Fig. 1 for bract morphology). Bract samples at the different developmental stages (Figs. 2, 3 and 4) were collected from plants growing on NCSU campus and JC Raulston Arboretum for all species except for *C. canadensis*. Materials of *C. canadensis* were collected from plants grown in NCSU Phytotron introduced from several wild populations collected from Spruce Knob, West Virginia and Stoddard, New Hampshire. Materials were fixed in 3.0% glutaraldehyde solution overnight at 4°C. They were dehydrated through a series of ethanol gradients, critical point dried and observed under a Scanning Electron Microscope as previously described (Feng et al., 2011).

B-class gene isolation

To obtain the complete cDNAs of B-class genes in *Cornus*, a combined total RNA sample from different developmental stages was used. Inflorescence buds of six species from

the four major clades, *C. officinallis* (CC group), *C. florida* (BB group), *C. canadensis* (DW group), *C. macrophylla* (BW group), *C. sanguinea* (BW group) and *C. controversa* (BW group) at different developmental stages were collected from the JC Raulston Arboretum and North Carolina State University main campus, and stored in RNAlater (Ambion of Applied Biosystems, Foster City, CA, USA). Total RNA at different developmental stages was isolated following the protocol modified from the CTAB RNA isolation method (Chang et al. 1993). RNA samples from the same species were combined as the template for first strand cDNA synthesis using SMART RACE cDNA Amplification Kit (Clontech, Alameda, CA, USA) and SuperScript III first-strand synthesis system kit (Invitrogen, Carlsbad, CA, USA). Twelve primers were designed to amplify *CorPI* and *CorAP3* using PCR (Table 1). PCR products were cloned into sequencing vectors using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA, USA). Positive clones were sequenced in the Genome Sequencing & Analysis Core Facility at Duke University after sequencing reactions using the M13 primers and DNA sequencing kit (Applies Biosystems, Warrington WA1 4SR, UK).

Phylogenetic analyses

To confirm the homology of *CorPI* and *CorAP3* sequences obtained in *Cornus* on a broader scale, cDNA sequences of *PI*- and *AP3*- like genes representing major clades in Asterids, Rosids, and basal eudicots were downloaded from GenBank (Supplementary table ST1). These two sets of sequences were combined with the *CorPI* and *CorAP3* sequences (including partial cDNA sequences from the study of Zhang et al., 2008), respectively. The sequences were translated to the predicted amino acid sequences and then aligned in

MUSCLE (Edgar 2004). Since sequences in the C domain are highly divergent among the basal eudicots, Asterids, and Rosids, preliminary phylogenetic analyses were conducted by excluding the C-domain sequences (including only the conserved sequences from the MADS, I, and K domain alignments). The resulting phylogenetic relationships were used as a guide to reorder the sequences in the data matrix following their positions in the phylogenetic tree to assist the alignment of the C domains. The C domains of both PI- and AP3- like proteins were further aligned in MUSCLE (Edgar 2004), followed by manual adjustment. Codon-based DNA alignments were finally conducted according to the protein alignments.

Phylogenetic analyses were first performed for these two large datasets of DNA sequences, each consisting of 86 taxa, 828 sites and 56 taxa, 843 sites, respectively, using the Maximum Likelihood (ML) and Bayesian methods implemented in RAxML 7.6.2 (Stamatakis 2006) and MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003), respectively. For ML analyses, the GTRGAMMA model was employed to search for the best-scoring ML tree. Analyses of 1000 bootstrap replicates were performed to estimate support for the ML trees. For the Bayesian analyses, Modeltest 3.06 (Posada and Crandall 1998) was first used to determine the best models of sequence evolution for *CorPI* and *CorAP3* genes separately. The GTR +I+ G model was suggested to be the best-fit model for *PI*-like genes and GTR+G model for *AP3*-like genes among the 56 models under AIC (Akaike Information Criterion). These models were used to set priors for Bayesian analyses. Two independent runs each with four chains of 1,000,000 generations were performed in MrBayes 3.1.2. Convergence was evaluated using the average standard deviation of split frequencies between runs and the

posterior distributions visualized in Tracer 1.4 (http://tree.bio.ed.ac.uk/software/tracer/). The first 10% trees were discarded as burn-in. After confirming the homology of *CorPI* and *CorAP3*, two small datasets only including *CorPI*, *CorAP3* and their closely relatives in Cornales were generated. Phylogenetic analyses were then conducted to reconstruct the gene genealogy for *CorPI* and *CorAP3* using ML and Bayesian methods as described above.

Molecular evolutionary analysis

Using the gene genealogies as the basis, we performed molecular evolutionary analyses to examine the dynamics of evolutionary mechanism governing the gene evolution. We estimated and compared the ratio of nonsynonymous to synonymous substitution rates (*dN/dS* or ω) under different branch- or site-models using the codeml program implemented in PAML version 4.4 (Yang 1997, 2007) to characterize the selective processes that may act on the *CorPI* and *CorAP3* genes. A ω value significantly less than one indicates purifying selection, equal to one suggests neutral selection, and significantly larger than one implies strong positive selection.

Since an ancient duplication of *CorPI* was detected in the gene genealogy during *Cornus* early diversification (also see Zhang et al. 2008), we first estimated whether the two ancient *CorPI* copies could be under heterogeneous selection by using the branch-models (Yang 1998; Yang and Nielson 2000). We compared the one-ratio model, which constrains a single ω on all branches, with the two-ratio model and the three-ratio model, which assume different ω values on stem branches of *CorPI*-A, and /or *CorPI*-B. We then evaluated the models by using the likelihood ratio test (LRT) to determine if the model assuming different

ω on stem branches of *CorPI*-A and *CorPI*-B is the best. To further examine whether selection differed among the four lineages of *Cornus,* or has altered after gene duplication, we estimated the ω values for the stem branch of each lineage and for branches that followed a gene duplication event under the free-ratio model (which allow omega to differ among branches) for both *CorPI* and *CorAP3*. The model was compared to the one-ratio model by the likelihood ratio test (LRT) to determine whether the free-ratio model fits the data significantly better than the one-ratio model.

We also applied the branch-site models to test if particular amino acid residues in the duplicated *CorPI* copies as well as *CorAP3* may be subject to positive selection. The branch-site models allow ω to vary both among sites in the protein and across the specified branch on the phylogenetic tree. The Bayes empirical Bayes (BEB) (Yang et al. 2005) was used to identify amino acid sites under positive selection by providing the posterior probabilities for site classes with $\omega > 1$. The branch-site model was compared with its corresponding null model with $\omega = 1$ fixed and evaluated by the likelihood ratio test (LRT) to see if it was significantly better than the null model.

Prediction of co-evolution

Because the protein products of *PI* orthologs were reported to interact with that of *AP3* orthologs and form heterodimers in regulating floral development, we continued to examine how protein interaction may affect the fate of duplicated genes. We first applied CAPS (Coevolution Analysis using Protein Sequences) program (Fares and Travers 2006) to identify amino acid sites that may have been co-evolved in *PI* orthologs and *AP3* orthologs.

Two multiple alignments, one including *PI*-like genes from Rosids and Asterids (core eudicots) as well as *CorPI* generated in this study, and another including corresponding *AP3*-like genes (data-pair 1) were generated for co-evolution analysis using the program CAPS (http://bioinf.gen.tcd.ie/caps/). CAPS compared correlated variance of evolutionary rates at two sites corrected by the time since the divergence of two protein sequences. Since substitutions at two independent sites can not be directly compared, the program applied blocks substitution matrix (BLOSUM) to calculate the transition probability (Henikoff and Henikoff 1992). The BLOSUM values were further normalized by the time of divergence between sequences. As the *PI* and *AP3* orthologs included in this interacting analysis are from the Rosid and Asterid clades, which are estimated to have diverged about 116 to 127 million years ago (Wikström et al. 2001; Bell et al. 2010), we used the Poisson model for time correction. The program then identified coevolved mutation pairs in *PI* and *AP3* orthologs by measuring the correlation in the pairwise amino acid variability relative to the mean pairwise variability per site between them. To determine significance of these correlations, 10000 random samplings were performed and correlation coefficient was tested under a normal distribution using a confidence value of (α=0.01). Correlated pairs with correlation coefficients larger than 0.4 were recorded, plotted and mapped onto the amino acid sequences. To eliminate the effect of phylogeny on the coevolution analysis in the large data set (data-pair1) and to identify sites under structural, functional or physical interaction between CorPI and CorAP3 in *Cornus*, we analyzed a smaller dataset retaining only sequences of *CorPI* and *CorAP3* (data-pair 2) using the same parameter settings. Results

were then compared to that of data-pair 1. According to the authors of the program, coevolving amino acid sites detected irrespective of sequences removed are predicted to have structural, functional or physical interaction (Fares and Travers 2006).

Since the CAPS program can only detect coevolved pairs at sites with variable amino acids, the lack of coevolution in a protein pair determined by the program indicates either truly lack of correlation or conservation of the protein sequences across species in both genes. We calculated the Boltzmann-Shannon entropy to measure the protein sequence conservation and variability at each site statistically. A large entropy value indicates high variation in certain sites and zero means no amino acid variation.

Predictions of the secondary structure and the amino acid accessibility of CorPI and CorAP3

We applied the newly developed program NetSurfP (http://www.cbs.dtu.dk/services/NetSurfP/) to predict secondary structures and calculate relative surface accessibilities (RSA) for each amino acid residue in CorPI and CorAP3 (Petersen et al. 2009). RSA of twelve CorPI amino acid sequences including both CorPI-A and CorPI-B and six CorAP3 protein sequences from six dogwood species were estimated using NetSurfP program. The program calculated RSA by dividing the solvent accessible surface area (ASA) of a given residue in a polypeptide over the maximal possible solvent exposed area of this residue in a tri-peptide flanked with either glycine (Chothia 1976) or alanine (Ahmad et al. 2003) residues. The prediction reliability was evaluated simultaneously by a Z-score where high Z-scores indicate lower prediction errors (Petersen et al. 2009). The

calculated RSA values were plotted and compared among species and paralogs to determine whether amino acid mutations in these genes may have affected the protein structures.

RNA probe transcription and in situ hybridization

In situ hybridization was used to examine the spatial pattern and level of *CorPI* –A, *CorPI* –B and *CorAP3* expression in early developmental stages of bracts and flowers in *C.officinallis*, *C.florida* and *C.canadensis*, the species with 4-fold bracts. RNA probes were in vitro transcribed following a protocol previously described (Franks et al. 2002) using primers designed for each species (Table1). They were further hydrolyzed into pieces between 75 and 150 bp long through carbonate hydrolysis following the protocols of the Long lab (http://pbio.salk.edu/pbiol/in_situ_protocol.html). The whole *in situ* hybridization protocol for *Cornus* species was modified from protocols of Franks' lab, Soltis' lab, and Koes' lab (Franks et al. 2002; Kim et al. 2005; Sourer et al. 2008). The hybridization specificity of *CorPI*-A and *CorPI*-B RNA probes to the corresponding RNA transcripts was tested using RNA dot blot protocol (Sambrook and Russell 2006). 10-fold to 100-fold differences were observed between *CorPI*-A and *CorPI*-B probes after probe hydrolysis under the hybridization and wash conditions similar to *in situ* hybridization (data not shown). Samples of inflorescence buds at bract initiation stages and floral organogenesis stages (Stages I and IV, respectively, Feng et al. 2011) were fixed in formaldehyde for at least 8 hours at 4°C and then dehydrated in a series of cold gradual ethanol solutions, permeated with an analytical grade xylene (Fisher, Fair Lawn, NJ, USA) series and embedded in Paraplast®Plus (Fisherbrand, Houston, TX, USA). The tissues were sectioned at a thickness

of 8 μm using a microtome and mounted on slides. They were treated with proteinase K (0.5 ng/ul) at 37℃ for 30 minutes and hybridized with specific probes (0.8-1.5 ng/ul) at 50℃ overnight. Slides were washed in 0.2X SSC at 55℃ with agitation twice. Signals were detected using Western blue (Promega, Madison, WI, USA). The *in situ* experiments for each probe were repeated at least two times with multiple biological replicates in each experiment. Both positive and negative controls were included for each experiment. Analyses of *CorPI* paralogs were conducted in the same experiments to ensure accurate comparisons of expression levels and avoid bias from experimental conditions.

RT-PCR

Expression patterns of *CorPI* and *CorAP3* in late developmental stages of bracts and flowers (Stage V, Feng et al. 2011 Figs. 3, 4) were investigated by RT-PCR. Inflorescence buds, including bracts and flowers of *C. officinallis, C.florida* and *C. canadensis* representing three important late developmental stages of bracts were fixed in RNAlater. These three stages included unopened bracts (B1), expanding bracts (B2) and white bracts (B3) (stages 1, 3, 4 in *C.florida* and stages 3, 4, 8 in *C. canadensis*) (see Figs. 3, 4, 5, 6). For *C. officinallis*, which is without petaloid bracts, inflorescence buds with developing bracts in June (B1), mature bracts in October (B2) and opened bracts in February (B3) during anthesis were collected as a comparison. Flowers at these three stages are referred to as F1, F2 and F3 in all three species. After bracts and flowers were manually separated, total RNA from these tissues was extracted using the CTAB method (Chang et al., 1993). SuperScript III Reverse Transcriptase from Invitrogen was used to synthesize cDNAs. Two types of *CorAP3*

transcripts were found *C. florida* and *C. canadensis*, and three types were found in *C. officinallis*. In each species, one of the types is the full-sequence and the other types are truncated forms resulting from alternative splicing that caused frame-shift or early termination of transcription (see details in Results). Primers were, thus, designed to distinguish expression of the two *CorPI* paralogs and alternatively spliced forms of *CorAP3* in these species. Their sequences can be found in Table 2. The PCR primer pairs for distinguishing *CorAP3* forms have at least one primer spanning across the spliced region. PCR conditions were optimized for each *CorPI* paralog to prevent annealing of primers with any mismatches to obtain accurate results of *CorPI*-A and *CorPI*-B expression. Multiple pairs of PCR primers were initially tested and the pairs generating the best results were chosen for final analyses. The RT-PCR analyses were repeated for the same cDNA libraries (at least once) as well as different cDNA libraries from at least one more biological duplicate. Expression patterns of different types of *CorAP3* in three BW species (*C. controversa, C. sanguinea, and C. macrophylla*) were also examined using type-specific primers that were designed to stretch across the spliced regions.

Analysis of correlation between B-class gene evolution and petaloid bracts

Since activities of B-class genes likely play a role in bract petaloidy in *Cornus*, correlation between B-class gene expression/evolution and petaloid bract morphology is expected. We conducted character-mapping analyses to identify co-changes in the genes and bract morphology on phylogenetic branches to decipher possible links of evolutionary changes in B-class genes to novelty in bract morphology. Ten characters derived from

developmental pathways of petaloid bracts, molecular evolution and expression analyses of B-class genes in *Cornus* were analyzed (summarized in Table 3). These characters were traced over the phylogeny of the studied species, corresponding to the simplified phylogenetic tree of *Cornu*s derived from DNA sequences of multiple genes (Xiang and Thomas 2008; Xiang et al. 2011) (the studied species represent all the four major clades). The analysis was conducted using the likelihood method with the mk1 model (one-parameter Markov k-state model) in MESQUITE 2.74 (Maddison and Maddison 2007) to reconstruct the ancestral states of these characters. The reconstructed evolutionary history of each character was then superimposed on the same tree topology to show character state changes on a branch. Species in each clade that were not included in the study largely conserve the bract morphology. Slight variation in the BB group (e.g., one species, *Cornus disciflora*, with 4-fold bracts that are non-petaloid) would only switch the relative timing of the changes identified on the BB branch.

Results:

Bract development and bract petaloidy in Cornus

In *C. canadensis*, *C. florida* and *C. officinallis* there are two pairs of fold bracts at the base of an inflorescence bud (Fig. 1 A, B and C). *Cornus macrophylla* from the BW group has early deciduous, rudimentary bracts that are small and subtend low inflorescence branches (Fig. 1 D). Bracts from these four species initiate and start development during early inflorescence developmental stages in the spring (*C. officinallis*) or summer (other species) after anthesis. The bract growth ceases in the winter. In the following spring, the

bracts open and expand before anthesis. Thus, the whole developmental process in all species spans two years and is divided into an early phase (bract initiation and development before winter) and a late phase (bract expansion before anthesis in the following spring) in this report.

Two different bract initiation patterns in the early phase were observed among the four species studied. In *C. florida* and *C. officinallis*, the two pairs of bracts are initiated before the initiation of inflorescence meristems (Fig. 2, D, G; Feng et al. 2011). That is, two pairs of bract primordia are generated before the central meristems become a disk-like mature inflorescence meristem that is ready to generate inflorescence branch meristems (Fig. 2 D, E, G, H). Bracts in *C. macrophylla* and the two pairs of bracts in *C. canadensis* are generated after the initiation of primary inflorescence branch meristems from the central inflorescence meristems. Bract primordia appear to arise from the outer (abaxial) portion of each inflorescence branch meristem (Fig. 2 B, K). Shortly after bract initiation, trichomes develop on the abaxial surface of bracts in *C. macrophylla*, *C. officinallis* and *C. florida* (Fig. 2 B, E, H). In *C. florida* and *C. officinallis*, the bracts continue to grow and finally cover the whole inflorescence and protect the winter inflorescence buds. In *C. canadensis* and *C. macrophylla*, however, the young inflorescences with the bracts are protected by outer scales in the winter. Bracts are exposed only in the following spring when inflorescences are expanding. No trichomes are seen on the surface of *C.canadensis* bracts and the bracts never cover the whole inflorescence in the winter buds (Fig. 2 L).

Little growth is observed in bracts of *C. macrophylla* and *C. officinallis* in their late developmental phase. In *C. florida* and *C. canadensis,* bracts will expand, gradually lose chlorophyll and become white within about one month before anthesis. Six stages in *C. florida* and nine stages in *C. canadensis* were defined during bract late development (Fig. 3, Fig. 4). Around mid March, inflorescence buds in *C. florida* break open (stage 1, Fig. 3 A). After one week, bracts completely open (stage 2, Fig. 3 B). Then they will enlarge significantly (stage 3, Fig. 3 C). The expanded bracts will turn white after two weeks (stage 4, Fig. 3 D). The terminal flower (the apical-most true flower) will then change petal color from green to yellow (stage 5, Fig. 3 E) and finally open (stage 6, Fig. 3 F). In *C. canadensis*, winter buds swell in May in nature (stage 1, Fig. 4 A). Stems elongate and shoots grow out from the protective scales in the next two weeks (stage 2, Fig. 4 B, C). Leaves will expand and bracts will be visible at the tip at the end of the third week (stage 3, Fig. 4 D). At this stage, bracts are about 3 mm long. Bracts will expand to about 15 mm long after one week (stage 4, Fig. 4 E, F, G). Flowers will be visible at this stage (stage 4, Fig. 4 G). The color of the corolla of the central flower then changes from green to white (stage 5, Fig. 4 H). About two days later, terminal flowers in the first pair of inflorescence branches also become white. The outer bracts gradually change to white from the tip to the base (stage 6, Fig. 4 I). The same phenomenon happens in the second pair of inflorescence branches and inner bracts (stage 7, Fig. 4 J). Finally, bracts become completely white (stage 8, Fig. 4 K) and the terminal flowers open (stage 9, Fig. 4 L).

Epidermal morphologies of bracts between these two species were compared. On unexpanded bracts (stage 2 in *C. florida* and stage 3 in *C. canadensis*), epidermal cells of the adaxial surface of these two species are similar. They are small, round and tightly packed. Division planes are still seen in some of the sister cells (Fig. 5 A, Fig. 6 A). These cells then lengthen and widen to two fold or three fold in the expanding stage. They are still un-domed at this stage (stage 3 in *C. florida*, Fig. 5 B, stage 4 in *C. canadensis*, Fig. 6 B). In the whitening bracts the cells will expand upward and become dome-shaped and display a surface morphology characterized by a set of fine ridges arranged in a radial pattern. No stomata are seen on the adaxial surface of bracts (stage 4 in *C. florida*, Fig. 5 C, D, stage 8 in *C. canadensis*, Fig. 6 C, D). The abaxial surface of unexpanded bracts in *C. florida* is densely covered with two-branched trichomes. Epidermal cells underneath are more regular in shape than that in *C. canadensis* (Fig. 5 E). They are arranged in long smooth files. Cells can elongate in expanding bracts and become interlocking pavement-like cells in expanded bracts. There are occasionally stomata among these cells (Fig. 6 F, G). Epidermal cells on the abaxial side of unexpanded bracts in *C. canadensis* are irregularly shaped and interspersed with many stomata (Fig. 6 E). They enlarge rapidly and become smooth, wide, and pavement-like (Fig. 6 F, G).

Gene duplication in CorPI and alternative splicing in CorAP3

The sequences obtained in this study were confirmed to be homologous to *PI* and *AP3* sequences by phylogenetic analyses with other *PI*- and *AP3*-like sequences from Asterids, Rosids and basal eudicots (Supplementary figures SF1 and SF2). Our study

obtained 30 *CorPI* sequences and 16 *CorAP3* sequences covering the entire coding region for the six species of *Cornus*. The *CorPI* sequence lengths vary among species from 600 bp to 630 bp, while *CorAP3* sequences are slight larger and range from 678 bp to 684 bp. Several different forms of truncated *CorAP3* transcripts are also identified, which range from 546 bp to 611 bp.

Phylogenetic analysis of *CorPI* sequences uncovered an ancient duplication event that occurred before *Cornus* diversified into major lineages (Fig. 7; Fig. SF1) as also shown in Zhang et al. (2008). The sequences from the six species were sorted into two monophyletic clades: corresponding to *CorPI*-A and *CorPI*-B, with high posterior probabilities and ML bootstrap supports. *CorPI*-A clade contains all six species while *CorPI*-B clade contains only *C. florida* (BB), *C. canadensis* (DW), and *C. officinallis* (CC). cDNA of *CorPI*-B was not detected from species of the BW group due to the loss of the copy (Zhang et al. 2008). Two more recent duplications of *CorPI*-A were also detected within the CC and BW group (Fig. 7). Within the *CorPI*-A clade, the monophyly of each of the four morphological groups is highly supported, although relationships among the groups is unresolved except the sister relationship between the BB and DW groups that produce petaloid bracts (Fig. 7).

Compared to *CorPI*, only one copy of *CorAP3* is detected in *Cornus* (Fig. 8). However, alternatively spliced mRNA of *CorAP3* for the all six species are recognized based on their sequence similarity to the expected transcripts, genomic sequence data, and by comparing to the exon/intron positions in *Arabidopsis thaliana* (Fig. 9). Seven exons and six introns are identified in *Arabidopsis AP3* gene. Based on partial genomic sequences of

CorAP3 in *Cornus* species, which include the first four exons and three introns, the exon/introns positions in *Arabidopsis AP3* are conserved in *CorAP3*. We detected two types of *CorAP3* transcripts in five species and three types of *CorAP3* transcripts in *C. officinallis* differing in length. The shorter sequences in each species are not detected at the genomic DNA level, suggesting that they are generated after transcriptions of *CorAP3*. By comparing the sequence structures between the long and short copies, we identified four types of alternative splicing in *Cornus* (Fig. 9). Type I and II were detected in *C. officinallis* (types I and II, Fig. 9). In type I, the second exon is completely missing, resulting a premature stop codon in the I domain. The translated protein is predicted to consist of the MADS domain and partial I region (Fig. 9). Type II is missing 79 bp of the seventh exon and it also results in an in-frame stop codon. The truncated polypeptides will lose most of the C domain. Type II is also found in *C. canadensis* and *C. florida*. Type III transcript is found in *C. controversa*, which splices 138 bp out of the first exon. As a result, 40 amino acids of the MADS domain, and the first 6 amino acids of the I domain are missing in the predicted protein. In *C. sanguinea* and *C. macrophylla*, another type of alternative splicing is revealed. The fifth and sixth exons are completely deleted in the truncated transcript. It results in a 16 amino acid deletion in the K domain and a 13 amino acid deletion in the C domain (Fig. 9).

Divergent selection patterns in CorPI paralogs and CorAP3

Results from two-ratio models indicated that the selection force ω (*dN*/*dS* ratio) on *CorPI*-B stem is about three times greater than that on *CorPI*-A and the outgroups (0.71 vs 0.24 in model C, Table 4). The selection pressure on *CorPI*-B is significantly relaxed

compared to that on *CorPI*-A as a LRT test supported that this two-ratio model fits the data significantly better than the one-ratio model does (Model C vs Model A, Table 5). Heterogeneous selection on duplicated copies after gene duplication is also shown in BW group for *CorPI*-A by analysis using the free-ratio model. After gene duplication in the BW group, omega significantly increased in one lineage and significantly reduced in the other (0.55 vs 0.12, Fig. 7). Analysis with the free-ratio model further indicates purifying selection on *CorPI*-A during the early diversification of the genus, with relaxation in the DW group (0.35, 0.42, 0.12 and 0.13 in ancestors of BW, DW, BB and CC groups, respectively; Fig. 7). In the CC group, strong positive selection is identified after the duplication of *CorPI*-A in this lineage ($\omega >> 1$, Table 4, Fig. 7). The free-ratio model also suggests a much more relaxed selection in *CorPI*-B (0.68, Fig. 7).

Analyses using the branch-site model allowing the ω ratio to vary among sites in the protein sequences and across the assigned branch in the phylogenetic tree (model A) detected no positively selected sites in CorPI-A, but nine sites under potential positive selection in CorPI-B, one (site 186) supported by high posterior probabilities (Table 6, Figs. 7, 10). Amino acid substitutions at this site involve exchanges between the polar amino acid (glutamine-Q) and conserved hydrophobic amino acids (leucine-L or valine-V). Results of the likelihood ratio test indicate that the model detecting positive selections for these sites (Model A) is significantly better than the null model where ω ratio was fixed as 1, supporting the positive selection in *CorPI*-B.

Purifying selection was also shown for *CorAP3* in analyses using the free ratio model (Fig. 8), but with a relaxation in the BW and BB groups (omega from 0.25 to 0.56 in BW and 0.4 in BB), and increased force in the DW group (from 0.25 to 0.18; Fig. 8). Likelihood ratio test for selection force difference in *CorAP3* of DW and BW under different branch models did not show statistical significance (Tables 7, 8). However, results from the branch-site model showed that the discrete model (M3) allowing ω values to vary among sites is significantly better than the one-ratio model under LRT test, suggesting that ω values vary among sites in *CorAP3* (Tables 9, 10). Results from two models allowing ω values >1 in some sites (M2a and M8) further revealed that 0.45% of *CorAP3* sites is under positive selection. LRT tests comparing these two models to the neutral models (M1a and M7, respectively) show that positive selection models are significantly better than the neutral models, providing strong evidence for adaptive evolution (Table 10). Three amino acids identified to be under positive selection, one (site 197) with supports of posterior probabilities > 90% (Table 10, Figs. 8, 10B). Amino acid substitutions at this site (site 197) are involved in exchange between polar amino acids (or ambivalent amino acids) and non-polar amino acids. Glycine (G, non-polar, neutral) is found in BW species and *C. canadensis*, alanine (A, non-polar, neutral) was found in *C. florida*, and aspartic acid (D, polar, negative) was found in *C. officinallis* (Figs. 8, 10B).

Estimation of coevolution and surface accessibility of CorPI and CorAP3 in Cornus

Correlated mutation pairs between *PI* and *AP3* orthologs in two data-pairs (data-pair1 and data-pair2) were identified using CAPS program. Two sites (sites 182 and 186, the latter

site under positive selection) in *PI* orthologs are consistently found in these two data pairs to co-evolve with one site (site 197 under positive selection) in *AP3* orthologs. This suggests that they might be under structural, functional or physical interaction (Figs. 10, 11A, 11B). By mapping these sites onto the peptide sequences, we find that they are within the C domains of *CorPI* and *CorAP3* (Fig. 10). The two sites in *CorPI* are both involved in hydrophobic-hydrophilic exchange between *CorPI*-A and *CorPI*-B. At site 182, the relatively hydrophilic amino acid tyrosine (Y) was found in *CorPI*-A while the relatively hydrophilic amino acid threonine (T) was fixed in *CorPI*-B protein in *C. florida*. At site 186, also predicted under positive selection with high posterior probabilities, the highly hydrophobic amino acid leucine (L) or valine (V) was found in *CorPI*-A, but highly hydrophilic amino acid glutamine (Q) was identified in *CorPI*-B in *C. florida*. Consistently, the estimated interacting site (site 197, also under positive selection) in *CorAP3* has also experienced hydrophobic-hydrophilic substitution among species. Glycine (G) is fixed in the BW and DW (as in *C.canadensis*) groups at site197 while Alanine (A) and Aspartic acid (D) are identified in the BB (*C. florida*) and CC (*C. officinallis*) groups.

The NetSurfP program indicates that amino acid substitutions in both *CorPI* and *CorAP3* do not affect the secondary structure prediction (data not shown). When the results of relative surface accessibility (RSA) for all sites from the analyses are plotted to compare among the proteins, the graph shows that protein accessibilities are highly conserved in the other domains but vary in the C-domain among species and copies, especially in CorPI-B and CorAP3 in *C. florida* (Figs. 12, 13). RSAs of amino acids in six sites in the C domain (sites

177, 179, 180, 186,188, and 190; site 186 was predicted under positive selection and interacting with site 197 in CorAP3) in CorfloPI-B (CorPI-B of *C.florida*) are different from RSAs of the C domain in other CorPI proteins (Fig. 12; Supplementary table ST2). Seven amino acid sites in the C domain of CorfloAP3 (CorAP3 of *C.florida*) may change their relative surface accessibilities compared to the amino acids at the same sites of other CorAP3 (Fig. 13; Supplementary table ST3).

Expression pattern of CorPI and CorAP3 during flower and bract development

Expression pattern of *CorPI* and *CorAP3* in *Cornus* in various stages and tissues are summarized in Table 11. As expected, *in situ* hybridization experiments showed that expressions of *CorPI* and *CorAP3* are detected in petals and stamens in all the species examined during early floral organogenesis. They are highly expressed during initiation of petals and stamens, but showed reduced expression levels when the petals and stamens are well-developed (Fig. 14). Clones from the pooled cDNA from these flower early stages reveal both copies of *CorPI* in *C. florida* and *C. canadensis*, and only *CorPI*-A mRNA in *C. officinallis* and the three BW group species included in this study (data not shown). Expression of *CorPI*-A is much higher than that of *CorPI*-B in *C. canadensis* while *CorPI*-B is the dominant transcript in *C. florida* based *on in situ* hybridization (Fig. 14). Up-regulation of *CorPI* and *CorAP3* is observed again in flower late development (F1, F2 and F3) by RT-PCR (Fig.15, Table 11). Using primers that can distinguish *CorPI*-A and *CorPI*-B in RT-PCR, results indicated that both copies of *CorPI* are highly expressed in *C. florida*, *C. canadensis* and *C. officinallis* flowers at F1 stage, although *CorPI*-B appears to be expressed

at a slightly higher level (Fig. 16). The expression level of *CorPI*-A greatly decreases in stages of F2 and F3 in *C. florida*, but maintains a similar level in the other two species (Fig. 16). For *CorAP3*, both primary (fully spliced) and alternative splice forms are detected in three stages of flower late development using RT-PCR in all six species examined (Fig. 16). Fully spliced *CorAP3* are strongly expressed, while type II, type III and type IV transcripts are weakly expressed. Type I transcript of *CorAP3* is only found in *C. officinallis* and it is highly expressed in all three flower late stages in this species (F1, F2 and F3, Fig. 16, Table 11).

In bracts, expression of *CorPI* and *CorAP3* is not observed by *in situ* hybridization during bract initiation and early development in all species examined (Fig. 14). For *C. florida*, *C. canadensis* and *C. officinallis*, which bear two pairs of bracts, we also examined the expression pattern of *CorPI* and *CorAP3* in the three bract late stages (B1, B2 and B3). Strong signals of *CorPI*-B as well as *CorAP3* are detected in *C. florida* bracts at the B1 stage. Their expression levels are down-regulated when bracts are fully expanded and become petaloid (Fig. 16). In *C. canadensis* and *C. officinallis,* only weak expression of fully spliced *CorAP3* is detected in bracts, and no expressions of *CorPI* is observed in developing bracts (Fig. 16). With a lower stringency in annealing temperatures and different primers, and more cycles (40), *CorAP3* expression was detected in all three species in all three stages B1-B3, except in B1 of *C. canadensis*, at the lower level than in flowers; strong *CorPI* expression was detected in B1 stage of *C. florida* and weak expression in B2 and B3 and in all stages of the other two species (Fig. 15).

Correlation of B-class gene evolution and petaloid bracts

Likelihood mapping of the ten characters related to the evolution and expression in bracts of B-class genes and petaloid bract morphology show seven character state changes co-occurred in the BB lineage (e.g., *C. florida*) with large petaloid bracts, including *CorfloPI*-A expression from none to weak (Character 3), *CorfloPI*-B expression from none to strong (Character 4), increased strong expression of *CorAP3* in bracts at B1 stage (Character 5), a change from Valine (V) to Leucine (L) in *CorPI*-A at the site 186 identified to interact with *CorAP3* (Character 6), a change from Glycine (G) to Alanine (A) in *CorAP3* at the site 197 identified to interact with *CorPI* (Character 8), and a change from Tyrosine (Y) to Threonine (T) in *CorPI*-B at the site 182 that interacts with *CorAP3* (Character 10, Fig. 17).

Discussion:

Coevolution between CorPI and CorAP3 affects the fate of CorPI duplicates in Cornus

Divergence in expression and/or function of duplicated genes has been hypothesized to play a vital role in phenotypic evolution (Ohno 1970; Force et al. 1999; Innan and Kondrashov 2010). Following duplication, paralogs may maintain the ancestral function (for functional redundancy) or change the ancestral function due to mutations and shifts of natural selection. One of the duplicated copies may become a pseudogene (lose its function) as a result of accumulation of deleterious mutations (nonfunctionalization), or the two paralogs split the ancestral functions between them (subfunctionalization). Alternatively, one or both paralogs evolve new, divergent functions (neofunctionalization) by fixation of beneficial

mutations (Force et al. 1999; Lynch and Force 2000; Moore and Purugganan 2005). Finally, two functions of an ancestral gene can be improved after the duplication as the result of release of functional constraint in the single gene ("escape from adaptive conflict", Des Marais and Rausher 2008). Our molecular evolutionary analyses of *CorPI* and *CorAP3* combining data from expression patterns reveal differences in selection regimes and expression patterns between the *CorPI* paralogs. Our finding of correlation between spatiotemporal expression patterns of *CorPI* paralogs and mutations at co-evolving sites under positive selection in *CorPI* and *CorAP3* supports that co-evolution in B-class proteins affects the fate of *CorPI*-duplicates in the genus.

Our data suggest that *CorPI* paralogs in *Cornus* are likely selected for divergent partnership with *CorAP3* in early floral development among lineages (e.g., *CorPI*-A---*CorAP3* in CC, DW, and BW groups; *CorPI*-B---*CorAP3* in BB group). In *C. florida* with petaloid bracts, partnership between co-mutated *CorPI*-B and *CorAP3* is favored for early floral development (Fig. 14) and extends its expression to the bracts. This partnership remains dominantly expressed in the course of floral development in the species. Mutations at the interacting sites of the two proteins in *C. florida* were predicted to result in changes in surface accessibility in the C-domain of these proteins, a domain important for higher order protein interactions (Egea-Cortines et al. 1999; Honma and Goto 2001; Lamb and Irish 2003). Thus the mutations could have possibly altered the interaction behaviors of the B-class complex with other proteins or cofactors in *C. florida*. In *C. officinallis* (CC) and *C. canadensis* (DW) group, *CorPI*-A and *CorAP3* are preferentially co-expressed in early

flower development stages (Fig. 14) and not expressed in bracts (Figs. 14, 16); co-expression of *CorPI*-B, *CorPI*-A and *CorAP3* occurred only in late flower development, in contrast to *C. florida* (BB group) where *CorPI*-B and *CorAP3* are strongly co-expressed in both early and late flower development stage and in bracts development. This suggests divergent functions of the *PI* paralogs at different stages and lineages, and that both *CorPI* paralogs retain functions in floral development, but only *CorPI*-B appears to function in bract development in *C. florida*. The results together further support that the B-class gene is selected for divergence in both spatial and temporal subfunctionalization of *CorPI* paralogs in *Cornus* after the ancient duplication.

Proteins *PI* and *AP3* are known to interact to form heterodimers to determine the petal identity in model plants. It is expected that the two genes are co-evolving. However, the details of coevolution between the two genes/proteins remain poorly understood (Geuten and Irish 2010;van Dijk and van Ham 2010). Study of Viaene et al. (2009) indicated that negative selective pressure on *PI* paralog codons is relaxed in the Asterid clade, whereas substitutions at sites putatively involved in protein–protein interaction show positive selection, allowing for a change in the interaction behavior of the *PI* paralogs after duplication. Study in Solanaceae also revealed relaxed purifying selection in *PI* paralogs and detected a site in the C domain under positive selection (Geuten and Irish 2010). In our study of *Cornus*, we similarly detected relaxed purifying selection in *CorPI* paralogs following gene duplication. We did not detect positive selection or co-evolving sites in the *CorPI*-*CorAP3* interaction domains (K1 and K2 domains; Yang et al. 2003; Yang and Jack 2004). However, we

detected positive selection at the co-evolving sites in the C-domain of *CorPI* and *CorAP3*, but not the same site detected in Geuten and Irish's study (2010). The predicted changes in surface accessibility of the *CorPI*-B-*CorAP3* complex in *C. florida* brought about by the amino acid mutations favored by natural selection at the co-evolving sites suggests a potential link between the alteration of *CorPI-CorAP3* complex surface features that may have affected higher order protein interactions, which may be important to expand their functions heterotopically in the nearby organs (bracts) in *C. florida.* Future analyses of protein interactions and functions will be helpful to test this hypothesis.

Independent origin of petaloid bracts in Cornus

Our developmental study indicated that petaloid bracts in *C. florida* and *C. canadensis* might not be homologous. Bracts in *C. florida* initiate outside of the inflorescence meristem and subtend the whole inflorescence, while bracts in *C. canadensis* initiate inside the inflorescence meristem and subtend the inflorescence branches. They also differ in bract late development. Bracts in *C. florida* become white before the central flower changes color, while whitening of bracts in *C. canadensis* occurs after color change of central flowers. It suggests that bract petaloidy in *C. canadensis* may require the maturation of flowers. We found that removal of developing flowers between stage 3 and 4 in *C. canadensis* resulted in a cessation of bract expansion and acceleration of whitening in *C. canadensis*, while bract expansion and whitening in *C. florida* were not affected by flower removal (data not shown). This also suggests that immature flowers may be required for bract expansion but suppress bract whitening in *C. canadensis*.

Our expression study also suggested independent origin of bract petaloidy in *Cornus*. RT-PCR showed a difference in B-class gene expression in petaloid bracts in *C. florida* and *C. canadensis*. Strong expression of both *CorPI*-B and *CorAP3* was detected in developing bracts in *C. florida*, while no significant expression of *CorPI* and *CorAP3* was recognized in *C. canadensis* (Fig. 16). Our result was consistent with the expression study from Maturen (2008), who studied the expression level of A- B- and E-class genes in *C. florida* and *C. canadensis* using real-time PCR. Maturen showed that A-, B- and E- class genes were all significantly expressed in *C. florida*, but not in *C. canadensis.* However, the study did not distinguish the expression of the paralogous copies (Maturen 2008).

Our data further showed that origin of petaloidy in *C. florida* is correlated with amino acid mutations in interacting sites between *CorfloPI* and *CorfloAP3*, while such correlation was not observed in *C. canadensis* bract petaloidy (Fig. 17). These hypotheses can be evaluated by future functional analyses of B-class genes.

Role of B class gene in bract petaloidy in Cornus

The lack of expression of B-class genes in bract initiation and early development suggest that petaloidy is not defined at early stages. Strong expression of *CorPI*-B and *CorAP3* is detected in *C. florida* during bract late development (B1; Fig. 16), while only traces of expression of *CorPI* and *CorAP3* is detected at similar stages in *C. canadensis* (Figs. 15, 16), suggesting the ABCE model is only recruited in bract late development and responsible for the bract petaloidy in *C. florida*. Surface morphology of petaloid bracts in these two species is highly similar, where the adaxial surfaces in both bracts are petal-like

and the abaxial surfaces are leaf-like (Figs. 5, 6), suggesting petaloidy in bracts in these two species may recruit the same genetic pathway in cell differentiation during bract late development. But genetic pathways regulating cell differentiation may be different in *C. florida* and *C. canadensis*. In *C. florida*, evidence suggests the recruitment of the ABCE model which triggers the expression of down-stream genes responsible for cell differentiation, while in *C. canadensis*, a different genetic pathway is involved for cell differentiation in petaloid bracts. Different molecular mechanisms underlying heterotopic petaloidy that may or may not require B class gene functions have also been observed in other plant groups. In a study on *Aristolochia,* (Aristolochiaceae, basal angiosperm), expression of an *AP3*-like gene was detected only at the late stage of petaloid perianth development, congruent with our finding in *C. florida*, which suggests that class B genes may only contribute to the late cell differentiation in petaloid organs (Jaramillo and Kramer 2004). In contrast, a study in *Asparagus officinallis* (Asparagaceae) reported no expression of class B orthologs in the outer whorl of tepals although both whorls of tepals were petaloid and identical (Park et al. 2003, 2004). This evidence together with our research in *Cornus* on heterotopic petaloidy, suggests that petaloid organs can evolve independently in different lineages, even within a single genus, through different genetic pathways and molecular mechanisms.

Acknowledgement:

The authors would like to give special thanks to S. Azhakanandam for technical assistance with *in situ* hybridization and Dr. DE Boufford for assistance in collecting *C. canadensis* in New Hemisphere. We thank Dr. DE Soltis, Dr. PE Soltis and Dr. RE Koes for helpful discussion, Jian Zhang and Juan Liu for providing comments for this manuscript. We acknowledge generous support from the JC Raulston Arboretum, the phytotron of College of Agriculture and Life Sciences at North Carolina State University and Center for Electron Microscopy at North Carolina State University. The research was supported by a NCSU Multi-disciplinary Research grant and NSF grant (IOS-1024629).

References:

Adamczak R, Porollo A, Meller J. 2004. Accurate prediction of solvent accessibility using neural networks-based regression. Proteins-Structure Function and Bioinformatics 56:753-767.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. Am. J. Bot. 97:1296-1303.

Bowman JL. 1997. Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. J. Biosci. 22:515-527.

Chang S, Puryear J, Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11: 113-116.

Chothia C. 1976. Nature of accessible and buried surfaces in proteins. J. Mol. Biol. 105:1-14.

Coen ES, Meyerowitz EM. 1991. The war of the whorls - genetic interactions controlling flower development. Nature 353:31-37.

Des Marais DL, Rausher MD. 2008. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. Nature 454:762-U85.

Doyle JJ. 1994. Evolution of a plant homeotic multigene family - toward connecting molecular systematics and molecular developmental genetics. Syst. Biol. 43:307-328.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Egea-Cortines M, Saedler H, 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 J. 18:5370-5379.

Eyde RH. 1988. Comprehending cornus - puzzles and progress in the systematics of the dogwoods. Bot. Rev. 54:233-351.

Eyde RH. 1987. The case for keeping cornus in the broad linnaean sense. Syst. Bot. 12:505-518.

Fares MA, Travers SAA. 2006. A novel method for detecting intramolecular coevolution: Adding a further dimension to selective constraints analyses. Genetics 173:9-23.

Feng CM, Xiang QY, Franks RG. 2011. Phylogeny-based developmental analyses illuminate evolution of inflorescence architectures in dogwoods (*Cornus* s. l., Cornaceae). New Phytol.

In press.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531-1545.

Franks RG, Wang CX, Levin JZ, Liu ZC. 2002. SEUSS, a member of a novel family of plant regulatory proteins, represses floral homeotic gene expression with LEUNIG. Development 129:253-263.

Freeling M, Thomas BC. 2006. Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. Genome Res. 16:805-814.

Geuten K, Irish V. 2010. Hidden variability of floral homeotic B genes in Solanaceae provides a molecular basis for the evolution of novel functions. Plant Cell 22:2562-2578.

Goto K, Meyerowitz EM. 1994. Function and regulation of the *Arabidopsis* floral homeotic gene *Pistillata*. Genes Dev. 8:1548-1560.

Henikoff S, Henikoff JG. 1992. Amino-acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. U. S. A. 89:10915-10919.

Hill TA, Day CD, Zondlo SC, Thackeray AG, Irish VF. 1998. Discrete spatial and temporal cis-acting elements regulate transcription of the arabidopsis floral homeotic gene APETALA3. Development 125:1711-1721.

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

Innan H, Kondrashov F. 2010. The evolution of gene duplications: Classifying and distinguishing between models. Nature Reviews Genetics 11:97-108.

Irish VF. 2006. Duplication, diversification, and comparative genetics of angiosperm MADS-box genes. Advances in Botanical Research: Incorporating Advances in Plant Pathology, Vol 44 44:129-161.

Jaramillo MA, Kramer EM. 2004. APETALA3 and PISTILLATA homologs exhibit novel expression patterns in the unique perianth of aristolochia (aristolochiaceae). Evol. Dev. 6:449-458.

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

Kim S, Koh J, Ma H, Hu Y, Endress PK, Hauser BA, Buzgo M, Soltis PS, Soltis DE. 2005. Sequence and expression studies of A-, B-, and E-class MADS-box homologues in eupomatia (eupomatiaceae): Support for the bracteate origin of the calyptra. Int. J. Plant Sci. 166:185-198.

Kim ST, Yoo MJ, Albert VA, Farris JS, Soltis PS, Soltis DE. 2004. Phylogeny and diversification of B-function MADS-box genes in angiosperms: Evolutionary and functional implications of a 260-million-year-old duplication. Am. J. Bot. 91:2102-2118.

Kramer EM, Jaramillo MA, Di Stilio VS. 2004. Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. Genetics 166:1011-1023.

Kramer EM, Di Stilio VS, Schluter PM. 2003. Complex patterns of gene duplication in the APETALA3 and PISTILLATA lineages of the ranunculaceae. Int. J. Plant Sci. 164:1-11.

Kramer EM, Dorit RL, Irish VF. 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.

Lamb RS, Irish VF. 2003. Functional divergence within the APETALA3/PISTILLATA floral homeotic gene lineages. Proc. Natl. Acad. Sci. U. S. A. 100:6558-6563.

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

Litt A, Kramer EM. 2010. The ABC model and the diversification of floral organ identity. Semin. Cell Dev. Biol. 21:129-137.

Lynch M, Force A. 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics 154:459-473.

Maddison WP, Maddison DR. 2007. Mesquite: a modular system for evolutionary analysis. Version 2.74. http://mesquiteproject.org.

Maturen NM. 2008. Genetic analysis of the evolution of petaloid bracts in dogwoods. University of Michigan.

McGonigle B, Bouhidel K, Irish VF. 1996. Nuclear localization of the arabidopsis APETALA3 and PISTILLATA homeotic gene products depends on their simultaneous expression. Genes Dev. 10:1812-1821.

Melzer R, Theissen G. 2009. Reconstitution of floral quartets in vitro involving class B and class E floral homeotic proteins. Nucleic Acids Res. 37:2723-2736.

Mondragon-Palomino M, Theissen G. 2009. Why are orchid flowers so diverse? reduction of evolutionary constraints by paralogues of class B floral homeotic genes. Annals of Botany 104:583-594.

Mondragon-Palomino M, Theissen G. 2008. MADS about the evolution of orchid flowers. Trends Plant Sci. 13:51-59.

Moore RC, Purugganan MD. 2005. The evolutionary dynamics of plant duplicate genes. Curr. Opin. Plant Biol. 8:122-128.

Munster T, Wingen LU, Faigl W, Werth S, Saedler H, Theissen G. 2001. Characterization of three GLOBOSA-like MADS-box genes from maize: Evidence for ancient paralogy in one class of floral homeotic B-function genes of grasses. Gene 262:1-13.

Murrell ZE. 1993. Phylogenetic-relationships in cornus (cornaceae). Syst. Bot. 18:469-495.

Nei M. 2005. Selectionism and neutralism in molecular evolution. Mol. Biol. Evol. 22:2318-2342.

Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. 39:121-152.

Ohno S. 1970. Enormous diversity in genome sizes of fish as a reflection of natures extensive experiments with gene duplication. Trans. Am. Fish. Soc. 99:120-&.

Park JH, Ishikawa Y, Ochiai T, Kanno A, Kameya T. 2004. Two GLOBOSA-like genes are expressed in second and third whorls of homochlamydeous flowers in asparagus officinalis L. Plant and Cell Physiology 45:325-332.

Park JH, Ishikawa Y, Yoshida R, Kanno A, Kameya T. 2003. Expression of AODEF, a B-functional MADS-box gene, in stamens and inner tepals of the dioecious species asparagus officinalis L. Plant Mol. Biol. 51:867-875.

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. 2000. B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature 405:200-203.

Petersen B, Petersen TN, Andersen P, Nielsen M, Lundegaard C. 2009. A generic method for assignment of reliability scores applied to solvent accessibility predictions. Bmc Structural Biology 9.

Posada D, Crandall KA. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics 14:817-818.

Purugganan MD. 1997. The MADS-box floral homeotic gene lineages predate the origin of seed plants: Phylogenetic and molecular clock estimates. J. Mol. Evol. 45:392-396.

Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF. 1995. Molecular evolution of flower development - diversification of the plant mads-box regulatory gene family. Genetics 140:345-356.

Riechmann JL, Wang MQ, Meyerowitz EM. 1996. DNA-binding properties of arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucleic Acids Res. 24:3134-3141.

Rijpkema AS, Vandenbussche M, Koes R, Heijmans K, Gerats T. 2010. Variations on a theme: Changes in the floral ABCs in angiosperms. Semin. Cell Dev. Biol. 21:100-107.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Sablowski R. 2010. Genes and functions controlled by floral organ identity genes. Semin. Cell Dev. Biol. 21:94-99.

Sambrook J, Russell DW. 2006. Dot and Slot Hybridization of Purified RNA. Cold Spring Harb. Protoc. doi:10.1101/pdb.prot4052

Schwarzsommer Z, Hue I, Huijser P, Flor PJ, Hansen R, Tetens F, Lonnig WE, Saedler H, Sommer H. 1992. Characterization of the antirrhinum floral homeotic MADS-box gene *deficiens* - evidence for DNA-binding and autoregulation of its persistent expression throughout flower development. EMBO J. 11:251-263.

Souer E, AB Rebocho, M Bliek, E Kusters, RAM de Bruin and R Koes. 2008, Patterning of inflorescences and flowers by the F-Box protein DOUBLE TOP and the LEAFY homolog ABERRANT LEAF AND FLOWER of Petunia. Plant Cell 20: 2033–2048.

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

Stellari GM, Jaramillo MA, Kramer EM. 2004. Evolution of the APETALA3 and PISTILLATA lineages of MADS-box-containing genes in the basal angiosperms. Mol. Biol. Evol. 21:506-519.

Theissen G, Saedler H. 2001. Plant biology - floral quartets. Nature 409:469-471.

Theissen G, Kim JT, Saedler H. 1996. Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. J. Mol. Evol. 43:484-516.

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H. 2000. A short history of MADS-box genes in plants. Plant Mol. Biol. 42:115-149.

Trobner W, Ramirez L, Motte P, Hue I, Huijser P, Lonnig WE, Saedler H, Sommer H, Schwarzsommer Z. 1992. *Globosa* - a homeotic gene which interacts with *deficiens* in the control of antirrhinum floral organogenesis. EMBO J. 11:4693-4704.

van Dijk ADJ, van Ham RCHJ. 2010. Conserved and variable correlated mutations in the plant MADS protein network. BMC Genomics. 11:607-631.

Vandenbussche M, Zethof J, Royaert S, Weterings K, 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.

Viaene T, Vekemans D, Irish VF, Geeraerts A, Huysmans S, Janssens S, Smets E, Geuten K. 2009. Pistillata-duplications as a mode for floral diversification in (basal) asterids. Mol. Biol. Evol. 26:2627-2645.

Wikstrom N, Savolainen V, Chase MW. 2001. Evolution of the angiosperms: Calibrating the family tree. Proceedings of the Royal Society B-Biological Sciences 268:2211-2220.

Xiang QY, Thomas DT. 2008. Tracking character evolution and biogeographic history through time in cornaceae - does choice of methods matter? Journal of Systematics and Evolution 46:349-374.

Xiang QY, Thomas DT, Xiang QP. 2011. Resolving and dating the phylogeny of Cornales – effects of taxon sampling, data partitions, and fossil calibrations. Molecular Phylogenetics and Evolution, in press.

Xiang QY, Manchester SR, Thomas DT, Zhang WH, Fan CZ. 2005. Phylogeny, biogeography, and molecular dating of cornelian cherries (cornus, cornaceae): Tracking tertiary plant migration. Evolution 59:1685-1700.

Xiang QYJ, Thomas DT, Zhang WH, Manchester SR, Murrell Z. 2006. Species level phylogeny of the genus cornus (cornaceae) based on molecular and morphological evidence - implications for taxonomy and tertiary intercontinental migration. Taxon 55:9-30.

Yang YZ, Jack T. 2004. Defining subdomains of the K domain important for protein-protein interactions of plant MADS proteins. Plant Mol. Biol. 55:45-59.

Yang YZ, Fanning L, Jack T. 2003. The K domain mediates heterodimerization of the arabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. Plant Journal 33:47-59.

Yang ZH. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. 15:568-573.

Yang ZH. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. Computer Applications in the Biosciences 13:555-556.

Yang ZH, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol. Biol. Evol. 17:32-43.

Yang ZH, Wong WSW, Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22:1107-1118.

Yang Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586-1591.

Zahn LM, King HZ, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, dePamphilis CW, Ma H. 2005. The evolution of the SEPALLATA subfamily of MADS-box genes: A preangiosperm origin with multiple duplications throughout angiosperm history. Genetics 169:2209-2223.

Zhang W. 2008. Molecular evolution of floral homeotic B-class genes in the dogwood genus Cornus (Cornaceae) – gene duplication, selection, and coevolution. North Carolina State University.

Zhang W, Xiang QY, Thomas DT, Wiegmann BM, Frohlich MW, Soltis DE. 2008. Molecular evolution of PISTILLATA-like genes in the dogwood genus Cornus (cornaceae). Mol. Phylogenet. Evol. 47:175-195.

Table 1. Primers designed in the study for cloning B-class genes in *Cornus*

Primer name	Oligonucleotide sequences
3'-RACE CDS Primer A	AAGCAGTGGTATCAACGCAGAGTAC(T)30VN
5'-RACE CDS Primer A	(T)25VN
UPM Long	ctaatacgactcactatagggcAAGCAGTGGTATCAACGCAGAGT
UPM Short	ctaatacgactcactatagggc
ap3 uGSP F29	GGACAGGTGAGAGTTTGAACGATCTGAGC
ap3 uGSP R29	GCTCAGATCGTTCAAACTCTCACCTGTCC
ap3 5'dep F	RGAGAGARWATGGCTCGVGG
ap3 3'dep R	AGGTGCTATTMAAGCAGKGC
PI uGSP F29	AGTCTGGGAAGAGGTTGTGGGATGCTAAG
PI uGSP R29	CTTAGCATCCCACAACCTCTTCCCAGACT
PI 3uR20	CAAGTCTACATTCTGTCCTG
PI 5UF23	TCATGGGGAGAGGAAAGATAGAG

Table 2 Primers designed in the study for RT-PCR analyses in *Cornus*

Gene name	Primer name	Oligonucleotide sequences	Tm (°C)
CorfloPI-A	FAF	GAGAA TTTGAGCAATGAAATTGAT	65
	FAR	GATGGAATCCATCCTCCATCT	
CorfloPI-B	FBF	ACAACGGGCTTGATAGCATTG	65
	FBR	CTGCTGATATGTTTTCTCCATTTCA	
CorcanPI-A	CAF	CACTCGATAATGGTCTTGCAAT	65
	CAR	CTCTCACTCTCTGGTGGAATT	
CorcanPI-B	CBF	GAGACT GGGCTTGTGGGTGTT	68
	CBR	CTCTGATGATGATATTCTTTCTCCA CAT	
CoroffPI-A	OAF	TGCATGAGTACTGCAGCCCTGCA	68
	OAR	CCTGAGCTCAATCCGCATGC	
CoroffPI-B	OBF	GCATGAGTACTGCAGCCCTTCC	68
	OBR	GGTGCCTTAGCTCAATCTGCATCT	
Fully spliced *CorcanAP3*	AP3 cana LSF	GGAAATTAGGCAAAGGACAGG	62
	AP3 cana LR	CCTCCACCTGAGAATCCAAG	
Type II *CorcanAP3*	AP3 cana LSF	GGAAATTAGGCAAAGGACAGG	65
	AP3 cana SR	GAGGCCCTCCACAAATTCAT	
Fully spliced *CorfloAP3*	AP3 flo. LF	GGAAATTAGGCAAAGGACG	62
	AP3 flo. LR	CTCCACCTGTGAATCCAAGA	
Type II *CorfloAP3*	AP3 flo. SF	TTAGGCAAAGGACGGGTGA	62
	AP3 flo. SR	CTAGCGGCTCCACGAATTCA	
Fully spliced *CoroffAP3*	AP3 off. LSF	ACTTTGGGGATCGATCTGTG	62
	AP3 off. LR	CGTAATCGGCTCCATTATCA	
Type II *CoroffAP3*	AP3 off. LSF	ACTTTGGGGATCGATCTGTG	62
	Ap3 off. SR	GAGGATCTCCACGAATTCATTT	
Type I *CoroffAP3*	AP3 off. IF	TGCACCACAGAATGCAAGAG	60
	AP3 off. IR	CGTAATCGGCTCCATTATCAA	
Fully spliced *CorconAP3*	AP3 con LF	CTTCAAGAAGGCCAATGAGC	62
	Ap3 con LSR	TTTCGTGAAGACGAAGCAGA	
Type III *CorconAP3*	AP3 con SF	TCAACCAACAGAACGAAGACA	65
	Ap3 con LSR	TTTCGTGAAGACGAAGCAGA	
Fully spliced *CormacAP3*	AP3 mac LSF	GCGATGCGAAGGTCTCTATC	62
	AG3 mac LR	TTATCTCGAAATCATGGAGGAG	
Type IV *CormacAP3*	AP3 mac LSF	GCGATGCGAAGGTCTCTATC	62
	AP3 mac.sang SR	CCTCTCTTATTATCTCCTTTCGATCA	
Fully spliced *CorsanAP3*	AP3 sang LSF	CTTTGGGGATTGATCTGTGG	62
	AP3 sang LR	TCGAAATTGTGGAGGAGGTT	
Type IV *CorsanAP3*	AP3 sang LSF	CTTTGGGGATTGATCTGTGG	62
	AP3 mac.sang SR	CCTCTCTTATTATCTCCTTTCGATCA	

Table 2 Continued

26S	26S 12 26S 2782R	GTCCTAAGATGAGCTC GGTAACTTTTCTGACACCTC	55

Table 3. Characters and character states of bract development and B class gene expression and evolution that are variable among species and analyzed for evolutionary history on Mesquite 2.74

Species	1	2	3	4	5	6	7	8	9	10
C.canadensis (DW)	1	1	0	0	1	0	0	0	1	0
C.florida (BB)	0	1	1	1	2	1	0	1	1	1
C.officinallis (CC)	0	0	0	0	1	0	?	2	0	?
C.macrophylla (BW)	1	0	0	0	0	0	?	0	0	?
C.sanguinea (BW)	1	0	0	0	0	0	?	0	0	?
C.controversa (BW)	1	0	0	0	0	0	?	3	0	?

1, bract origin: bract initiated from outside of inflorescence meristem, 0; bract initiated from primary inflorescence branch meristem, 1;
2, bract petaloidy: no petaloidy, 0; petaloidy, 1;
3, expression of *CorPI*-A in unexpended bracts (B1): not expressed, 0; weakly expressed, 1;
4, expression of *CorPI*-B in unexpended bracts (B1): not expressed, 0; strongly expressed, 1;
5, expression of *CorAP3* in unexpended bracts (B1): not expressed, 0; weakly expressed, 1; strongly expressed, 2;
6, amino acid of CorPI-A at the site 186 under positive selections and predicted to co-evolve with CorAP3: V, 0; L, 1;
7, amino acid of CorPI-B at the site 186 under positive selections and predicted to co-evolve with CorAP3: Q, 0;
8, amino acid of CorAP3 at the site 197 under positive selections and predicted to interact with CorPI: G, 0; A, 1; D, 2; P; 3
9, amino acid of CorPI-A at the site 182 predicted to co-evolve with CorAP3: Y, 0; F, 1;
10, amino acid of CorPI-B at the site 182 predicted to co-evolve with CorAP3: Y, 0; T, 1

Table 4. Estimation of dN/dS ratio (ω) for *CorPI*-A (ω_1), *CorPI*-B (ω_2) and outgroups (ω_0) under different branch models using PAML 4.4

Models	p	lnL	K	ω_0	ω_1	ω_2
A. One-ratio: $\omega_0 = \omega_1 = \omega_2$	66	-2203.62	1.46	0.238	0.238	0.238
B. Two-ratio: $\omega_0 = \omega_2$, ω_1	67	-2203.33	1.41	0.252	0.387	0.252
C. Two-ratio: $\omega_0 = \omega_1$, ω_2	67	-2202.22	1.40	0.244	0.244	0.710
D. Three-ratio: ω_0, ω_1, ω_2	68	-2201.85	1.41	0.238	0.385	0.714
E. Free ratio	129	-2189.99	1.42	See Fig.7		

Table 5. Results of likelihood ratio tests (LRTs) for significance of dN/dS ratio (ω) differences between CorPI-A (ω_1), CorPI-B (ω_2) and outgroups (ω_0) in PAML 4.4

Null hypothesis	Assumption made	Models compared	2XΔl
I. $\omega_0 = \omega_1 = \omega_2$	ω_2 free	A and C	2.8* (df=1)
II. $\omega_0 = \omega_1 = \omega_2$	ω_1, ω_2 free	A and D	3.64 (df=2)
III. $\omega_0 = \omega_1 = \omega_2$	Free ratio	A and E	43.26 (df=63)

* indicates significantly fitted model at >90% level.

Table 6. Results from analyses of positive selections in *CorPI*-A and *CorPI*-B using branch-site model in PAML4.4

Branch	Model	lnL	Positive selected sites	2XΔl
CorPI-A	Model A Null model	-2160.07 -2160.07	None	0
CorPI-B	Model A Null model	-2157.42 -2159.15	K1 (L^{92}, R^{100}, Q^{107}), C (V^{178}, E^{180}, Q^{*186}, E^{188}, S^{202}, D^{208})	$3.46^{\#}$ (df=1)

The detected positive selected sites are mapped to functional domains (see Fig. 10 A), K1: the first amphipathic α-helices in K domain, C: C-terminal domain. Letters within parentheses follow amino acids sequences of *CorPI*-B in *C.florida*. Numbers in superscript indicate the amino acid position in CorPI alignment. An asterisk * indicates sites potentially under positive selection under site models with posterior probabilities > 90%. # indicates significantly fitted model at >90% level.

Table 7. Estimation of dN/dS ratio (ω) for *CorAP3* in DW group (ω_1), ancestor of BW and DW groups (ω_2) and the rest of groups (ω_0) under different branch models using PAML 4.4

Models	p	L	K	ω_0	ω_1	ω_2
A. One-ratio: $\omega_0 = \omega_1 = \omega_2$	16	-2021.20	1.87	0.309	0.309	0.309
B. Two-ratio: $\omega_0 = \omega_2, \omega_1$	17	-2020.23	1.86	0.325	0.175	0.175
C. Three-ratio: $\omega_0, \omega_1, \omega_2$	18	-2020.23	1.86	0.307	0. 176	0.290
D. Free ratio	33	-2016.77	1.89	See Fig.8		

Table 8. Results of likelihood ratio tests (LRTs) for significance of dN/dS ratio differences for *CorAP3* in DW group (ω_1), ancestor of BW and DW groups (ω_2) and the rest of groups (ω_0) under different branch models using PAML 4.4

Null hypothesis	Assumption made	Models compared	2XΔl
I $\omega_0 = \omega_1 = \omega_2$	ω_1 free	A and B	1.94 (df=1)
II. $\omega_0 = \omega_1 = \omega_2$	ω_1, ω_2 free	A and C	1.94 (df=2)
III. $\omega_0 = \omega_1 = \omega_2$	Free ratio	A and D	8.86 (df=17)

Table 9. Results from analyses of positive selections in *CorAP3* using site model in PAML4.4

Models	p	lnL	Estimates of parameters	Positively selected sites
M0: one-ratio	12	-1654.82	ω=0.21, K=1.52	None
M1a: Nearly Neutral	13	-1648.73	p_0=0.90, ω_0=0.13, K=1.49	None
M2a: Positive Selection	15	-1644.89	p_0=0.91, p_1=0.084 (p_2=0.0045), ω_0=0.14, ω_2=31.76, K=1.55	I (Q^{66}), C (A^{*197}, G^{211})
M3: Discrete	16	-1644.86	p_0=0.857, p_1=0.138 (p_2=0.0045), ω_0=0.12, ω_1=0.76, ω_2=31.62, K=1.55	C (G^{**211})
M7: β	13	-1649.70	p=0.388, q=1.350, K=1.51	None
M8: β&ω	15	-1644.94	p_0=0.995 (p_2=0.0045), p=0.602, q=2.19, ω=31.61, K=1.55	I (Q^{66}), C (A^{*197}, G^{211})

P is the number of free parameters in the ω distribution. Parameters in parentheses are not free.

Position of positive selected sites was mapped into functional domains (see Fig. 10 B), I: intervening domain, C: C-terminal domain; Letters within parentheses followed amino acids sequences of *CorAP3* in *C.florida;* numbers in superscript indicated the amino acid position in *CorAP3* alignment

* indicates sites potentially under positive selection under site models with posterior probabilities > 90%.

** indicates ites sites potentially under positive selection under site models with posterior probabilities > 99%.

Table 10. Results from comparison of three pairs of site models using likelihood ratio tests (LRTs) in PAML4.4

Models compared	df	$\chi 2$ critical value (5%)	$\chi 2$ critical value (1%)	2XΔl
M0 and M3	4	9.49	13.28	19.91**
M1a and M2a	2	5.99	9.21	7.68*
M7 and M8	2	5.99	9.21	9.52**

* indicates significantly fitted model at >95% level.
** indicates significantly fitted model at >99% level.

Table 11. Summary of expression pattern of *CorPI*-A, *CorPI*-B and *CorAP3* in various developmental stages of flowers and bracts in *Cornus.*

Stages	Genes	*C.can* (DW)	*C.flo* (BB)	*C.off* (CC)	*C.mac* (BW)	*C.san* (BW)	*C.con* (BW)
Early flowers[1]	*CorPI*-A	+++	+	+++	?	?	?
	CorPI-B	+	+++	-	?	?	?
	CorAP3	++	++	++	?	?	?
Late flowers[2]	*CorPI*-A	++	++	++	+++	+++	+++
	CorPI-B	+++	+++	+++	-	-	-
	CorAP3	+++	+++	+++	+++	+++	+++
Expanding bracts[3]	*CorPI*-A	-	+	-	-	-	-
	CorPI-B	-	+++	-	-	-	-
	CorAP3	+	+++	+	-	-	-

[1]Eerly flowers refer to floral initiation and organogenesis, data based on *in situ* hybridization; [2]Late flowers refer to F1, F2 and F3 stages, data based RT-PCR; [3]data based on RT-PCR at B1 stages. *C.can* = *C.canadensis*; *C.flo* = *C.florida*; *C.off* = *C.officinallis*; *C.mac* = *C.macrophylla*; *C.san* = *C.sanguinea*; *C.con* = *C.controversa*; DW: dwarf dogwoods; BB: big-bracted dogwoods; CC: cornelian cherry dogwoods; BW: blue- or white- fruited dogwoods.

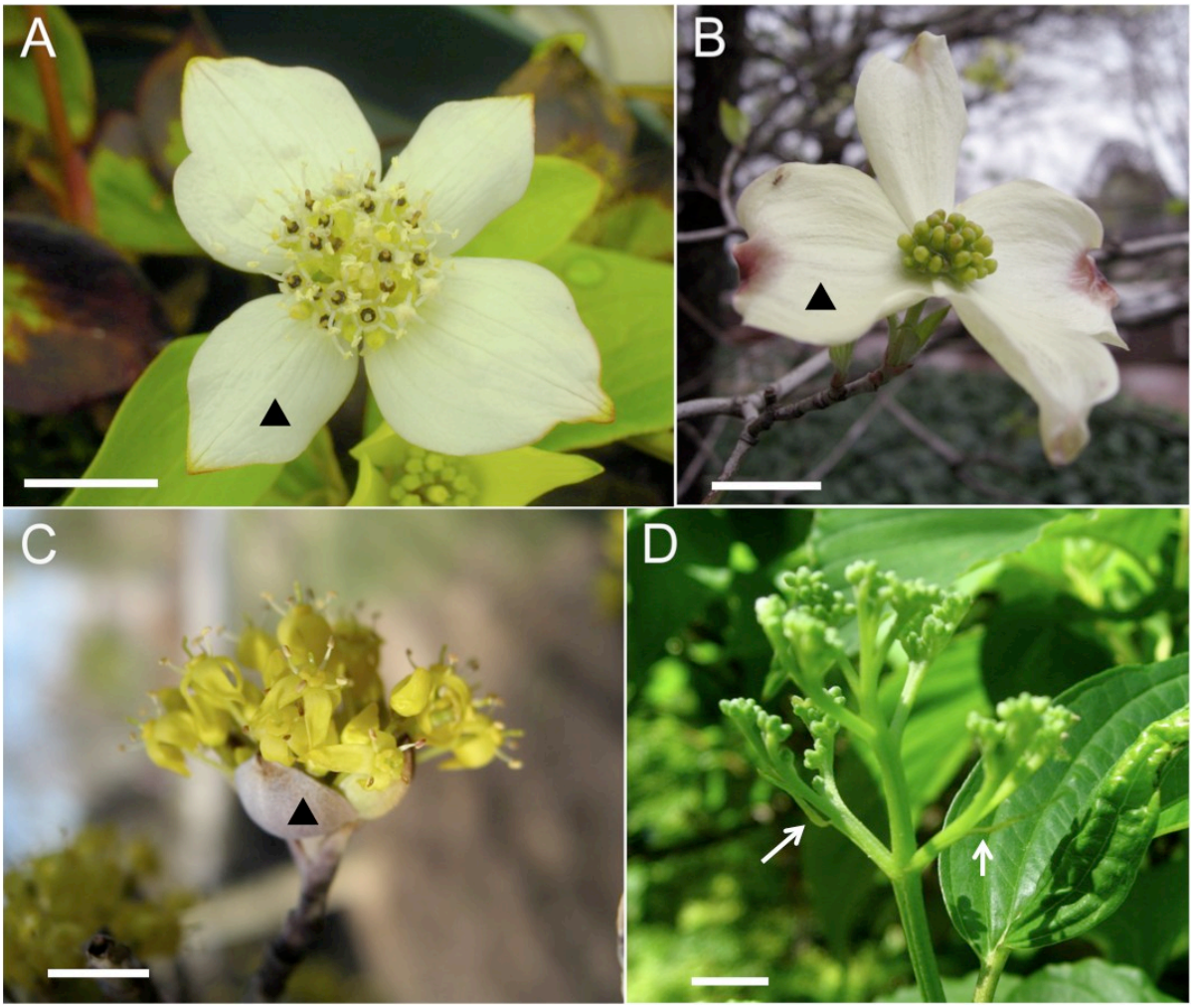
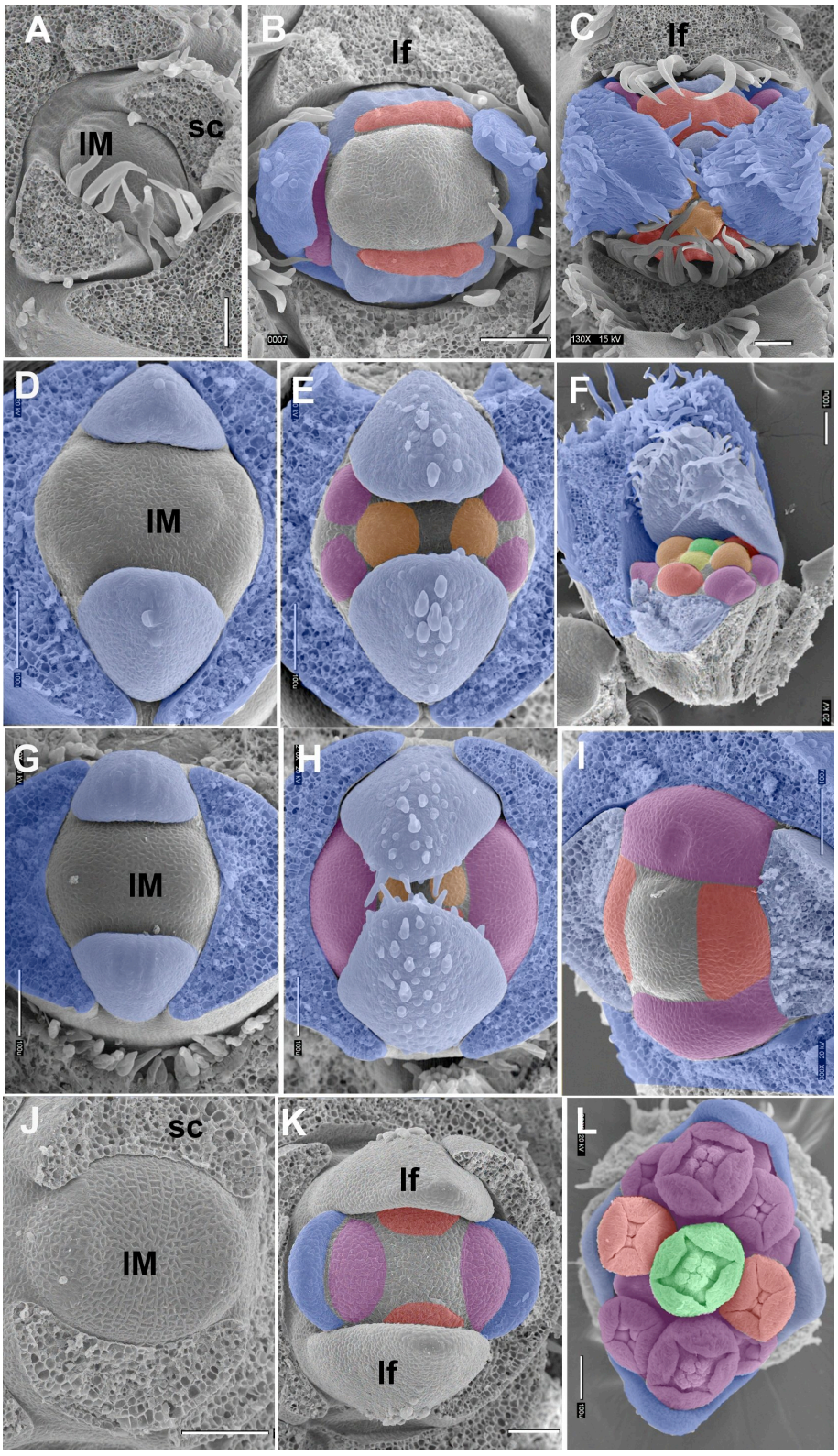


Fig. 1 Bract diversity in dogwood (*Cornus*). A: *C. canadensis* (DW group); B: *C. florida* (BB group); C: *C.officinallis* (CC group); D: *C. macrophylla* (BW group); Black triangle: 4-fold bracts; White arrow: early deciduous rudimentary bracts; BW, blue- or white-fruited dogwoods; DW, dwarf dogwoods; BB, big-bracted dogwoods; CC, cornelian cherry; Scale bars: 1 cm.

Fig. 2 Initiation and early development of bracts in dogwood (*Cornus*). A-C: bract initiation and development in *C. macrophylla*; D-F: bract initiation and development in *C. officinallis*; G-I: bract initiation and development in *C. florida*; J-L: bract initiation and development in *C. canadensis*; Blue--bract primordia; Purple--the first group of primary inflorescence branch meristem; Salmon—the second group of primary inflorescence branch meristem; Orange—the third group of primary inflorescence branch meristem; Yellow—the fourth group of primary inflorescence branch meristem; Green—the terminal flower; IM-inflorescence meristem; sc: scales; lf: leaf; Scale bars: 100 nm.


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

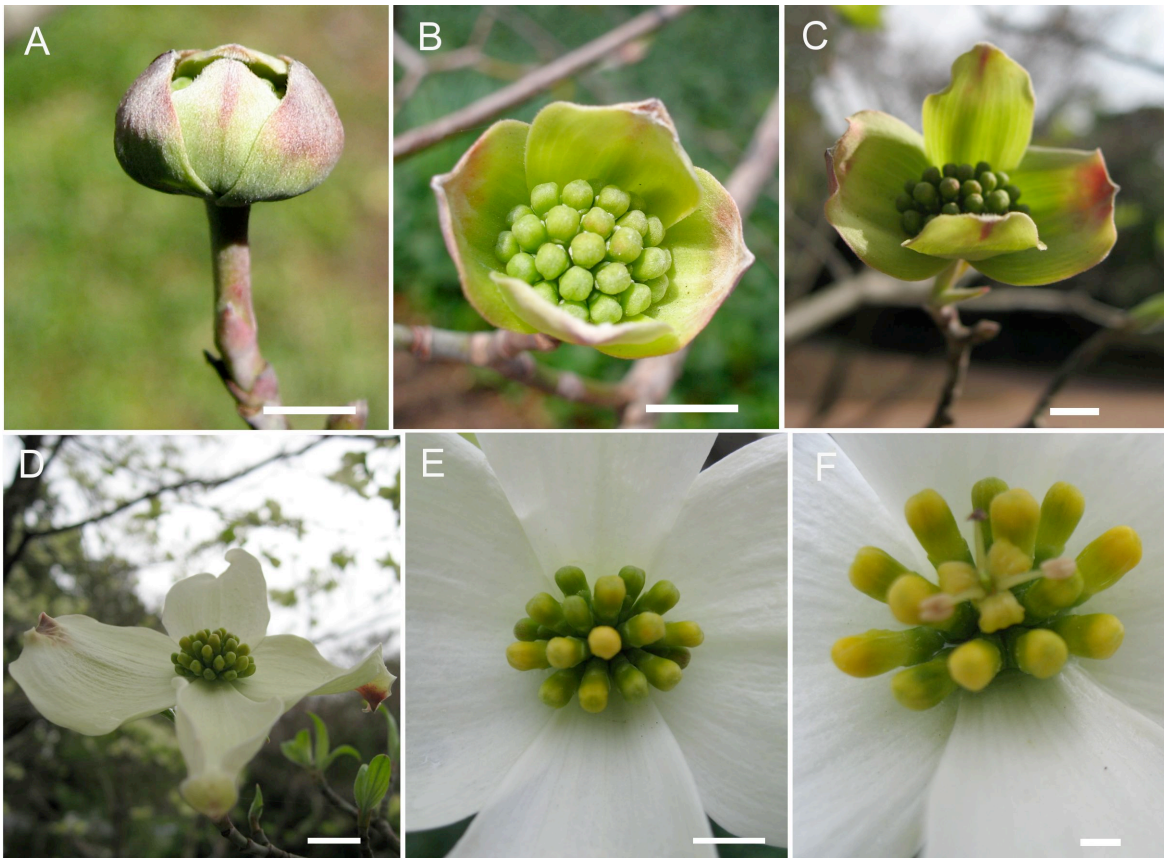


Fig. 3 Bract late development in *C. florid*a. A: Stage 1, bud burst; B: Stage 2, bract open; C: Stage 3,bract expand; D: Stage 4, bract whiten; E: Stage 5, flower color change from green to yellow-green; F: Stage 6, first flower open. Scale bars in A-C, E: 0.5 cm; Scale bar in D: 1.0 cm; Scale bar in F: 0.1 cm.

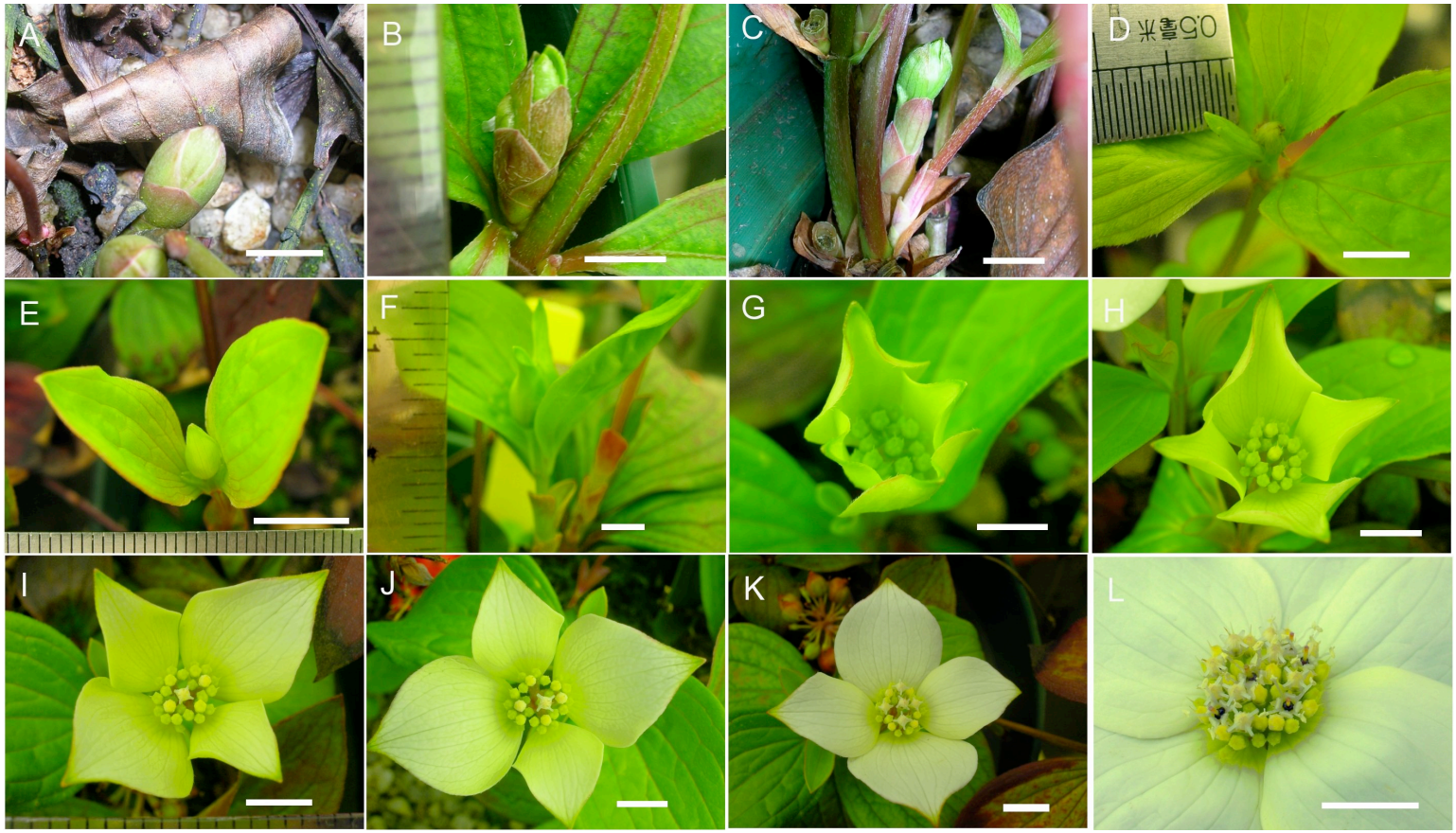


Fig. 4 Bract late development in *C. canadensis.* A: Stage 1, bud enlarge; B: Stage 2, scale open; C: Stage 3, leaves out of scales; D: Stage 4, leaves open, bract visible; E: Stage 5, bract length 3-5 mm long; F: Stage 6, bract expand; G: Stage 7, central flower color change from green to white; H: Stage 8, outer bract whiten; I: Stage 9, inner bract whiten; J: Stage 10, bract complete white; K: Stage 11, first flower open; Scale bar in E: 1 cm, others 0.5 cm.

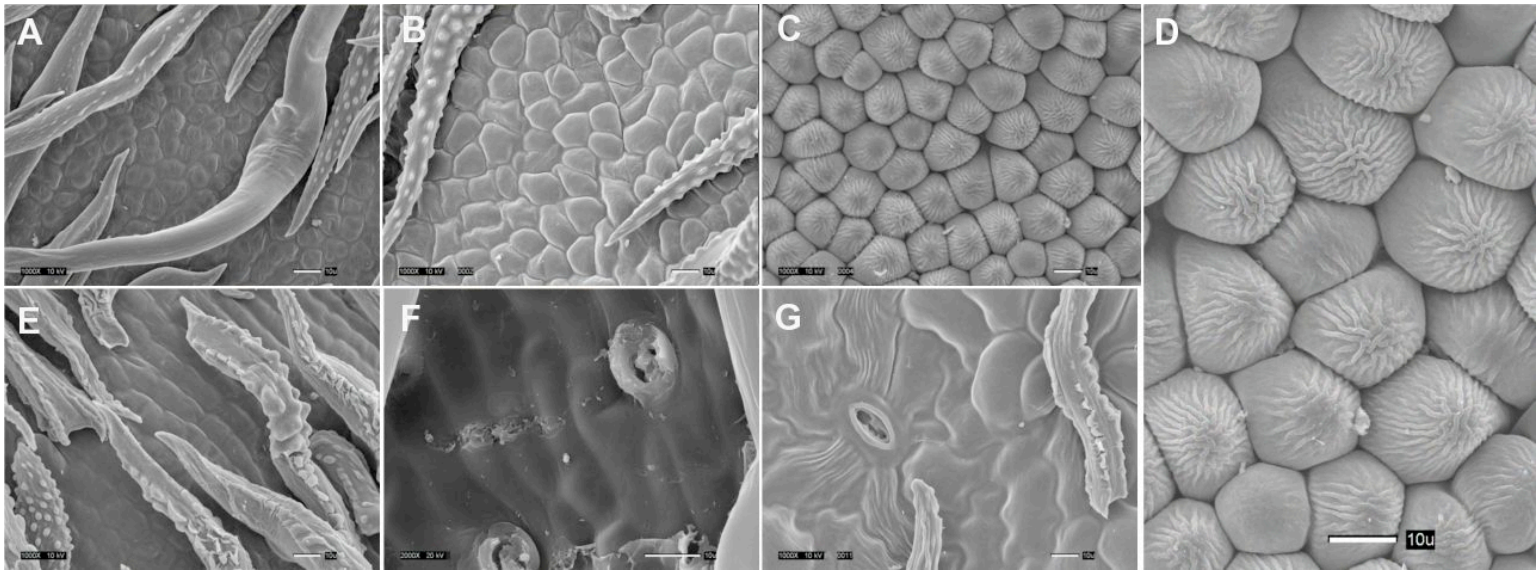


Fig. 5 Epidermal morphology of bracts during late development in *C.florida.* A-D: adaxial surface of bracts at stages 2, 3, 4, respectively; E-G: abaxial surface of bracts at stages 2, 3, 4, respectively; Scale bars: 100 nm.

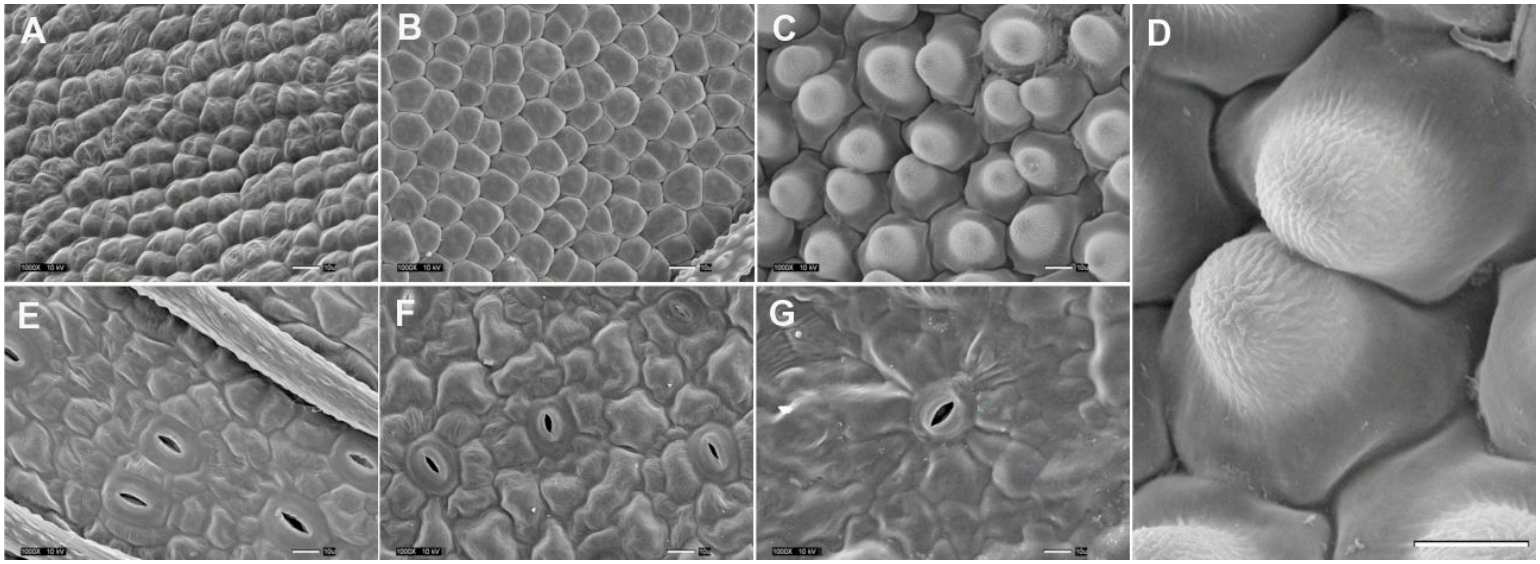


Fig. 6 Epidermal morphology of bracts during late development in *C.canadensis.* A-D: adaxial surface of bracts at stages 5, 6, 10, respectively; E-G: abaxial surface of bracts at stages 5, 6, 10, respectively; Scale bars: 100 nm.

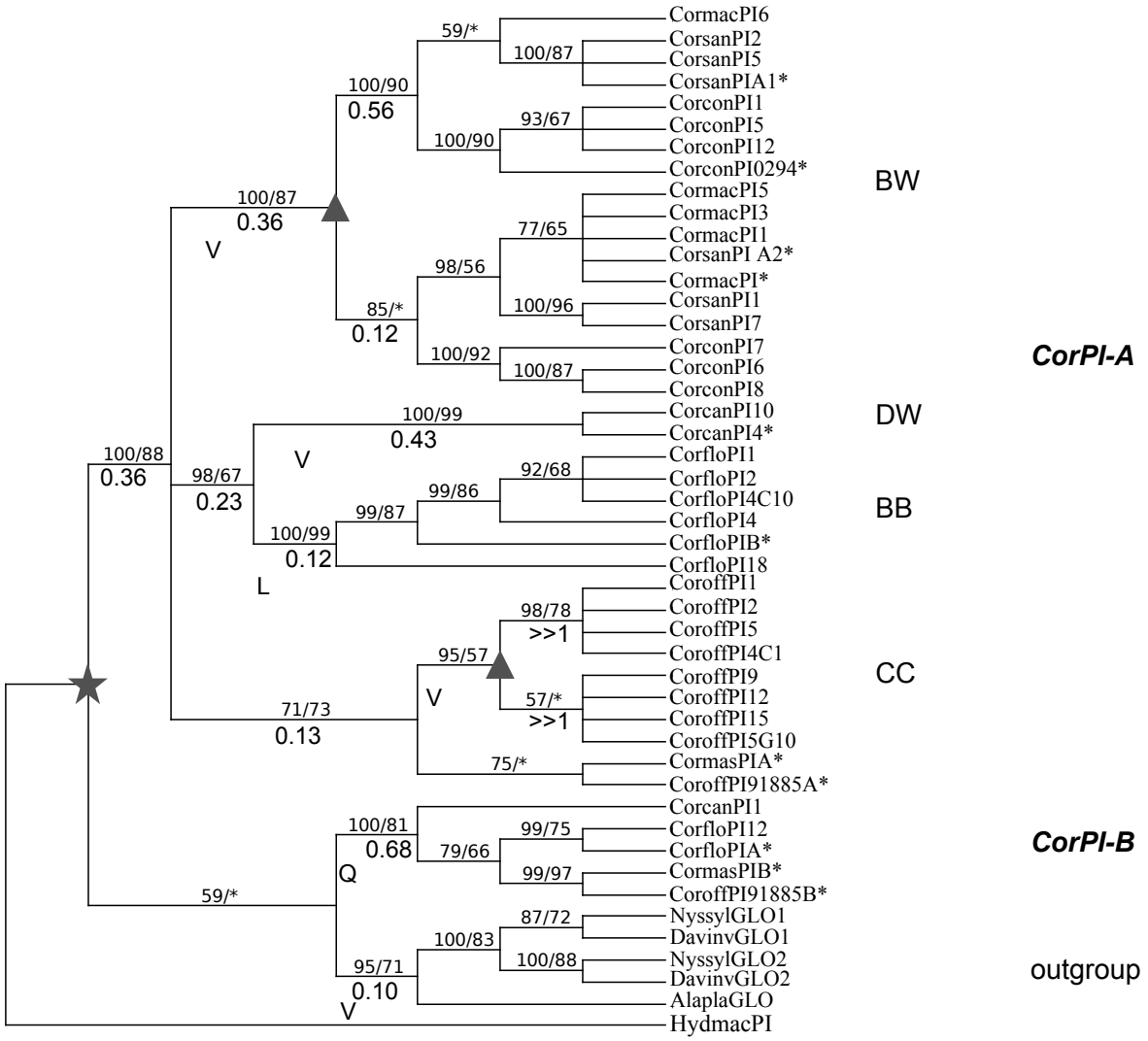


Fig. 7 Genealogy of *CorPI* inferred from Bayesian and Maximum Likelihood (ML) analyses. Numbers above branches are Bayesian posterior probabilities and bootstrap supports from ML analyses; Numbers below branches are ω values from free-ratio model analysis using PAML 4.4; Ticks on branches indicate the amino acid at the site 186 which is under positive selection using PAML 4.4; BW, blue- or white-fruited dogwoods; DW, dwarf dogwoods; BB, big-bracted dogwoods; CC, cornelian cherry. Star indicates an ancient gene duplication event; triangles indicate small-scale gene duplication events within each dogwood group. Sequences marked by an asterisk are partial sequences of the 5'end generated in a previous study (Zhang et al. 2008), sequences without marks are generated in this study.

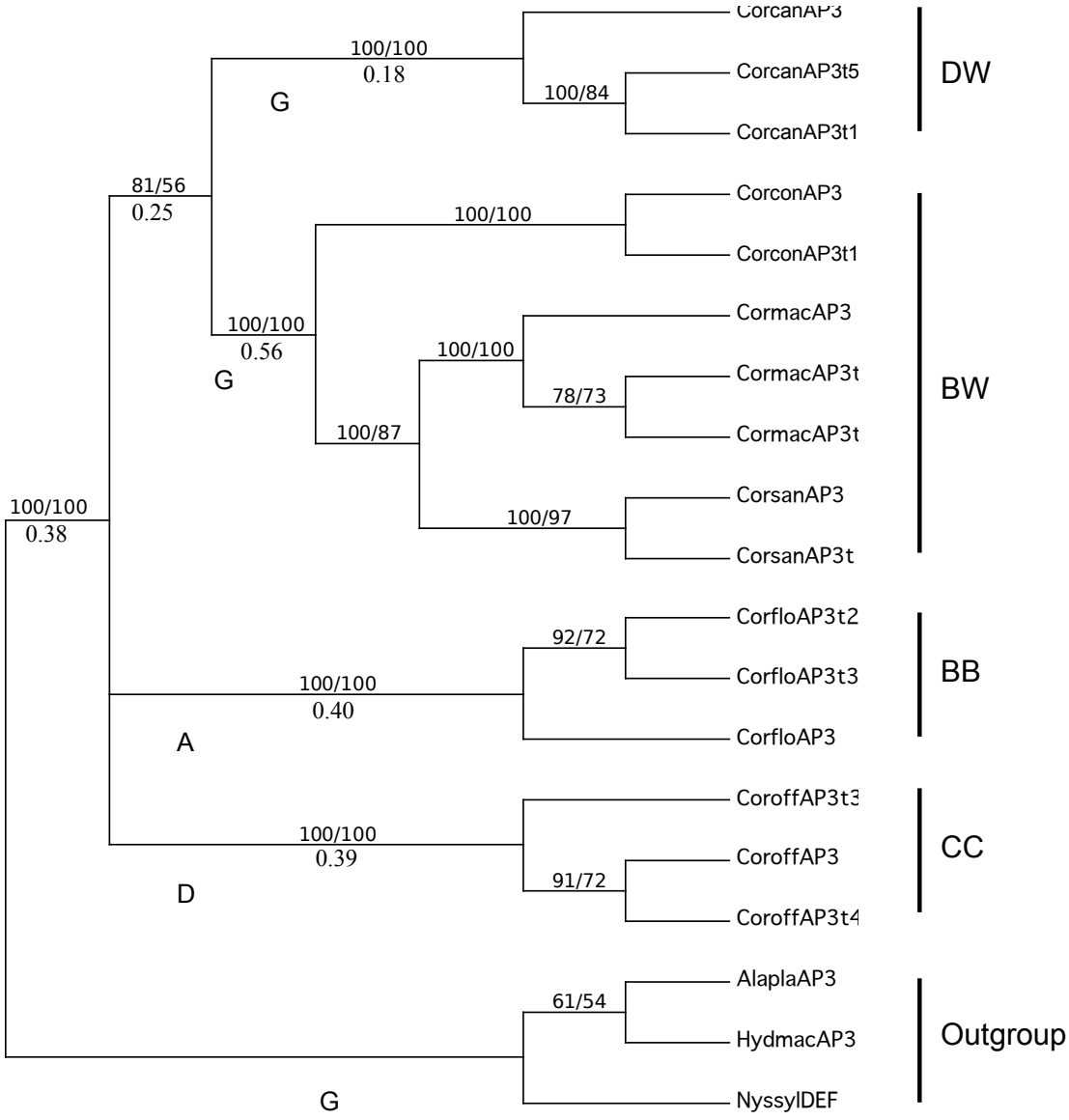


Fig. 8 Genealogy of *CorAP3* genes inferred from Bayesian and Maximum Likelihood (ML) analyses. Numbers above branches are Bayesian posterior probabilities and bootstrap supports for ML analyses; Numbers below branches are ω values from free-ratio model analysis using PAML 4.4; Ticks on branches indicate the amino acid at the site 197 which is under positive selection using PAML 4.4; BW, blue- or white-fruited dogwoods; DW, dwarf dogwoods; BB, big-bracted dogwoods; CC, cornelian cherry. Underlines indicate truncated transcripts.

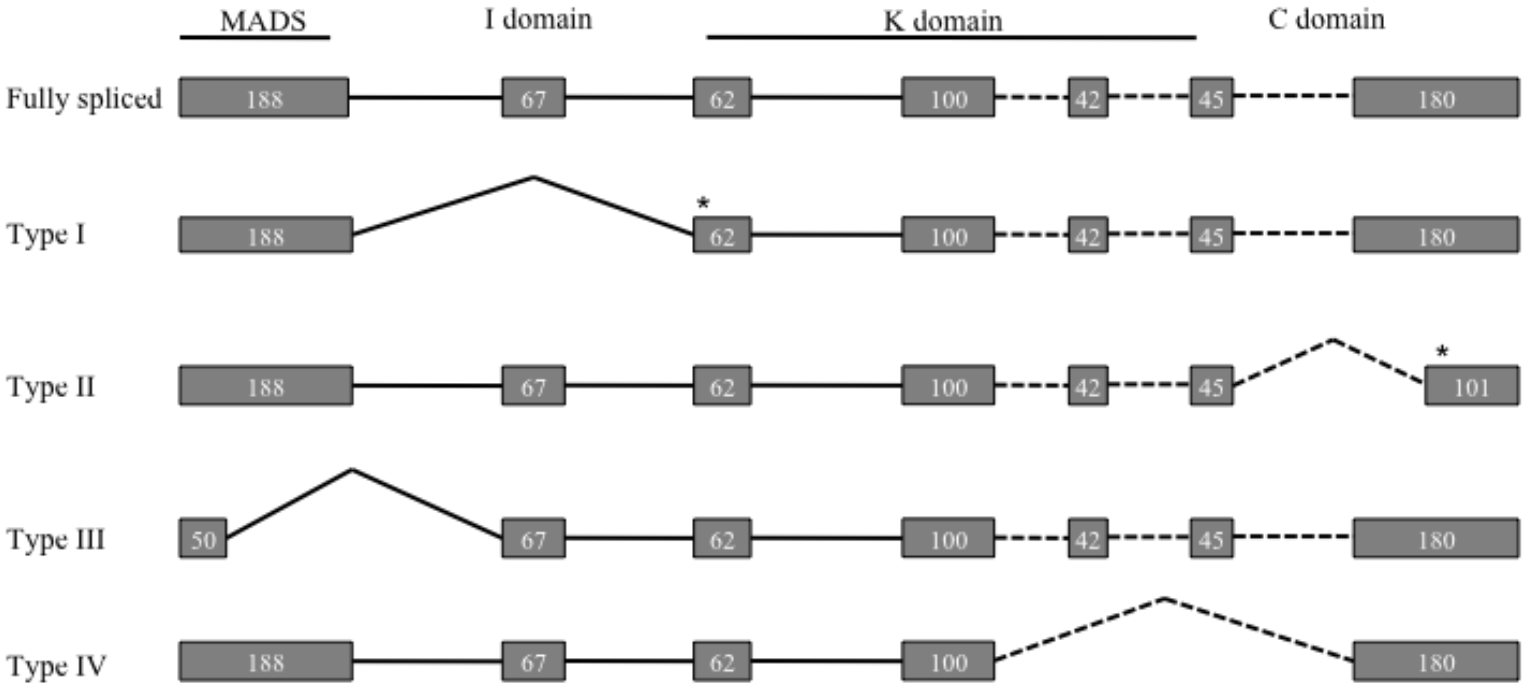


Fig. 9 Alternative spliced forms of *CorAP3* detected in *Cornus*. Exons are showed as open box. Numbers inside the boxes indicate length of exons in base pair numbers. Intervening solid lines between exons indicate confirmed introns from the genomic DNA sequences. Dashed lines indicate unconfirmed introns. The asterisks "*" indicate the position of premature stop codon.

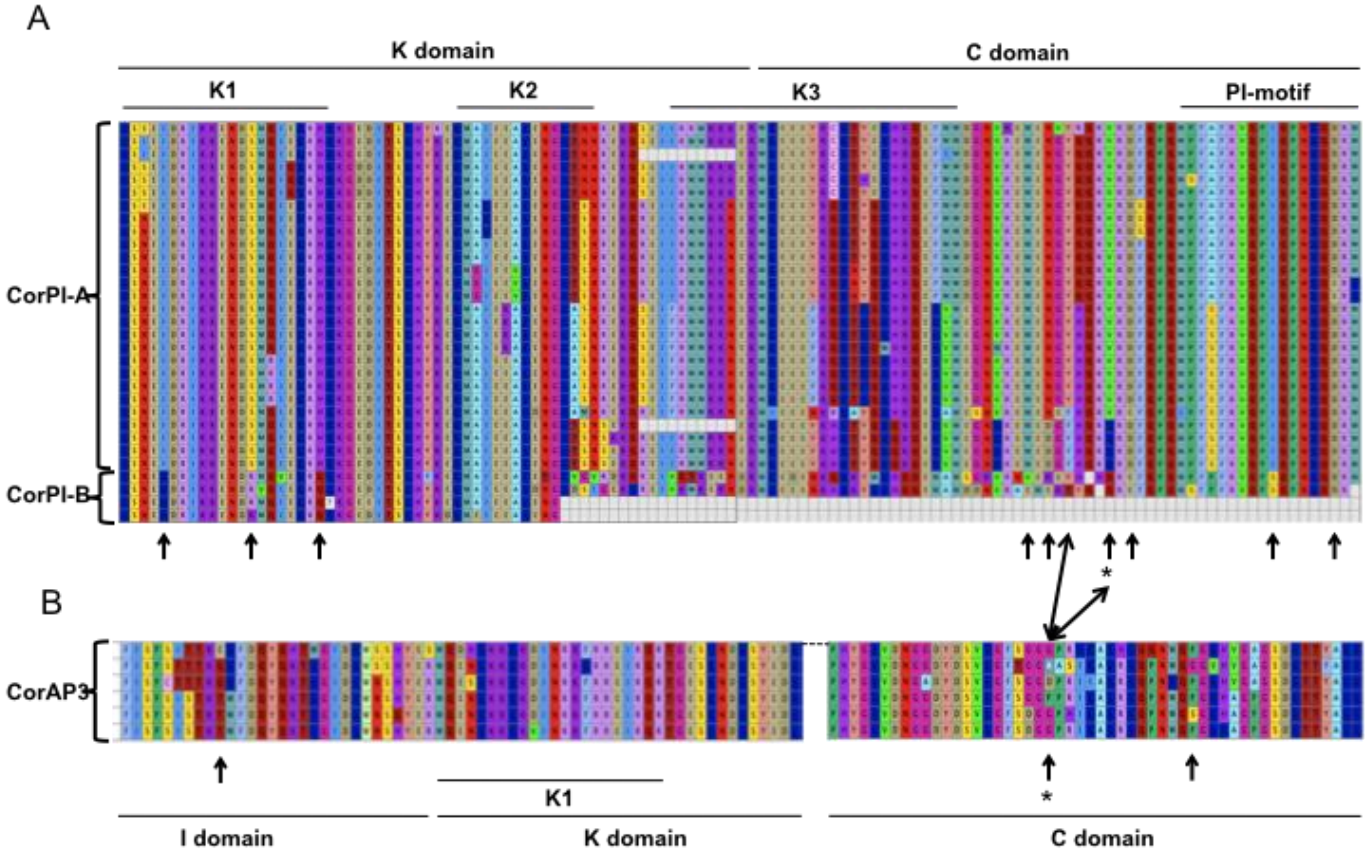
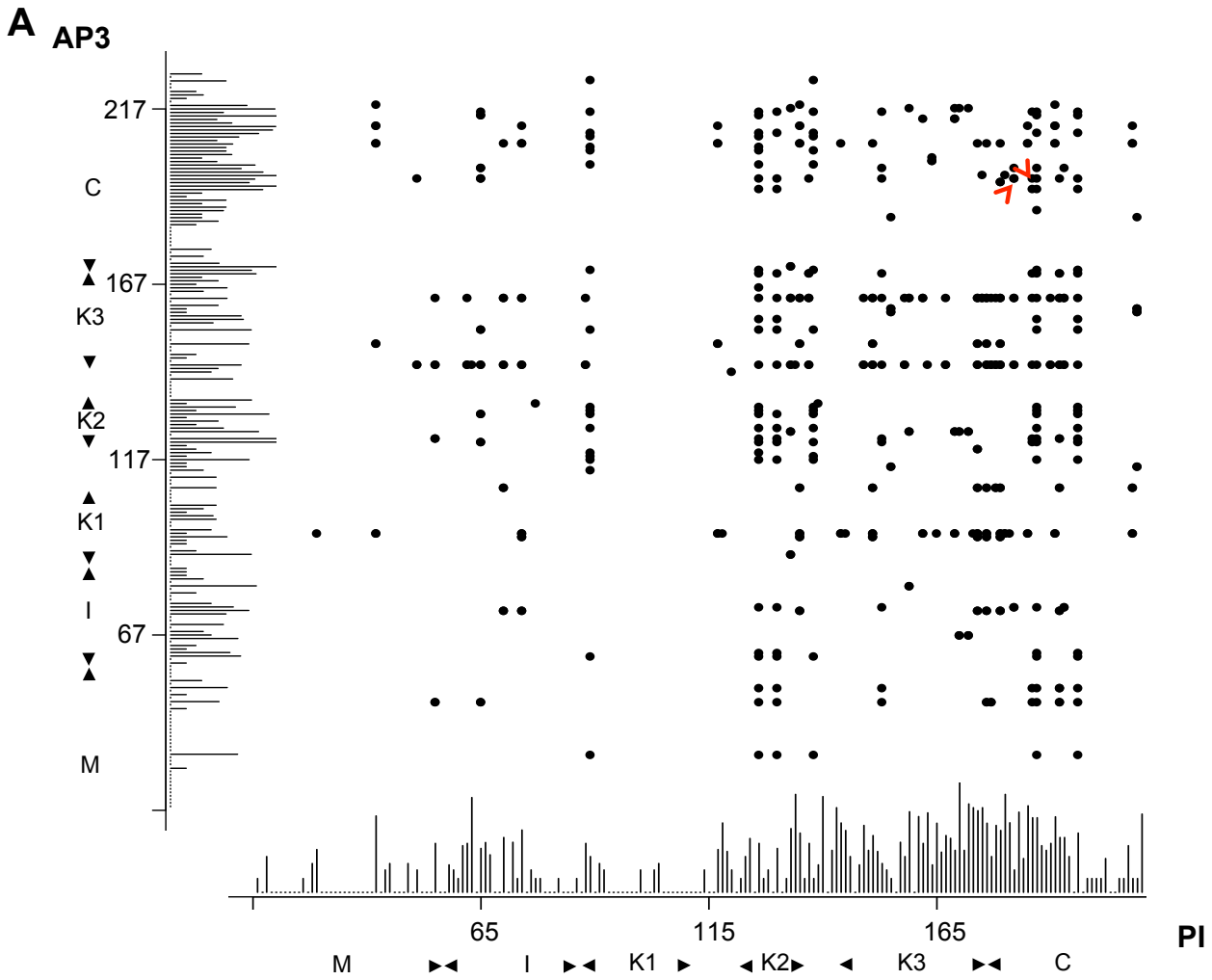
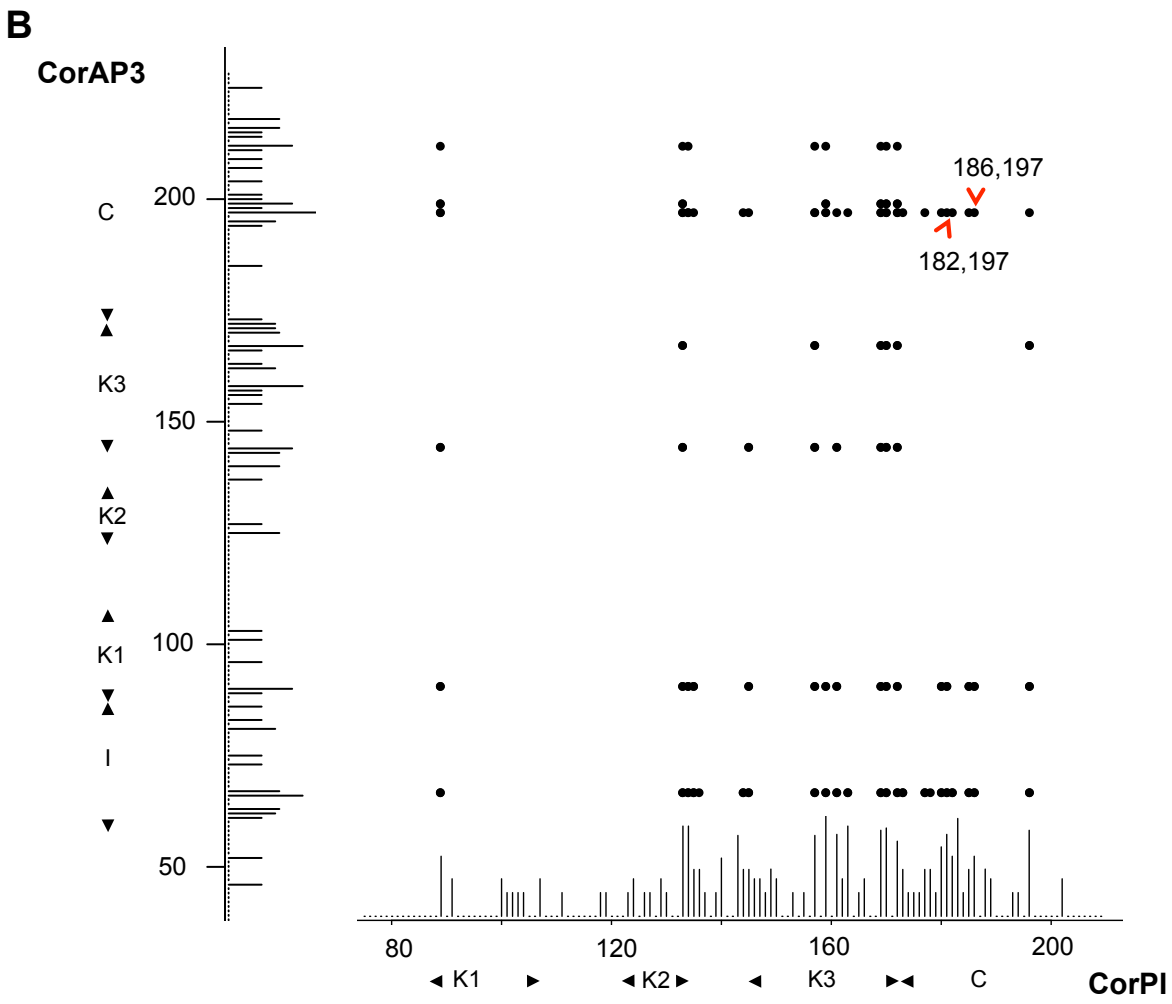


Fig. 10. Aligned amino acid sequences of *CorPI*-A and *CorPI*-B. Identical amino acid is shown in same color; black arrows indicate positive selection sites inferred from branch-site model (* indicates P >90%); Arrows indicate the sites predicted to interact with *CorAP3* using co-evolution analysis with the program CAPS.

Fig. 11 Amino acid sites identified to be ”co-evolving” between PI- and AP3- orthologies by analysis using CAPS. A: Plot of “coevolving” sites in PI- and AP3- orthologs; B: Plot of coevolving sites in CorPI and CorAP3 in *Cornus*. Arrows indicate the two co-evolving sites identified in both analyses of A and B. Numbers along the X-axis and Y-axis indicate the amino acid positions of PI ortholog and AP3 ortholog, respectively; Bar graphs above X-axis as well as at the right of Y-axis show the entropy value (reflecting the variation level of amino acids at the site in the data matrix) in each site of PI- ortholog and AP3- ortholog, respectively; each dot means the coevolving site between PI and AP3 orthologs; M: MADS domain; I: intervening domain; K: keritin-like domain; K1, K2 and K3: three amphipathic α-helices; C: C-terminal domain.


A
AP3
217
C
167
K3
K2
117
K1
I
67
M
65
115
165
PI
M
I
K1
K2
K3
C


B
CorAP3
C
K3
K2
K1
I
200
150
100
50
186,197
182,197
80
120
160
200
K1
K2
K3
C
CorPI

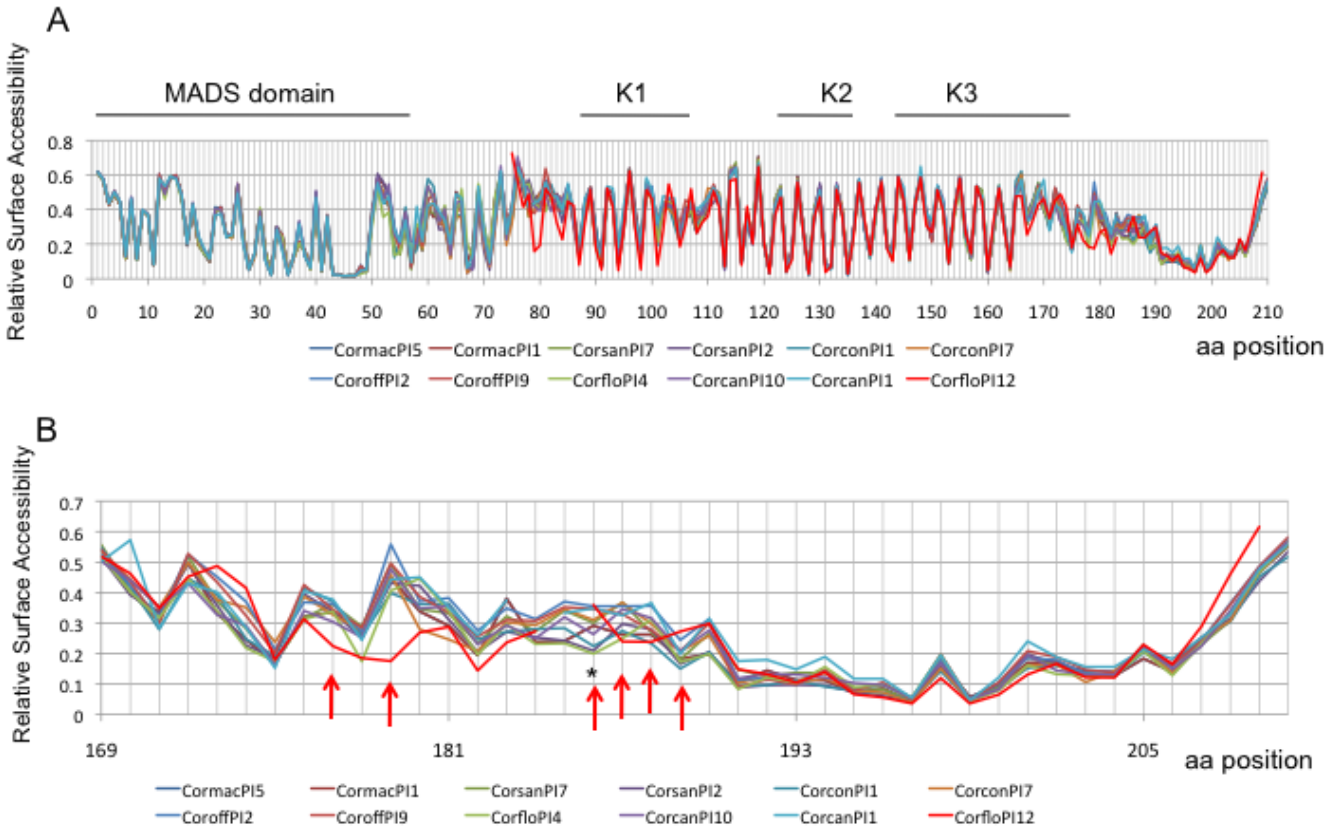


Fig. 12 Plot of relative surface accessibility of each amino acid residue in CorPI based on calculation of NetSurfP program. A shows plot of the whole CorPI proteins; B shows plot of the partial C domaion of CorPI proteins (from site 169 to site 210); Red arrows point amino acid sites showing altered relative surface accessibility in CorfloPI-B (CorfloPI12, red line); asterids marks the site (site 186) under positive selection and co-evolving with CorAP3.

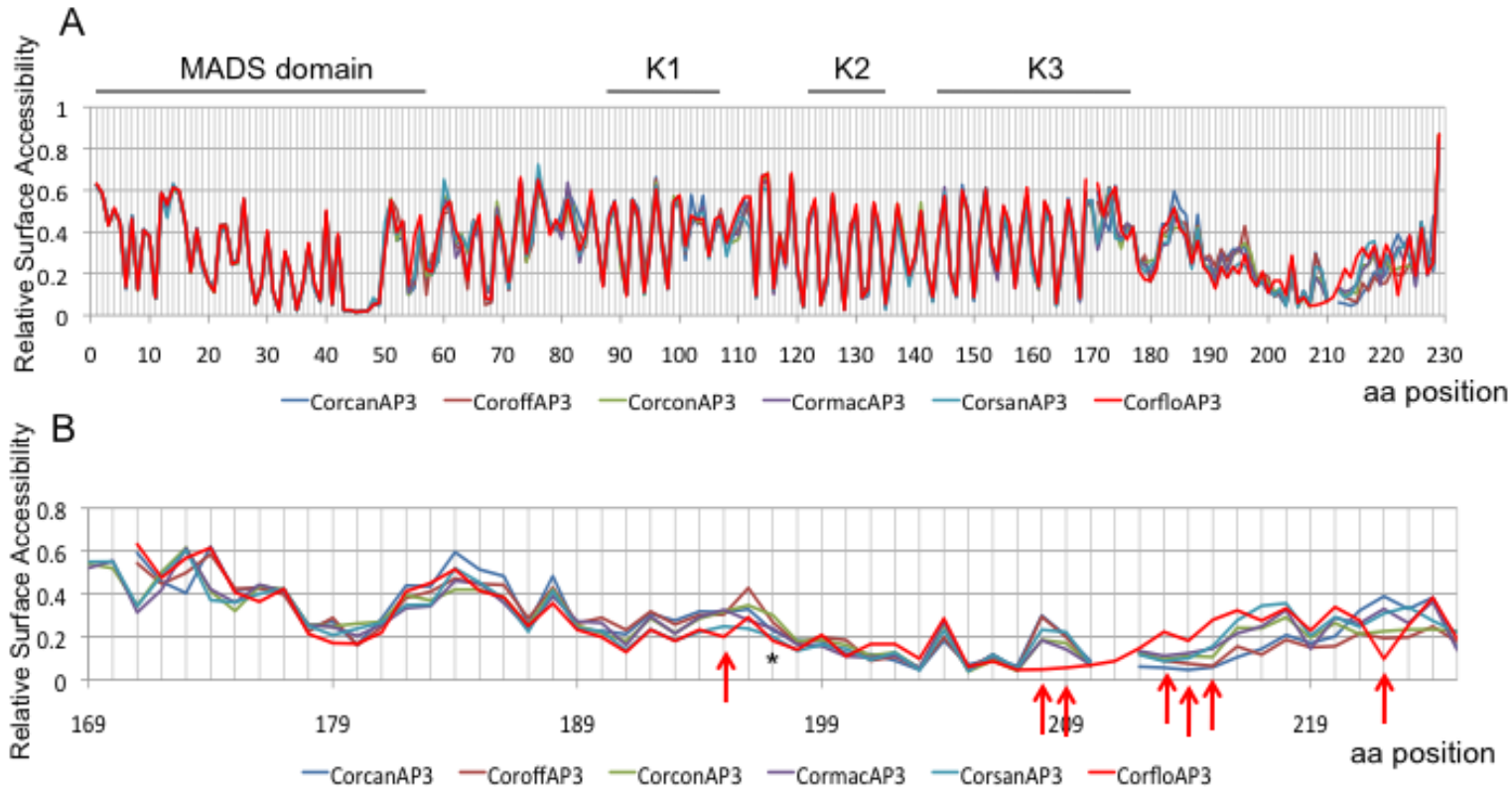


Fig. 13 Plot of relative surface accessibility of each amino acid residue in CorAP3 based on calculation of NetSurfP program. A shows plot of the whole CorAP3 proteins; B shows plot of the partial C domaion of CorAP3 proteins (from site 169 to site 225); Red arrows point amino acid sites showing altered relative surface accessibility in CorAP3 (red line); asterids marks the site (site 197) under positive selection and co-evolving with CorPI.

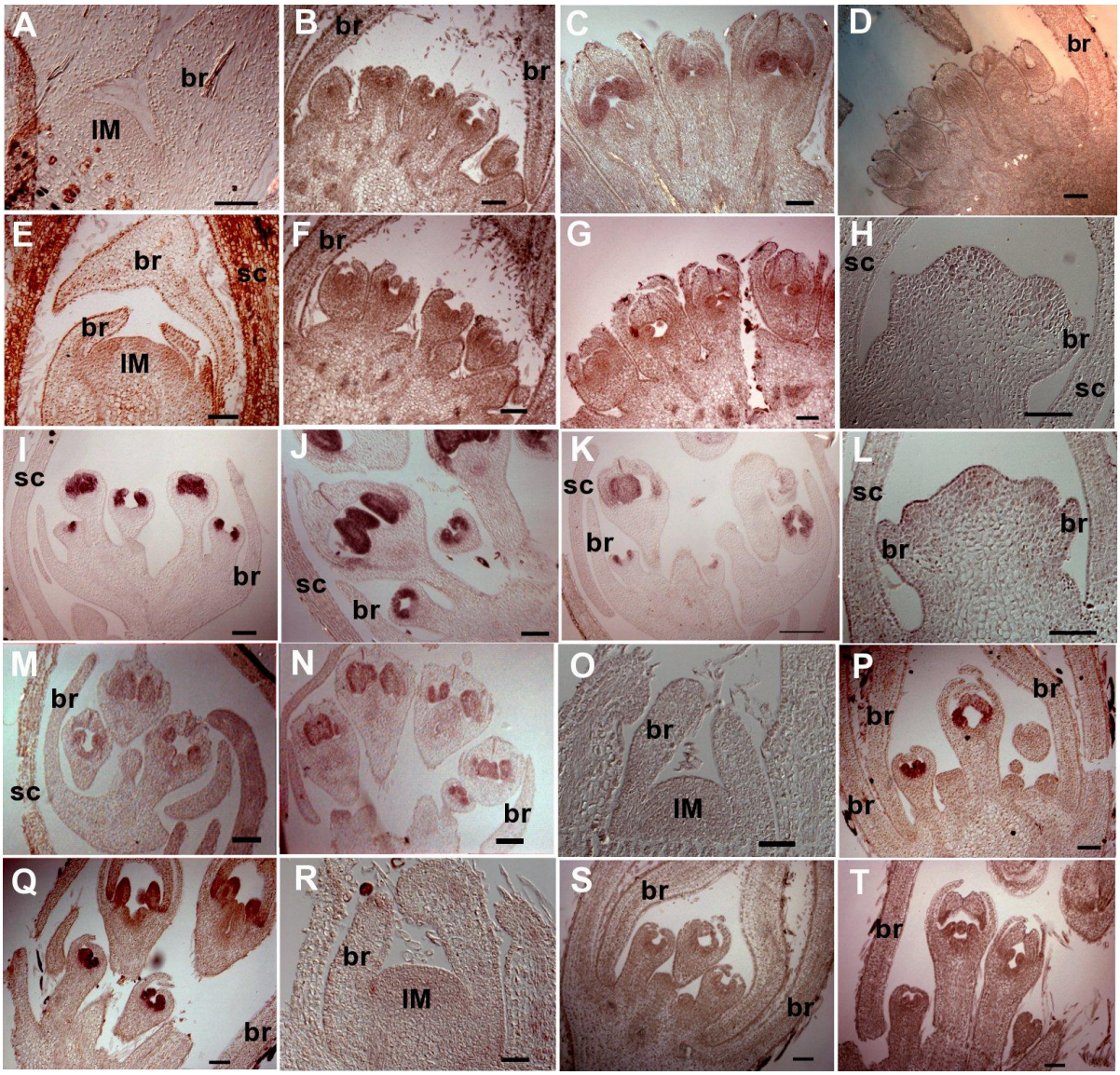
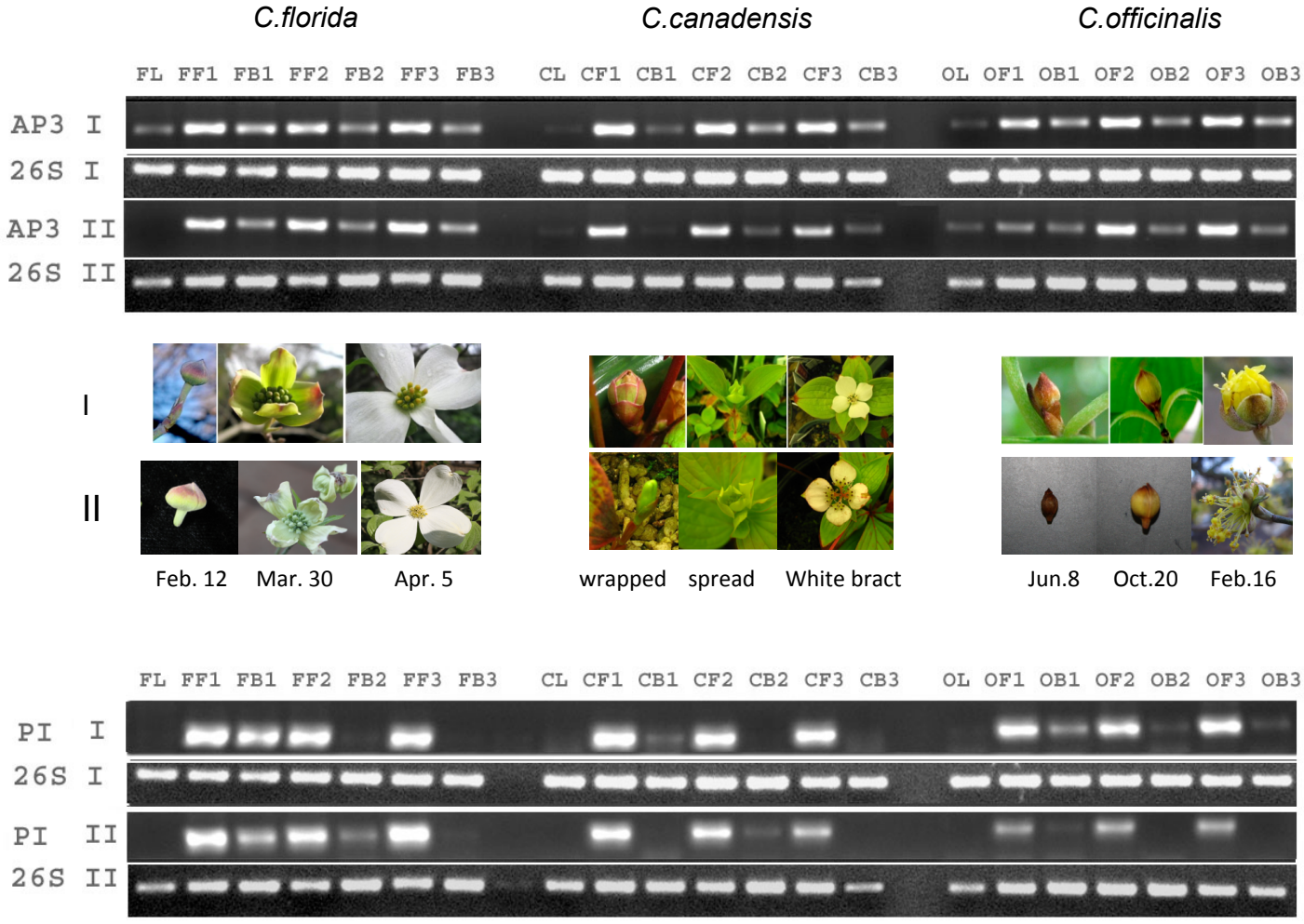


Fig. 14 Expression patterns of *CorPI* and *CorAP3* in *Cornus* bract and flower early development using in situ RNA hybridization. A-C: Expression of *CorPI*-B in *C.florida* developing bracts and flowers; D: Expression of *CorPI*-A in *C.florida* developing flowers; E-G: Expression of *CorAP3* in *C.florida* developing flowers; H-J: Expression of *CorPI*-A in *C.canadensis* developing flowers; K: Expression of *CorPI*-B in *C.canadensis*; L-N: Expression of *CorAP3* in *C.canadensis* developing flowers; O-Q: Expression of *CorPI*-A in *C.officinallis* developing flowers; R-T: Expression of *CorAP3* in *C.officinallis* developing flowers. br: bract; IM: inflorescence meristem; sc: scale; Scale bars: 100 um.

Fig. 15. Expression of *CorPI* and *CorAP3* in *Cornus* bract and flower late development using RT-PCR. I, II in gel panels indicate two experimental repeats, and I and II in plant image panel indicate two biological repeats. L: leaf; B1-B3: three bract late developmental stages representing unopened bracts, expanding bracts and whitened bracts, respectively; F1-F3: three flower late developmental stages associated with the three bract developmental stages. 26S rDNA is used for internal control.


C.florida
C.canadensis
C.officinalis
FL FF1 FB1 FF2 FB2 FF3 FB3
CL CF1 CB1 CF2 CB2 CF3 CB3
OL OF1 OB1 OF2 OB2 OF3 OB3
AP3 I
26S I
AP3 II
26S II
I
II
Feb. 12
Mar. 30
Apr. 5
wrapped
spread
White bract
Jun.8
Oct.20
Feb.16
FL FF1 FB1 FF2 FB2 FF3 FB3
CL CF1 CB1 CF2 CB2 CF3 CB3
OL OF1 OB1 OF2 OB2 OF3 OB3
PI I
26S I
PI II
26S II

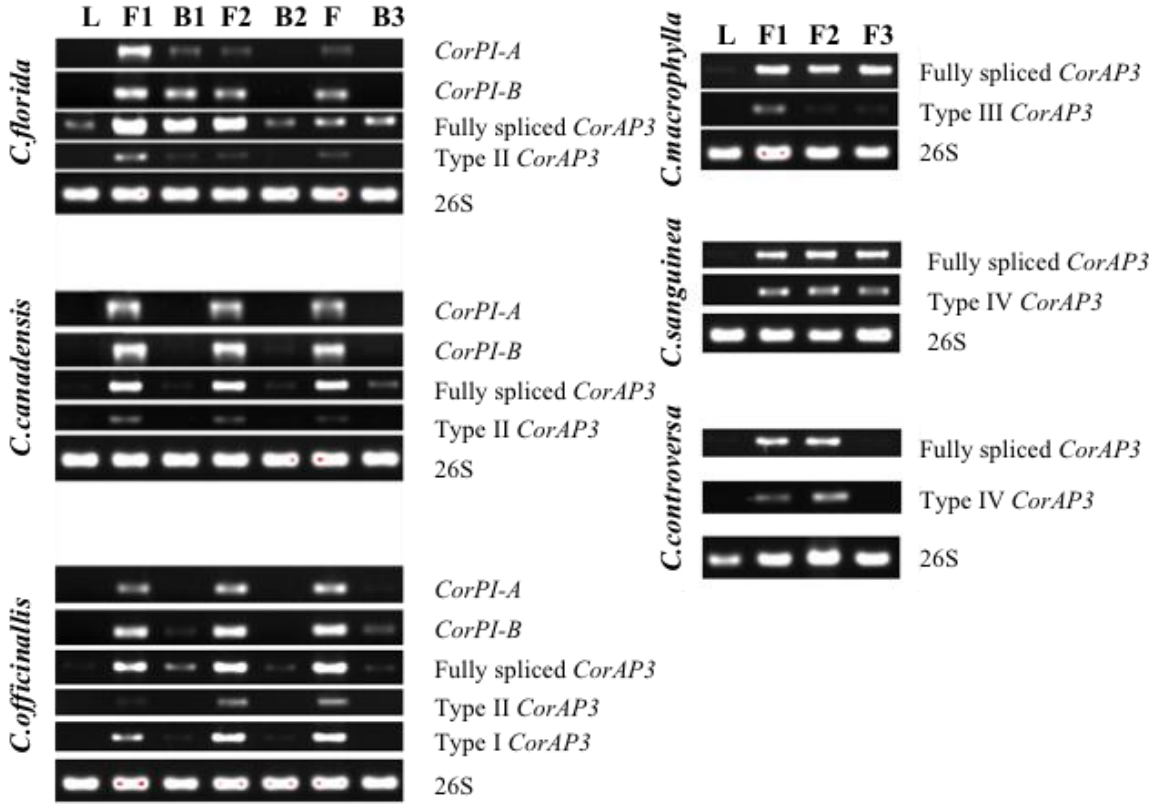
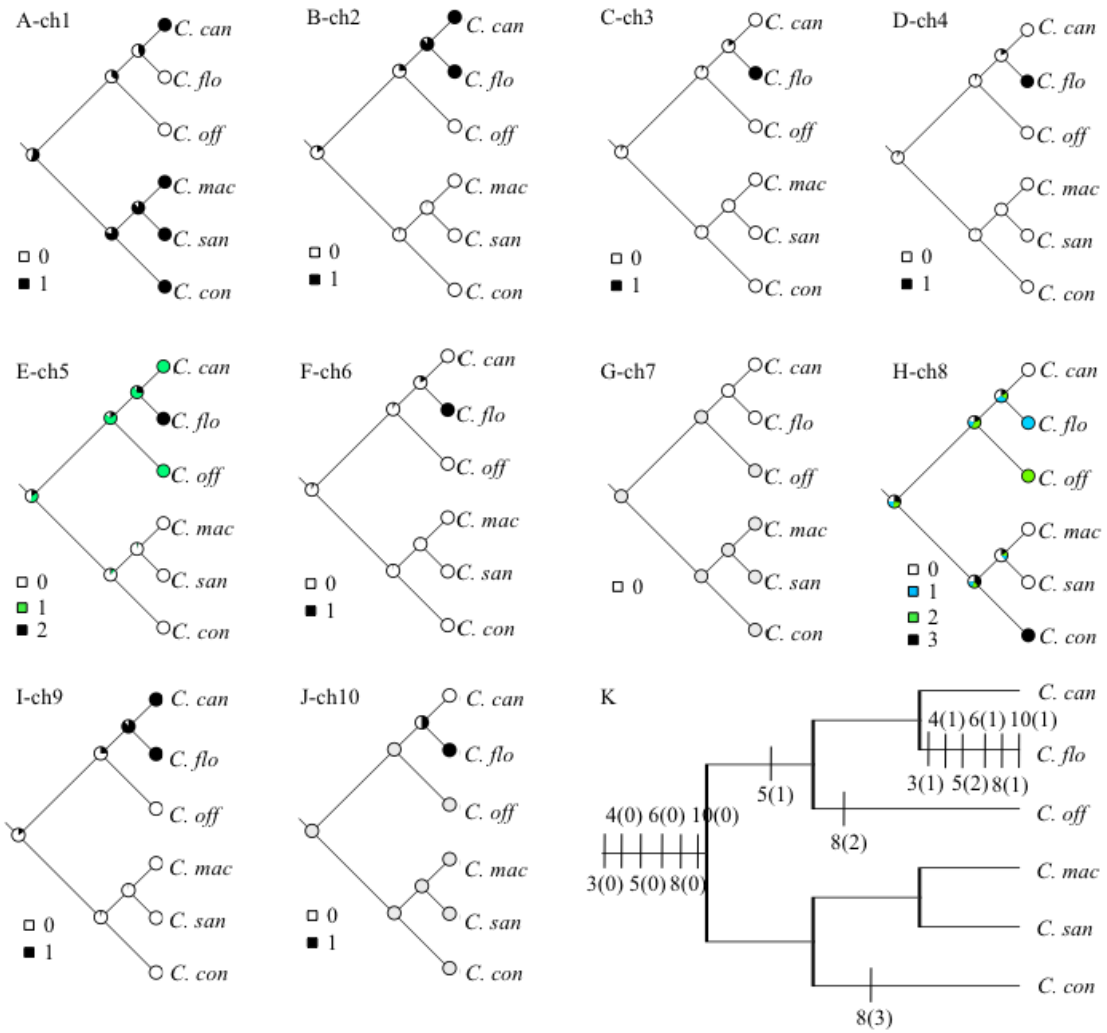


Fig. 16 Expression of *CorPI*-A, *CorPI*-B, and *CorAP3* in *Cornus* bract and flower late development using RT-PCR. L: Leaf, B1-B1 corresponds to three stages of Bracts in Fig. 13. F1-F3 corresponds to three stages of flower/inflorescences in Fig. 13. Types I to type IV *CorAP3* refer to alternatively spliced *CorAP3* transcripts in Fig. 9. 26S rDNA is used for internal control.

Fig. 17. Evolutionary changes in B-class genes and bract morphology associated with *Cornus* cladogenesis, inferred from ML analyses in Mesquite 2.74. The tree topology used in the analysis was simplified from the phylogeny of *Cornus* derived from DNA sequences of multiple genes (Xiang and Thomas 2008; Xiang et al. 2011). A-J: evolutionary trends for ten characters (ch1 to ch10); K: summary of evolutionary changes for the co-related characters in *C.florida*, ticks on branches represent the evolutionary events. Numbers by the ticks indicate the characters and numbers within the parenthesis are states of characters. Information on characters and character states is referred to Table 11.


A-ch1
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
B-ch2
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
C-ch3
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
D-ch4
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
E-ch5
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
2
F-ch6
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
G-ch7
C. can
C. flo
C. off
C. mac
C. san
C. con
0
H-ch8
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
2
3
I-ch9
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
J-ch10
C. can
C. flo
C. off
C. mac
C. san
C. con
0
1
K
C. can
4(1) 6(1) 10(1)
C. flo
3(1) 5(2) 8(1)
C. off
5(1)
8(2)
4(0) 6(0) 10(0)
3(0) 5(0) 8(0)
C. mac
C. san
C. con
8(3)

ST1. List of MADS-box class-B gene homologues for phylogenetic analysis.

Order	Family	Species	PI homolog
Cornales	Cornaceae	*Cornus canadensis*	CorcanPI4 (EU439931)
	Cornaceae	*Cornus florida*	CorfloPI_A (EU439925), CorfloPI_B (EU439926)
	Cornaceae	*Cornus officinallis*	CoroffPI918_85A (EU447705), CoroffPI918_85B (EU447706)
	Cornaceae	*Cornus mas*	CormasPI_A (EU439934), CormasPI_B (EU439935)
	Cornaceae	*Cornus macrophylla*	CormacPI (EU439933)
	Cornaceae	*Cornus sanguinea*	CorsanPIA1 (EU439928), CorsanPIA2 (EU439927)
	Cornaceae	*Cornus controversa*	CorconPI02_94 (EU447709)
	Alangiaceae	*Alangium platinifolium*	AlaplaGLO (GQ141107)
	Nyssaceae	*Nyssa sylvatica*	NyssylGLO1 (GQ141111), NyssylGLO2 (GQ141109)
	Nyssaceae	*Davidia involucrata*	DavinvGLO1(GQ141110), DavinvGLO2(GQ141108)
	Hydrangeaceae	*Hydrangea macrophylla*	HydmacPI (AF230711)
Ericales	Symplocaceae	*Symplocos chinensis*	SymchiGLO1(GQ141127),SymchiGLO2(GQ141149)
	Diapensiaceae	*Galax urceolata*	GalurcGLO1 (GQ141151), GalurcGLO2 (GQ141119)
	Actinidiaceae	*Sauravia zahlbruckneri*	SauzahGLO1(GQ141122), SauzahGLO2(GQ141141)
	Actinidiaceae	*Sauravia zahlbruckneri*	SauzahGLO3 (GQ141142)
	Actinidiaceae	*Actinidia chinensis*	ActchiGLO1 (GQ141123), ActchiGLO2 (GQ141124)
	Clethraceae	*Clethra tomentosa*	CletomGLO (GQ141143)
	Theaceae	*Stewartia pseudocamellia*	StepseGLO1 (GQ141140), StepseGLO3 (GQ141139)
	Theaceae	*Camellia japonica*	CamjapGLO1(GQ141126), CamjapGLO2 (GQ141138)

ST1. Continued.

Ericales	Ericaceae	*Erica hiemalis*	ErihieGLO1 (GQ141147), ErihieGLO2 (GQ141148)
	Ebenaceae	*Diospyros digyna*	DiodigGLO (GQ141136)
	Myrsinaceae	*Hymenandra wallichii*	HymwalGLO (GQ141135)
	Lecythidaceae	*Napoleona vogelii*	NapvogGLO (GQ141117)
	Roridulaceae	*Roridula gorgonias*	RorgorGLO (GQ141146)
Gentianales	Rubiaceae	*Coffea arabica*	CofaraGLO (GU332285)
	Gentianaceae	*Eustoma grandiflorum*	EusgraGLO1 (EF569228)
Apiales	Araliaceae	*Hedera helix*	HedhelGLO (HQ005416)
Asterales	Asteraceae	*Gerbera hybrida*	GerhybGLO1 (AJ009726)
	Asteraceae	*Helianthus annuus*	HelannPI (AY173069)
Solanales	Solanaceae	*Petunia hybrida*	FBP1 (M91190)
	Solanaceae	*Nicotiana tabacum*	NictabGLO (X67959)
Lamiales	Verbenaceae	*Verbena hybrida*	BerhybGLO (AB234909)
	Plantaginaceae	*Antirrhinum majus*	GLO (AB516403)
	Plantaginaceae	*Misopates orontium*	MisoroPI (AM162211)
Brassicales	Brassicaceae	*Arabidopsis thaliana*	PI (D30807)
	Brassicaceae	*Capsella bursa-pastoris*	CapburPI_A (EU551762)
	Brassicaceae	*Lepidium satvum*	LepsatPI (FJ411022)
Malpighiales	Salicaceae	*Populus deltoides*	PopdelPI (EU029172)
Cucurbitales	Cucurbitaceae	*Cucumis sativus*	CucsatSUM26 (AF043255)
Ranunculales	Ranunculaceae	*Aquilegia vulgaris*	AquvulPI (EF489475)
Ranunculales	Papaveraceae	*Papaver somniferum*	PapsomPI1 (EF071994), PapsomPI2 (EF071995)
Proteales	Nelumbonaceae	*Nelumbo nucifera*	NelnucPI (GU048646)

ST1. Continued.

Species	AP3 homolog	TM6 homolog
Cornus canadensis	To be submitted to GenBank	
Cornus florida	To be submitted to GenBank	
Cornus officinallis	To be submitted to GenBank	
Cornus mas	To be submitted to GenBank	
Cornus macrophylla	To be submitted to GenBank	
Cornus sanguinea	To be submitted to GenBank	
Cornus controversa	To be submitted to GenBank	
Alangium platinifolium	AlaplaDEF (GQ141169)	AlapalTM6 (GQ141152)
Nyssa sylvatica	NyssylDEF (GQ141170)	NyssylTM6 (GQ141155)
Davidia involucrata		DavinvTM6 (GQ141153)
Hydrangea macrophylla	HydmacAP3 (AB454439)	HydmacTM6 (AB454440)
Galax urceolata	GalurcDEF (GQ141187)	GalurcTM6 (GQ141164)
Sauravia zahlbruckneri		SauzahTM6 (GQ141167)
Sauravia zahlbruckneri		SauzahTM6 (GQ141167)
Clethra tomentosa	CletomDEF (GQ141180)	CletomTM6 (GQ141166)
Stewartia pseudocamellia	StepseDEF (GQ141182)	
Camellia japonica	CamjapDEF (HM773024)	CamjapTM6 (HM748646)
Erica hiemalis	ErihieDEF (GQ141186)	
Diospyros digyna	DiodigDEF (GQ141173)	
Hymenandra wallichii	HymwalDEF (GQ141176)	HymwalTM6 (GQ141159)
Napoleona vogelii	NapvogDEF (GQ141171)	
Coffea arabica	CofaraDEF (GU332287)	
Eustoma grandiflorum	EusgraDEF1 (EF569227), EusgraDEF2 (FJ469778)	
Gerbera hybrida	GerhybDEF2 (AJ009725), GerhybDEF3 (FJ817421)	GerhybTM6 (AJ009724)
Helianthus annuus	HelannAP3 (AY185363), HelannHAM2 (EF612597)	
Helianthus annuus	HelannHAM63 (EF612598)	
Petunia hybrida		TM6 (DQ539417)
Nicotiana tabacum	NictabDEF (X96428)	
Antirrhinum majus	DEF (X62810)	
Arabidopsis thaliana	AP3 (M86357)	

ST1. Continued.

Capsella bursa-pastoris	CapburAP3a (EU551764), CapburAP3b (EU551765)	
Lepidium satvum	LepsatAP3 (FJ411019)	
Cucumis sativus		CucsatMADS1 (AY944060)
Aquilegia vulgaris	AquvulAP3_1 (EF489478), AquvulAP3_2 (EF489477)	
Aquilegia vulgaris	AquvulAP3_3 (EF489476)	
Papaver somniferum	PapsomAP3_1 (EF071993), PapsomAP3_2(EF071992)	

ST2. Results of surface accessibility prediction for CorPI using NetSurfP-1.1

	CormacPI6		CormacPI5		CormacPI1		CorsanPI7		CorsanPI2		CorconPI1		CorconPI7	
Position	aa	SA	aa	SA	aa	SA	aa	SA	aa	SA	aa	SA	aa	SA
171	M	E	M	E	M	E	M	E	M	E	M	B	M	E
172	D	E	D	E	D	E	D	E	D	E	D	E	D	E
173	G	E	G	E	G	E	G	E	G	E	G	E	G	E
174	N	B	N	B	N	B	N	B	N	B	N	B	N	E
175	V	B	V	B	V	B	V	B	V	B	V	B	V	B
176	K	E	R	E	R	E	R	E	R	E	R	E	R	E
177	E	E	E	E	E	E	E	E	E	E	E	E	E	E
178	M	B	M	B	M	B	M	B	M	B	M	B	M	B
179	E	E	E	E	E	E	E	E	E	E	E	E	E	E
180	N	E	N	E	N	E	N	E	N	E	N	E	N	B
181	V	E	G	B	G	B	G	E	G	E	G	E	G	B
182	Y	B	Y	B	Y	B	Y	B	Y	B	Y	B	Y	B
183	R	B	Q	E	Q	E	Q	E	Q	B	Q	B	Q	E
184	Q	E	Q	B	Q	B	Q	B	Q	B	Q	E	Q	E
185	R	B	R	B	R	B	R	E	R	B	R	E	R	E
186	V	B	V	E	V	E	V	B	V	B	V	B	V	B
187	R	E	R	B	R	B	R	E	R	E	R	E	R	E
188	D	E	D	B	D	B	D	E	D	E	D	B	D	E
189	F	B	S	B	S	B	F	B	F	B	F	B	F	B
190	Q	B	Q	B	Q	B	Q	E	Q	B	Q	B	Q	B
191	P	B	P	B	P	B	P	B	P	B	P	B	P	B
192	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
193	M	B	M	B	M	B	M	B	M	B	M	B	M	B
194	P	B	P	B	P	B	P	B	P	B	P	B	P	B
195	F	B	F	B	F	B	F	B	F	B	F	B	F	B
196	A	B	A	B	A	B	A	B	A	B	A	B	A	B
197	F	B	F	B	F	B	F	B	F	B	F	B	F	B
198	R	B	R	B	R	B	R	B	R	B	R	B	R	B
199	V	B	V	B	V	B	V	B	V	B	V	B	V	B
200	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
201	P	B	P	B	P	B	P	B	P	B	P	B	P	B
202	I	B	I	B	I	B	I	B	I	B	I	B	I	B
203	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
204	P	B	P	B	P	B	P	B	P	B	P	B	P	B
205	N	B	N	B	N	B	N	B	N	B	N	B	N	B

ST2. Continued.

206	L	B	L	B	L	B	L	B	L	B	L	B	L	B
207	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
208	D	E	D	E	D	E	D	E	D	E	D	E	D	E
209	R	E	R	E	R	E	R	E	R	E	R	E	R	E
210	M	E	M	E	M	E	M	E	M	E	M	E	M	E

	CorcanPI10		CorcanPI1		CoroffoPI2		CoroffPI9		CorfloPI4		CorfloPI12	
Position	aa	SA	aa	SA	aa	SA	aa	SA	aa	SA	aa	SA
171	M	E	M	B	M	E	M	E	M	E	M	E
172	E	E	S	E	E	E	E	E	E	E	D	E
173	S	E	G	E	G	E	G	E	G	E	D	E
174	N	B	K	B	N	E	N	B	N	B	N	E
175	V	B	V	B	V	B	V	B	L	B	V	B
176	R	E	R	E	R	E	R	E	R	E	R	E
177	E	B	N	E	E	E	E	E	E	E	A	B
178	M	B	V	B	M	B	M	B	M	B	E	B
179	E	E	E	E	E	E	E	E	E	E	M	B
180	N	E	K	E	N	E	N	E	D	E	E	B
181	E	E	E	E	G	E	G	E	G	E	K	E
182	F	B	Y	B	Y	B	Y	B	F	B	T	B
183	H	B	H	E	H	E	H	B	H	B	Y	B
184	Q	B			Q	E	Q	E	Q	B	Q	B
185	R	E	H	E	R	E	R	E	R	B		
186	V	B	Q	E	V	E	V	E	L	B	Q	E
187	R	E	R	E	R	E	R	E	R	B	R	B
188	E	E	E	E	D	E	D	B	D	E	E	B
189	F	B	F	B	F	B	F	B	F	B	F	B
190	Q	E	Q	E	Q	E	Q	E	Q	B	Q	E
191	P	B	P	B	P	B	P	B	P	B	P	B
192	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
193	I	B	M	B	M	B	M	B	M	B	M	B
194	P	B	P	B	P	B	P	B	P	B	S	B
195	F	B	F	B	F	B	F	B	F	B	F	B
196	S	B	P	B	S	B	S	B	S	B	P	B
197	F	B	F	B	F	B	F	B	F	B	F	B
198	R	B	R	B	R	B	R	B	R	B	R	B
199	V	B	V	B	V	B	V	B	V	B	V	B

ST2. Continued.

200	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
201	P	B	P	E	P	B	P	B	P	B	P	B
202	I	B	S	B	I	B	I	B	I	B	S	B
203	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
204	P	B	P	B	P	B	P	B	P	B	P	B
205	N	B	N	B	N	B	N	B	N	B	N	B
206	L	B	L	B	L	B	L	B	L	B	L	B
207	Q	B	Q	B	Q	B	Q	B	Q	B	Q	E
208	D	E	D	E	D	E	D	E	D	E	D	E
209	R	E	R	E	R	E	R	E	R	E	R	E
210	M	E	M	E	M	E	M	E	M	E		

Only partial C domain (sites from 171-210) is shown; SA: surface accessibility; E: exposed; B: buried; underline indicates sites under positive selection and interacting with CorAP3 proteins; Yellow highlights sites that are estimated to be different in surface accessibility between CorPI-A and CorPI-B in *C.florida*.

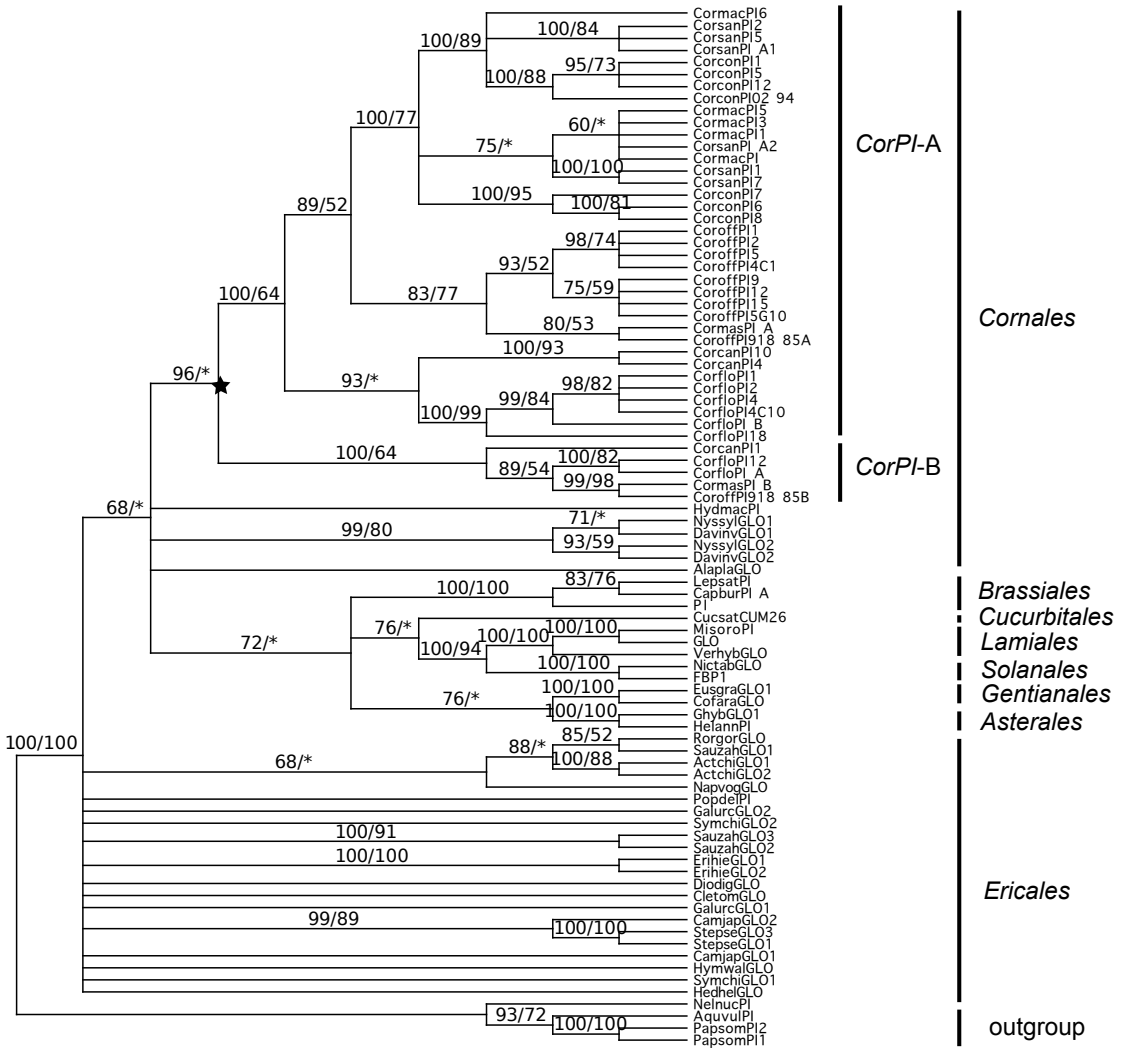
ST3. Results of surface accessibility prediction for CorAP3 using NetSurfP-1.1

	CorcanAP3		CorfloAP3		CoroffAP3		CorconAP3		CormacAP3		CorsanAP3	
position	aa	SA	aa	SA	aa	SA	aa	SA	aa	SA	aa	SA
165	L	E	L	E	L	E	L	E	L	E	L	E
166	H	E	H	E	N	E	H	E	H	E	H	E
167	E	E	E	E	E	E	D	E	D	E	N	E
168	F	B	F	B	F	B	F	B	F	B	F	B
169	E	E	E	E	E	E	E	E	E	E	E	E
170	/	/	/	/	/	/	I	E	I	E	I	E
171	I	E	T	E	I	E	T	E	I	E	I	E
172	R	E	K	E	K	E	R	E	R	E	R	E
173	E	E	G	E	E	E	E	E	E	E	E	E
174	E	E	E	E	E	E	E	E	E	E	E	E
175	D	E	D	E	D	E	D	B	D	E	D	E
176	P	E	P	E	P	E	P	E	P	E	P	E
177	H	E	H	E	H	E	H	E	H	E	H	E
178	Y	B	Y	B	Y	B	Y	B	Y	B	Y	B
179	G	B	G	B	G	B	G	B	G	B	G	B
180	L	B	L	B	L	B	L	B	L	B	L	B
181	V	B	V	B	V	B	V	B	V	B	V	B
182	D	E	D	E	D	E	D	E	D	E	D	B
183	N	E	N	E	N	E	N	E	N	E	N	E
184	G	E	G	E	G	E	G	E	G	E	G	E
185	G	E	G	E	A	E	G	E	G	E	G	E
186	D	E	D	E	D	E	D	E	D	E	D	E
187	Y	B	Y	B	Y	B	Y	B	Y	B	Y	B
188	D	E	D	E	D	E	D	E	D	E	D	E
189	S	B	S	B	S	B	S	B	S	B	S	B
190	V	B	V	B	V	B	V	B	V	B	V	B
191	L	B	L	B	L	B	L	B	L	B	L	B
192	G	E	G	B	G	E	G	E	G	E	G	B
193	F	B	F	B	F	B	F	B	F	B	F	B
194	S	E	T	E	S	E	S	E	S	E	S	B
195	G	B	G	B	G	B	G	E	D	E	D	B
196	G	E	G	E	G	E	G	E	G	B	G	B
197	G	B	A	B	D	B	P	E	G	B	G	B
198	P	B	A	B	P	B	P	B	P	B	P	B
199	R	B	S	B	R	B	R	B	H	B	R	B

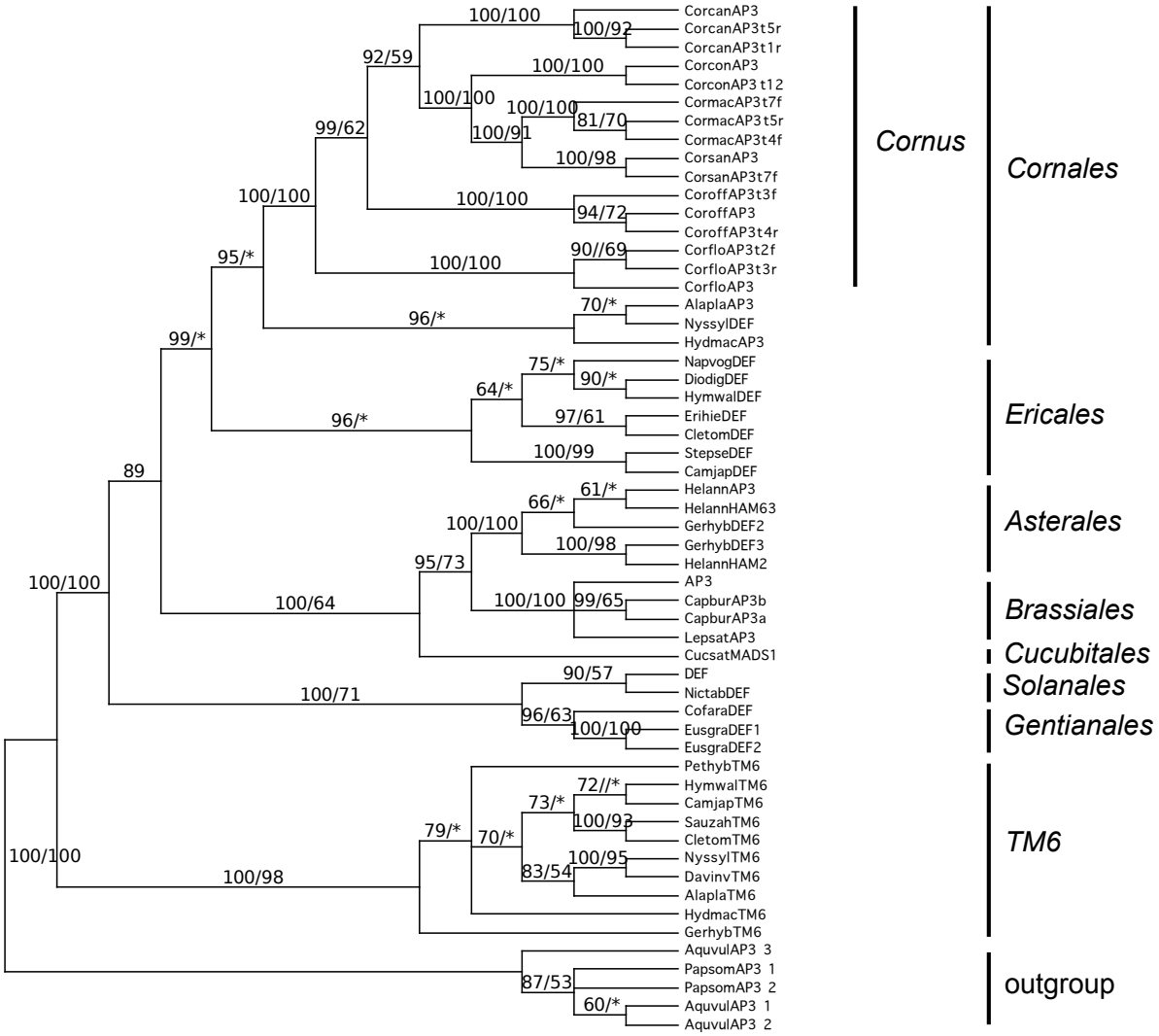
ST3. Continued.

200	L	B	I	B	I	B	I	B	I	B	I	B
201	L	B	L	B	I	B	L	B	L	B	L	B
202	A	B	A	B	A	B	A	B	A	B	A	B
203	L	B	L	B	L	B	L	B	L	B	L	B
204	H	B	R	E	R	B	R	E	R	E	R	E
205	L	B	L	B	L	B	L	B	L	B	L	B
206	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
207	Q	B	P	B	P	B	P	B	P	B	P	B
208	N	E	N	B	N	E	N	B	N	B	N	B
209	M	E	M	B	L	B	M	B	M	B	M	E
210	Q	B	Q	B	Q	B	Q	B	Q	B	Q	B
211	/	/	L	B	/	/	/	/	/	/	/	/
212	P	B	G	B	P	B	P	B	S	B	P	B
213	G	B	G	B	G	B	G	B	G	B	G	B
214	L	B	V	B	L	B	L	B	L	B	L	B
215	H	B	H	B	H	B	H	B	H	B	L	B
216	V	B	V	E	V	B	A	E	A	B	A	E
217	G	B	G	B	G	B	G	B	G	B	G	E
218	A	B	A	E	A	B	P	E	P	E	P	E
219	G	B	G	B	G	B	G	B	G	B	G	B
220	S	B	S	B	S	B	S	B	S	B	S	B
221	D	E	D	B	D	B	D	B	D	B	D	B
222	L	E	L	B	L	B	L	B	L	E	L	B
223	T	B	T	B	T	B	T	B	T	B	T	B
224	T	E	T	E	T	B	T	B	T	E	T	B
225	Y	B	F	B	Y	B	Y	B	Y	B	Y	B
226	A	E	A	E	A	E	A	E	A	E	A	E
227	L	B	L	B	L	E	L	E	L	E	L	E
228	L	E	L	B	L	B	L	B	L	B	L	B
229			E	E	E	E	E	E	E	E	E	E

Only partial C domain (sites from 169 to 229) is shown; SA: surface accessibility; E: exposed; B buried; underline indicates the site predicted to be under positive selection and to interact with CorPI proteins; Yellow highlights sites of CorAP3 that are estimated to be different in surface accessibility between *C.florida* and other five dogwood species.



SF1. Genealogy of *PI* homologs inferred from Bayesian and Maximum Likelihood (ML) analysis.
* indicates less than 50 bootstrap supports. Star indicates ancient duplication of *CorPI.*



SF2. Genealogy of *AP3* homologs inferred from Bayesian and Maximum Likelihood (ML) analysis.
* indicates less than 50 bootstrap supports.