The Pennsylvania State University
The Graduate School
Department of Biology

INVESTIGATING EVOLUTION OF PLANT DEVELOPMENT
IN BASAL ANGIOSPERMS

A Dissertation in
Biology
by
Barbara Joanne Bliss

© 2008 Barbara Joanne Bliss

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2008
The dissertation of Barbara Bliss was reviewed and approved* by the following:

Hong Ma  
Professor of Biology  
Dissertation Co-Advisor  
Co-chair of Committee

Claude dePamphilis  
Professor of Biology  
Dissertation Co-Advisor  
Co-chair of Committee

Paula McSteen  
Assistant Professor

Siela Maximova  
Research Associate

Doug Cavener  
Professor of Biology  
Head of the Department of Biology

*Signatures are on file in the Graduate School
ABSTRACT

Understanding the evolution of modern plants requires integrating findings from several disciplines, including plant physiology and development, molecular genetics, and genomics. Observations from model and non-model plants are brought together in a phylogenetic framework to derive hypotheses about how plant development evolved to generate the abundant diversity we see today. Testing those hypotheses requires a plant model system with the appropriate phylogenetic perspective: that of a basal lineage. The greatest diversity of plants today is among the angiosperms (flowering plants), a lineage which arose only about 160 million years ago. The most successful of these are the monocot and core eudicot flowering plant lineages, from which current plant model experimental systems are derived. For questions about the evolution of angiosperm development, a plant model from among the basal lineages is required.

In addition to phylogenetic perspective, model systems possess features and degrees of availability, representation, and utility not found in other members of the taxa to which they belong. For all organisms, culturing requirements are central determinants of utility, but for studying the evolution of plant development, amenability to studies employing methods of genomics, genetics, molecular and developmental biology are also required. This dissertation describes the search for and selection of a proposed basal angiosperm experimental model, Aristolochia fimbriata, along with the development of initial technologies required for testing hypotheses about the evolution of plant
development. Culturing, hand pollination, genetic transformation, and *in vitro* micropropagation and regeneration methods are described herein.

Genes involved in flower form and architecture have been particularly important in the evolution of angiosperm diversity. The TCP gene family, so named for its founding members (*TEOSINTE BRANCHED1, CYCLOIDEA, PROLIFERATING CELL FACTOR*) has been shown to play important roles in evolution of form in both monocots and eudicots. Prior functional and phylogenetic analyses of this gene family revealed clades of TCP genes with two different kinds of gene function. Since then, additional sequence data from basal lineages and new studies providing insight into TCP gene function have become available. Together, these warrant an updated phylogeny and review of this important gene family. Preliminary phylogenetic analyses of the TCP gene family is described as a foundation for conducting future expression and functional analyses. *A. fimbriata* has floral and vegetative features that will facilitate evaluating the role of TCP genes in evolution of angiosperm form, and advance the use of this species as a basal angiosperm model system.
TABLE OF CONTENTS

LIST OF FIGURES ............................................................................................................ ix
LIST OF TABLES ............................................................................................................. xii
LIST OF ABBREVIATIONS ................................................................................................. xiv
ACKNOWLEDGEMENTS .................................................................................................... xvii
Chapter 1  Introduction ...................................................................................................... 1
  Innovations in land plants .............................................................................................. 2
    Bryophytes and tracheophytes: Controlling water and gravity .................................. 2
    Spermatophytes: Efficient reproduction ..................................................................... 6
    Angiosperms: “Flowering” plants ............................................................................. 8
  Angiosperm phylogeny .................................................................................................. 11
    Phylogenetic analysis ............................................................................................... 11
    Selection of sequence data ....................................................................................... 12
  Basal angiosperms ........................................................................................................ 14
    Tracheary elements ................................................................................................. 16
    Double fertilization ................................................................................................. 17
    Flower ....................................................................................................................... 18
  Floral developmental genetics .................................................................................... 23
    Floral organ identity model ..................................................................................... 23
    Genes involved in floral organ identity .................................................................. 25
    Broader applications ............................................................................................... 29
  Basal angiosperm model ............................................................................................ 30
    Transformation systems ......................................................................................... 31
    Regeneration systems ............................................................................................ 33
  Overview of dissertation .............................................................................................. 34
Chapter 2  Aristolochia fimbriata: A proposed experimental model for basal
  angiosperms ................................................................................................................. 37
  Abstract ....................................................................................................................... 41
  Introduction ................................................................................................................... 42
  Results ........................................................................................................................... 47
    Evaluation of potential models in basal angiosperm orders and families ............ 47
    Aristolochiaceae candidates considered ............................................................... 50
    Physical and life cycle features of Aristolochia fimbriata ................................... 52
    A phylogenetic perspective of genome sizes ...................................................... 55
    Methods for genetics ............................................................................................... 57
  Discussion ...................................................................................................................... 60
Aristolochia fimbriata has many characteristics of a valuable experimental model ................................................................. 60
Aristolochia fimbriata is well positioned for studies of the evolution of development ............................................................... 61
Aristolochia contains highly developed biochemical pathways offering insight into evolution of biochemical synthesis and coevolution with insects .................................................................................................................. 64
Aristolochia can provide insight into development of woodiness ........ 66
Aristolochia can reveal features of the ancestor common to monocots and eudicots ......................................................................................................................... 67
Growing genomic resources in Aristolochiaceae support further development of model system .................................................. 68

Conclusion ........................................................................................................................................................................... 71
Materials & Methods ............................................................................................................................................................. 72
Cultivation .................................................................................................................................................................................. 72
  Genome sizing .................................................................................................................................................................... 73
Phylogenetic analysis ............................................................................................................................................................. 73
Pollination experiments ......................................................................................................................................................... 74
Genetic transformation ............................................................................................................................................................. 75
Genomic PCR analysis ............................................................................................................................................................ 76
Genetic transformation ......................................................................................................................................................... 77
Supplemental Material ........................................................................................................................................................... 77
Acknowledgements ............................................................................................................................................................... 78

Cultivation Supplement ........................................................................................................................................................... 82
  Background ........................................................................................................................................................................ 82
  Perianth maturation ............................................................................................................................................................. 83
  Hand pollination methods .................................................................................................................................................. 84
  Pollination is prevented with pollination bags .................................................................................................................... 86
  Day two autogamous pollinations are most successful ........................................................................................................... 87
  Self-compatibility in Aristolochia elegans ........................................................................................................................ 88
  Seed germination methods ................................................................................................................................................ 88
  Effect of seed age on germination ..................................................................................................................................... 89

Chapter 3  In vitro propagation, shoot regeneration and rooting protocols for Aristolochia fimbriata, potential model plant for basal angiosperms .................................................. 95

Abstract ...................................................................................................................................................................................... 97
Introduction .................................................................................................................................................................................. 99
Materials and Methods ............................................................................................................................................................ 102
  Plant material ....................................................................................................................................................................... 102
  Micropropagation initiation and multiplication .................................................................................................................. 103
  In vitro rooting .................................................................................................................................................................... 106
  Shoot regeneration from whole leaf explants ........................................................................................................................ 107
  Shoot regeneration from petiole and stem sections ................................................................................................................. 108
Results .................................................................................................................................................................................... 110
Seeds storage experiments ................................................................. 191
Inducing flowering.............................................................................. 194
Hand pollination ................................................................................... 198
Recording pedigree .............................................................................. 205
  Example 1: ...................................................................................... 206
  Example 2: ...................................................................................... 207
Possible mutants .................................................................................. 208

Literature cited .................................................................................... 210
LIST OF FIGURES

Figure 1-1: Evolution of major innovations in embryophytes..............................................4
Figure 1-2: Angiosperm flowers............................................................................................10
Figure 1-3: A consensus phylogeny of angiosperms..............................................................16
Figure 1-4: Floral specializations...........................................................................................20

Figure 2-1: Angiosperm phylogeny based on Stevens (2007) (http://www.mobot.org/MOBOT/research/APweb/welcome.html). ..................44
Figure 2-2: Overall approach for selecting a basal angiosperm model system. .................49
Figure 2-3: Genome sizes in basal angiosperm families updated with Cui et al. (2006) and Bennett and Leitch (2005), shown on logarithmic scale. .................50
Figure 2-4: Diversity of flower and growth forms in Aristolochiaceae...............................53
Figure 2-5: Aristolochia fimbriata..........................................................................................54
Figure 2-6: Phylogenetic relationships among sampled Aristolochiaceae, with Tasmannia lanceolata (Cannelleales) as outgroup.................................56
Figure 2-7: Green fluorescent protein expression in Aristolochiaceae.................................60
Figure 2S-1: Aristolochia fimbriata flower. ...........................................................................84
Figure 2S-2: Hand pollination, A. fimbriata ..........................................................................85
Figure 2S-3: Aristolochia fimbriata genotype and perianth detail .......................................93
Figure 2S-4: Phylogram of Aristolochiaceae relationships, showing minimal variation in branch lengths within Aristolochiaceae. .......................94
Figure 3-1: Aristolochia fimbriata regeneration. ....................................................................105
Figure 3-2: Schematic representation of micropropagation and regeneration protocols of A. fimbriata .................................................................111
Figure 4-1: Supertribe 777, MUSCLE alignment of amino acids ........................................140
Figure 4-2: Supertribe 1148, MUSCLE alignment of amino acids .......................................141
Figure 4-3: Supertribe 2066, MUSCLE alignment of amino acids. ......................... 142

Figure 4-4: RaxML phylogenetic analysis of TCP domain-containing sequences from automated pipeline ................................................................. 143

Figure 4-5: Neighbor joining tree and expression pattern of TCP genes in rice and *Arabidopsis* from Yao et al. (2007). ....................................................... 145

Figure 4-6: Unrooted ML tree of 126 TCP proteins from land plants from (Navaud et al., 2007). ................................................................. 147

Figure 4-7: Classes 1, 2, 3 TCP sequences ............................................................................................................................ 148

Figure 4-8: Class 1 (PCF-like) TCP sequences.. ..................................................... 149

Figure 4-9: HMM logo PFAM03634 ........................................................................... 150

Figure 5-1: Evaluating Kanamycin (Kan) as selection agent........................................ 156

Figure 5-2: GFP expression in *A. fimbriata* roots..................................................... 158

Figure A-1: MultiPipMaker plots of organelle genome sequences aligned to chloroplast target genome sequences......................................................... 169

Figure A-2: Frequency of high identity regions in alignments of real and simulated mitochondrial to chloroplast genome sequences................................. 171

Figure A-3: Distribution of similarity scores for 100 bp windows from 17 pairwise alignments of mt genome sequence or simulated mt genome sequence to cp genome sequence ................................................................. 173

Figure A-4: Strict consensus trees from NJ analysis with PAUP* (500 bootstrap replicates) using aligned coding sequences from cp and mt genomes. ............... 176

Figure A-5: Successful amplification of *psbA* (A, C), *rbcL* (B, D), and *ycf1* (E, F), from the *Arabidopsis thaliana* (A, B, E) chloroplast and (C, D, F) mitochondrial genomes................................................................. 177

Figure A-6: Strict consensus trees from PAUP* (500 bootstrap replicates) analysis of aligned coding sequences from cp and mt genomes. ................................. 178

Figure B-1: Seeds planted in varying conditions on 9/28/2006................................. 183

Figure B-2: Comparing effects of light on germination of seeds planted on ½ MS media................................................................................................. 184

Figure B-3: Comparing seed germination on wet filter paper or medium................. 185
Figure B-4: Comparing effect of light on germination of seeds planted in wet toweling. ................................................................. 186

Figure B-5: Effects of light on germination of seeds planted in wet toweling. ........ 187

Figure B-6: Comparing effect of light on germination of seeds planted on soft media........................................................................................................................................................................ 188

Figure B-7: Toweling germination method. ......................................................................................................................... 189

Figure B-8: Short term effects of seed age on germination........................................ 192

Figure B-9: Longer-term effects of seed age on germination........................................ 193

Figure B-10: Inducing blooms................................................................................................................................................ 195

Figure B-11: A. fimbriata produces successive axillary blooms along the shoot...... 196

Figure B-12: A. fimbriata in bloom. .......................................................................................................................... 197

Figure B-13: Pollination bag........................................................................................................................................... 199

Figure B-14: Pollination bag attachment detail. ................................................................................................................. 200

Figure B-15: Day of anthesis. .......................................................................................................................... 201

Figure B-16: Gynostemium .......................................................................................................................... 202

Figure B-17: Self-pollination in a day two flower.................................................................................................................. 203

Figure B-18: Inferior ovary.......................................................................................................................... 204

Figure B-19: Abnormal seedlings.......................................................................................................................... 209
LIST OF TABLES

Table 2-1: Homologs of genes involved in development, cell wall biosynthesis, and response to biotic and abiotic stress revealed in preliminary sequencing of two cDNA libraries constructed from combined *A. fimbriata* tissues. ..........69

Table 2S-1: A summary of relevant basal angiosperm characteristics. ....................79

Table 2S-2: Cultivation features of 25 Aristolochiaceae taxa evaluated, with findings. .................................................................81

Table 2S-3: Evaluating self-compatibility in *A. fimbriata*. ................................86

Table 2S-4: Percent fruit set (sample size) in *A. elegans* resulting from hand pollinations of flowers on days 1-4 with autogamous, geitonogamous and xenogamous pollen from flowers on days 1-3. ........................................88

Table 2S-5: Genome sizes, vouchers, sources and accessions for sequence data used. ..........................................................................................................................90

Table 3-1: Media formulations for micropropagation and shoot organogenesis. ......105

Table 3-2: The effect of plant growth regulators on axillary shoot proliferation at 21 days after culture initiation. ..............................................................112

Table 3-3: Effect of indole-3-butyric acid (IBA) on rooting of shoots multiplied in vitro..............................................................113

Table 3-4: Direct organogenesis of *Aristolochia fimbriata* from whole leaf explants ........................................................................114

Table 3-5: Shoot regeneration of petiole and stem explants.............................116

Table 3-6: Shoot regeneration, rooting and acclimation of plants regenerated from petiole and stem explants from plants maintained in REN2 medium.........117

Table 4-1: *Arabidopsis thaliana* (At) and *Oryza sativa* ssp. *japonica* (Os) TCP genes from Yao et al. (2007) and Navaud et al. (2007) ........................................136

Table 4-2: Three supertribes of genomic sequences comprise Arabidopsis and rice genes of interest. .................................................................137

Table 4-3: Genomic sequences containing TCP domain (PFAM domain 03634). ....138

Table A-1: List of species and GenBank accessions for 25 organelle genome sequences. ........................................................................................................165
Table A-2: Primers to amplify chloroplast genes from *Arabidopsis* cp and mt genomes. ................................................................. 167

Table A-3: Locations of chloroplast-specific sequences in pairwise alignments of mt genome sequences with cp genome sequences. ........................................ 175

Table B-1: Pedigree record, example 1................................................................. 206

Table B-2: Pedigree record, example 2................................................................. 207

Table B-3: Pedigree and incidence of “tricots.” ..................................................... 208
LIST OF ABBREVIATIONS

6BA: 6-benzoamino-purine
AAGP: Ancestral Angiosperm Genome Project (NSF DEB 0638595)
AG: AGAMOUS
AGL6: AGAMOUS-Like 6
AP1: APETAL1
bp: base pairs
CaMV: cauliflower mosaic virus
cp: chloroplast
CUC: CUP-SHAPED COTYLEDON
cup: cupuliformis
CYC: CYCLOIDEA
DICH: DICHOTOMA
DIV: DIVARICATA
EGFP: enhanced green fluorescent protein
ERF: ethylene responsive factor
EST: expressed sequence tags
FGP: Floral Genome Project (NSF DBI-0115684)
GFP: green fluorescent protein
HGT: Horizontal gene transfer
IBA: indole-3-butyric acid
Kan: kanamycin
MADS: MCM, AGAMOUS, DEFICIENS, SRFC
MAX: MORE AXILLARY GROWTH
**MI:** micropropagation initiation medium

**MP:** micropropagation medium

**MS:** Murashige and Skoog (basal medium)

**mt:** mitochondrial

**mya:** million years ago

**NAA:** $\alpha$-naphthalene acetic acid

**nptII:** neomycin phosphotransferase

**orf:** open reading frame

**PCF:** proliferating cell factor

**PCNA:** proliferating cell nuclear antigen

**PCR:** polymerase chain reaction

**petG:** cytochrome b6/f complex subunit 5

**PGR:** plant growth regulator

**pip:** percent identity plot

**PGR:** plant growth regulator

**psaA:** photosystem I P700 apoprotein A1

**psbA:** photosystem II reaction center polypeptide D1

**psbD:** photosystem II reaction center protein D2

**QTL:** quantitative trait loci

**rbcL:** ribulose 1, 5–bisphosphate carboxylase/oxygenase, large subunit

**RAD:** RADIALIS

**REN1, REN2:** root elongation media one and two

**RI:** root initiation medium

**SI:** shoot induction medium
SR: shoot regeneration medium

TB1: TEOSINTE BRANCHED1

TCP: TB1, CYC, PCF

TDZ: thidiazuron

ycf1, ycf2: large orf common in angiosperm chloroplast genomes
ACKNOWLEDGEMENTS

I would like to thank my committee, for their individual support and collective service on my behalf. In the lab of Dr. Claude dePamphilis, I conceived of my project and brought it to bear, a feat I certainly could have done nowhere else. The hand of Dr. Hong Ma kept me guided firmly toward my goal; providing me with technical expertise and supporting my interest in classical and molecular genetics. Dr. Siela Maximova has been an unfailing inspiration, even when all else failed, and Dr. Paula McSteen has served to moderate and tell me plainly what the outcome should look like. My work at Penn State was supported in 2002-03 by a University Graduate Fellowship from the graduate school of Penn State University and a Braddock Fellowship from the Eberly College of Science. During 2003-08, the work was supported by NSF grants DBI0115684 (Floral Genome Project) and DEB 0120705 (Ancestral Angiosperm Genome Project) to Claude dePamphilis, Teaching Assistantships from the Department of Biology, and travel grants from the Floral Genome Project. I received another Braddock Fellowship from the Eberly College of Science in 2008.

I deeply appreciate my colleagues and mentors. My lab mates Jill R. Duarte, Yan Zhang, and most especially Joel McNeal shared their expertise. Jim Leebens-Mack, Kerr Wall, Paula Ralph, and Ali Barakat provided the eagle’s perspective, and Laura Zahn contributed the outsider’s viewpoint, encouragement, and friendship. My training has also been shaped by the excellent habits of Sharon Pishak and the focus and determination of Stefan Wanke.
I am grateful to my family and friends who recklessly urged me to embark on this endeavor, knowing I would not rest until I had done so. They have not ceased to remind me that despite circumstances encountered along the way, this was - and is - the right path for me. I especially want to acknowledge the sacrifices of my two children, who gave up so much of Mom just so that she could go to school to get a job that she liked. I hope they, too, will not hesitate to go to school to get a job that they like, and that if they are also parents when they do, that their children give them just as hard a time as they have given me. They have been the wind beneath my wings; which has been, indeed, the wind of a firestorm threatening me with certain immolation if I do not rise high, quickly enough.
Chapter 1
Introduction

Research into the evolution of plant development aims to understand the mechanisms which have given rise to the wide variety of plants we observe today (Raff, 2000). The most speciose extant plant lineage is that of the angiosperms, with over 250,000 species. In the angiosperms, the innovations of flower, fruit, and double fertilization supported such a rapid radiation of diversity that Darwin described it as an “abominable mystery” (1903). Although the origins of the angiosperms have seemed enigmatic, as the disciplines of developmental biology (from early embryology), genetics, molecular biology, molecular evolution, and modern genomics are brought to bear, Darwin’s “abominable mystery” is giving way to a rudimentary understanding of processes involved in the evolution of angiosperm development (Friedman et al., 2004).
Innovations in land plants

"Ontogeny (embryology or the development of the individual) is a concise and compressed recapitulation of phylogeny (the paleontological or genealogical series) conditioned by laws of heredity and adaptation." (Haeckel, 1909)

The evolution of development that led to the flowering plant lineage was directed and limited by development that had occurred previously, and had proved to be advantageous. Development in an individual flowering plant does not strictly recapitulate the evolutionary history of angiosperms, but major changes in form and organization can be identified in the lineage leading to land plants (embryophytes) since they last shared a common ancestor with charophycean algae (Endress, 2006; Qiu et al., 2007). Thus, it is necessary to study the predecessors of flowering plants in order to recognize and construct hypotheses about what features evolved to give rise to the flowering plants. What follows is a brief review of the major vegetative and reproductive innovations in land plants leading to the success of the angiosperm lineage.

Bryophytes and tracheophytes: Controlling water and gravity

Early adaptive changes required for successful colonization of land arose around 480 million years ago (mya) in the Ordovician period, in an aquatic common ancestor the embryophytes shared with Characeae (Karol et al., 2001; Judd et al., 2002; Simpson, 2006). These early adaptations included apical development and phragmoplastic cell division, necessary for producing land plant architecture. Cellulosic cell walls and specialized gametangia (antheridia and archegonia) with encased egg cells (cortication)
contributed to the ability to survive drying conditions of land, particularly for reproductive structures (Figure 1-1).

Land plants appeared in the lower Silurian (approx 430 mya) and shared further adaptations, including a cuticle and jacketed gametangia to further protect the gametes and embryo from desiccation (Figure 1-1) (Qiu, 2008). The basal nonvascular embryophytes are known as the bryophytes and include the mosses, hornworts, and liverworts (Figure 1-1) (Nickrent et al., 2000). The genome sequence of the moss Physcomitrella patens (Physcomitrella) (Qiu et al., 2007), reveals a loss of genes required for success in an aquatic environment, and acquisition of new genes required for long-term success on land (Qiu et al., 2006; Rensing et al., 2008). The Physcomitrella genome sequence contains evidence of a single large-scale genome duplication event, which would have supported the birth of new genes by providing the raw material for innovation (Ohno, 1970; Howarth and Donoghue, 2005; Rensing et al., 2008). Sequence similarity with the modern angiosperms Arabidopsis thaliana (Arabidopsis) or Oryza sativa (Oryza) (Rensing et al., 2008) reveals acquisition or expansion of genes functioning in desiccation and UV tolerance, in hormonal pathways coordinating development (auxins, abscisic acid, and cytokinin signaling pathways), in control of morphogenesis (phytohormones and light receptors), and in vesicle trafficking (adenosine triphosphate binding cassette proteins) in the Physcomitrella genome, indicating further adaptations to the atmosphere and gravity.

Vascular tissue (tracheids) arose, approximately ten million years after plants colonized land, potentiating an adaptive radiation of vascular plants (tracheophytes) including vascular non-seed plants (club mosses, whisk ferns, horsetails, ferns), and
Figure 1-1: Evolution of major innovations in embryophytes. Nonvascular, basal embryophytes (blue). Tracheophytes include vascular non-seed plants (orange) and seed plants. Gymnosperms (green) and angiosperms (pink) are monophyletic.
vascular seed plants (cycads, gnetophytes, ginkgos, conifers, angiosperms). These tracheophytes attained much greater size and independence from water in the environment than had been possible for nonvascular plants (Judd et al., 2002; Qiu et al., 2007). The gametophyte became reduced and specialized for reproduction, eventually entirely dependent on and retained by the sporophyte (endosporic development). The sporophytes became the dominant generation in the tracheophytes, superseding the dominant gametophyte generation of nonvascular plants (Figure 1-1) (Judd et al., 2002; Simpson, 2006). Expressed sequence tags (ESTs) from the basal tracheophyte and lycopod Selaginella moellendorffii (Selaginella) include genes for several functions required for vascularization and adaptation to land not found in Physcomitrella, including genes involved in lignification, cell division control, intracellular transport, responses to sulfur starvation, dehydration, and viral infection (Weng et al., 2005).

The basalmost lineage of tracheophytes, the lycophytes (club mosses), diverged from the lineage which later led to the euphyllophytes (whisk ferns, horsetails, ferns, and seed plants) around 400 mya (Qiu et al., 2007). The lycophytes lack lateral leaf traces and lateral root branches, and have microphylls instead of euphyls (Kenrick and Crane, 1997) mya (Figure 1-1). Later diverging tracheophytes, including the pteridophytes (ferns, fern allies) and spermatophytes (seed plants) share euphyls (true leaves) as well as a 30 kB inversion in the chloroplast genome (Raubeson and Jansen, 1992).
Spermatophytes: Efficient reproduction

The spermatophytes, in which the seed serves as the unit of dispersal, first appeared near the end of the Devonian period, about 360 mya (Judd et al., 2002; Simpson, 2006). The seed consists of an embryonic sporophyte enclosed with food stores in a protective coating (seed integument) derived from sterile, parental, sporophyte tissues. The seed allows the dormant sporophyte generation to disperse over long distances and survive harsh conditions of drought and cold. Sperm, which is delivered to eggs by water droplets in non-seed plants, is carried to eggs by pollination in spermatophytes. The spermatophyte megasporangium typically, and ancestrally retains just one megaspore after the abortion of three of the four haploid products of meiosis (Friedman and Williams, 2004). The female gametophyte contained therein is enveloped by sterile sporophytic tissue forming the integument, which has a single opening at one end (the micropyle).

Microspores, immature male gametophytes in the form of pollen grains, develop in microsporangia on the same plant (monoecy) or on another plant (dioecy). The seed plant microspore germinates, giving rise to a haustorial pollen tube which grows and extracts nutrition from the sporophytic tissue surrounding the megasporangium on its way to deliver sperm to the ovule (siphonogamy). In addition to the shared seed characteristic of spermatophytes, seed plants also produce wood via secondary growth of xylem from a bifacial cambium (Figure 1-1) (Judd et al., 2002).

Seed plants comprise the most diverse vascular plant lineage, with about 270,000 extant species. The major extant spermatophytes lineages include cycads (Cycadales),
ginkgos (Ginkgoales), conifers (Coniferales), gnetophytes (Gnetales), and flowering plants (angiosperms) (Figure 1-1) (Judd et al., 2002; Simpson, 2006). In the first four seed plant lineages, megasporangia are exposed except for the coverage provided by megasporophylls, hence they are collectively known as the “gymno-“ (naked) “-sperms” (seeds). In the gymnosperms, pollen grains are delivered (presumably originally by wind pollination) near the micropyle, where the microspore germinates and slowly grows toward the egg.

Relative to the angiosperms, the four gymnosperm lineages contain greater diversity than the angiosperms, but only a fraction of the species (Raven et al., 1999; Judd et al., 2002; Simpson, 2006). The cycad lineage is limited to about 130 species, which feature limited secondary xylem, dioecy, and compound leaves. They have lost the axillary branching that appeared earlier in land plants, and have derived coralloid roots which host nitrogen-fixing cyanobacteria. Cycads have derived a specialized pollination system employing irregular thermogenesis in the cone, and mutualistic relationships with its pollinators. Cycad seeds are usually colorful and fleshy, attracting animal dispersal agents. Although the basalmost cycad has several ovules borne on leaf-like megasporophylls, more derived cycads bear two ovules on peltate megasporophylls in strobili.

Ginkgos form a monospecific taxon, with secondary growth and fan-shaped leaves. Ginkgos are also dioecious, bear ovules in pairs on axillary stalks, and produce seeds with a fleshy, smelly outer integument and hard inner integument. Both cycads and ginkgos have giant swimming sperm, which are interpreted as an ancestral feature in seed plants (Raven et al., 1999; Judd et al., 2002; Simpson, 2006; Roemer et al., 2008).
Conifers, gnetophytes and angiosperms have nonmotile sperm, thus the pollen tube must grow farther to deliver them directly to the female gametophyte.

Conifers comprise approximately 600 living species with well-developed wood and, often, needle-like leaves adapted for drought. Conifers release pollen from strobili bearing microsporangia on microsporophylls (Raven et al., 1999; Judd et al., 2002; Simpson, 2006). Seed-producing strobili have megasporophylls bearing receptive ovules, in which one remaining haploid cell gives rise to the female gametophyte, which then produces one or more eggs. Polyembryony is common in conifers, due either to separate fertilization events or by subdivision of a single embryo in early development.

Gnetophytes contain three genera and about 80 species, including Ephedra, Gnetum, and Welwitschia mirabilis. Gnetophytes share unusual features including vessels with pore-like perforation plates, a form of double fertilization that does not produce endosperm, both compound microsporangiate and compound macrosporangiate strobili, unique pollen shape and markings, and opposite leaves (Friedman, 1992; Carmichael and Friedman, 1996; Raven et al., 1999; Judd et al., 2002; Simpson, 2006). These relatively few species in the four basal seed plant lineages contain advantageous innovations (Figure 1-1) which can be seen as precursors or initial debuts of features later optimized and exploited in the angiosperms.

**Angiosperms: “Flowering” plants**

Angiosperms arose in the seed plant lineage about 160 mya (Bell et al., 2005; Leebens-Mack et al., 2005). Although they are known as the “flowering plants,” the
specific innovation that changed the condition of gymnospermy to angiospermy was the innovation of the carpel. The carpel has a lower portion (ovary) enclosing the ovules and an upper portion (stigma), often elongated (style), which is receptive to pollen (Figure 1-2 A). Angiosperm pollen is produced in stamens, which are microsporophyll-like organs bearing two thecae (anther), each containing two sporangia (Figure 1-2 A, C). The angiosperm flower is comprised of the staminate and carpellate reproductive structures, together with sterile tissues, including those of the perianth (Figure 1-2 A), but the sterile tissues can be highly variable. Wind-pollinated flowers may lack a perianth entirely (Figure 1-2 B, C, D). Lineage-specific flower and inflorescence modifications vary widely. In particular, the grass spikelet (Poaceae) (Figure 1-2 C), and spathe/spadix structures of Araceae (Figure 1-2 E), bear little resemblance to the typical angiosperm flower.

Pollen grains (microspores) delivered to the stigmatic surfaces of the angiosperm carpel germinate there, and the pollen tube grows toward the ovule to deliver two sperm cells through the micropyle (Raven et al., 1999; Simpson, 2006). Unlike consequences of double fertilization in the gymnosperms, double fertilization in angiosperms involves one sperm fertilizing the egg cell to form a diploid zygote, and the other fusing with polar nuclei to form endosperm. Angiosperm endosperm provides essential nutrition to the embryo, replacing that role of the female gametophyte in earlier spermatophyte lineages (Friedman and Williams, 2003; Friedman, 2008). Maturing embryos in the angiosperm seed are enclosed in two integuments, as are also found in gnetophyte seeds. The angiosperm seed is enclosed by additional sterile ovary and carpel tissue, forming the fruit, which is the angiosperm dispersal unit.
Angiosperms consist of two major lineages, and several less speciose ones. The
eudicot lineage contain about 75% of all angiosperm species, and diverged about 125
mya, while the monocots contain about 22% of all angiosperm species, and diverged about 120 mya (Magallon et al., 1999; Leebens-Mack et al., 2005). Angiosperms coevolved with insect pollinators and animal dispersal agents, and early angiosperm fossils show evidence of herbivory and insect pollination (Labandeira et al., 1994; Hu et al., 2008). However, the rapid radiation of 97% of all extant angiosperm species is not attributable to a few key innovations, but to several potentially interacting factors, including specialized pollination and dispersal modes, life history, habit, and geographical distribution (Davies et al., 2004; Sargent, 2004; Friedman and Barrett, 2008). The capacity of the angiosperm lineage to diversify in response to these factors was a function of the potential for evolution of development contained in the basal taxa.

**Angiosperm phylogeny**

A clear and accurate understanding of the evolutionary relationships among taxa informs comparisons of species and gene sequences, providing insight into relationships and lineage-specific phenotypes. Such an understanding is essential for constructing hypotheses about the evolution of development (Soltis and Soltis, 2003).

**Phylogenetic analysis**

A phylogenetic tree represents a hypothesis about the evolutionary relationships among taxa, inferred from the data used, and according to the model of evolution applied (Page and Holmes, 1998). Data used for phylogenetic analyses are derived from
observations of conserved features, including fossil evidence, morphological (e.g., embryological, vasculature, reproductive organs), palynological, karyological, and chemical characters. Character evolution is analyzed (originally, with parsimony methods) to make inferences about organism evolution. As technology has advanced, more and more molecular sequence data have become available (Soltis et al., 2005). Today, sequence data from the three plant genomes (chloroplast, mitochondrion, or nucleus) has largely replaced the use of physical and chemical features as the basis for phylogenetic analysis, and has brought resolution to long muddled questions, including the phylogenetic position of the monocots (Qiu et al., 1999; Doyle and Endress, 2000; Cai et al., 2006; Jansen et al., 2007; Doyle, 2008). The resulting hypothesis of organismal evolution provides a framework for analyzing the evolution of angiosperm development (Doyle and Endress, 2000; Soltis and Soltis, 2003).

**Selection of sequence data**

Sequences used for phylogenetic analysis must be conserved enough among the taxa of interest so that the sequences can be aligned, but they must have enough differences to be phylogenetically informative, distinguishing the taxa (Soltis and Soltis, 1998). For plants, chloroplast genome sequences are frequently used for phylogenetic analysis due to the small size of the genome, conserved function of the genes in the genome, and presence of most of these genes in single copy, in contrast with the multigene families found in the diploid nuclear genome (Soltis and Soltis, 1998). Chloroplast genes (protein-coding sequences as well as introns) vary in number of
nucleotides and rate of evolution, so different sequences can be chosen to provide resolution at different taxonomic levels (Soltis and Soltis, 1998).

If the sequence data are appropriate for the level of discrimination required, and the models of molecular evolution are appropriate for the data, the evolutionary relationships described from the analysis of the sequence data may be expected to accurately reconstruct organismal relationships. However, phylogenetic analysis of any given sequence will reflect the history of that sequence. Events other than vertical transmission will generate a topology that differs from that of the organismal history (Wendel and Doyle, 1998). Hybridization, introgression, and lineage-sorting obscure species-level phylogenies (Doyle and Davis, 1998; Rieseberg and Welch, 2002). Horizontal transfer has a similar effect, across even greater phylogenetic distances (Wendel and Doyle, 1998). Genes which have been transferred between the chloroplast and mitochondrial genomes include the photosystem II reaction center polypeptide D1 (psbA); the large subunit of ribulose bisphosphate carboxylase (rbcL); the cytochrome b6/f complex subunit 5 (petG), the photosystem II reaction center protein D2 (psbD), two conserved, large open reading frames frequently found in angiosperms (ycf1, ycf2), the photosystem I P700 apoprotein A1 (psaA), and an approximately 500 bp open reading frame (orf) (Appendix A).

Combining several genes for phylogenetic analysis limits the distortion any single gene can produce, and using a taxon-rich dataset reduces the potential for long-branch distortions of the resulting phylogeny (Leebens-Mack et al., 2005). The most current, well-supported hypothesis of evolutionary relationships among the angiosperms is based on analysis of 81 sequences from the chloroplast genomes of 64 taxa (Jansen et al.,
2007). That phylogeny (Figure 1-3), resolves ambiguities regarding the basalmost angiosperm lineages, and provides an exceptionally robust context in which to analyze evolution of features which historically may have been used as characters for phylogenetic analysis.

**Basal angiosperms**

Findings from basal angiosperms inform and polarize the interpretation of flowering plant features, and are necessary for inferring ancestral states and considering evolution of development in land plants (Endress, 2001a; Soltis et al., 2002; Soltis et al., 2005). The diverse habits, vegetative and floral forms, and specialized adaptations found in basal angiosperms display a level of diversity not found in more recently diverged lineages (Endress and Igersheim, 2000; Floyd and Friedman, 2001; Williams and Friedman, 2002; Friedman and Williams, 2004; Endress, 2006). Over one hundred macro- and microscopic morphological characters vary among the basal angiosperms and have been evaluated in a phylogenetic context (Doyle and Endress, 2000). In the context of well-resolved phylogenetic relationships, the diversity in these basal lineages can provide insight into the evolution of development of key innovations associated with angiosperm success. In particular, highly efficient tracheary elements, endosperm, and carpels with variably attached sterile tissues (including perianth and ovary tissues) are angiosperm innovations that became genetically fixed in later diverging lineages, after selective advantage of these features had been proved (Endress, 2001b).
Tracheary elements

Vessel elements are a unique tracheary element that have perforation plates at each end of the cell, providing greater efficiency of solute transport, (Raven et al., 1999). A kind of vessel element, with pore-like perforation plates, first appeared in the gymnosperm gnetophyte lineage, but has not been found in the conifers or other gymnosperms. Vessel elements with simple, scalariform, or reticulate perforation plates have been reported in the basal angiosperm families Hydatellaceae, Canellaceae, Piperaceae, and Saururaceae, but in many basal angiosperms, which are typically woody, only tracheids have been reported (Stevens, 2008).

However, closer examination using scanning electron microscopy revealed that vessel elements are not found in a bimodal state of presence or absence in the basal angiosperms (Carlquist and Schneider, 2002). A series of criteria can be used to describe the extent to which tracheary elements should be described as vessels or tracheids, based on cell dimensions, perforations, and associations with other cells (Carlquist and Schneider, 2002). Using these criteria, incipient vessel elements (in transition from tracheids) have been identified in all basal angiosperm families previously reported to be lacking them (Carlquist and Schneider, 2002; Stevens, 2008).

Figure 1-3: A consensus phylogeny of angiosperms. Adapted from Stevens (2008), incorporating added resolution based on whole plastid genome phylogenies of Jansen et al. (2007) and Moore et al. (2007). Species indicated are floral models. Blue shading designates basal angiosperm lineages, beige designates basal eudicot lineages, pink designates the rosid core eudicot lineages, yellow designates the asterid core eudicot lineages, green designates monocot lineages.
Double fertilization

Too sperm cells are delivered from the gymnosperm pollen tube, but typically only one sperm cell is functional. When double fertilization does occur, additional embryos are produced, not sterile, nutritive tissue (Carmichael and Friedman, 1996). In gymnosperms, the embryo obtains nutrition from the large female gametophyte which becomes enriched to nourish the embryo regardless of whether fertilization occurs (Friedman, 2008). Angiosperm double fertilization involves the fusion of one sperm nucleus with the egg nucleus and one sperm nucleus with a central cell nucleus. The associated production of endosperm in angiosperms provides them the advantage of deferring production of tissue for embryonic nutrition until after fertilization has taken place (Friedman et al., 2008).

Over 70% of extant angiosperms produce a seven-celled, eight-nucleate Polygonum-type female gametophyte from a single megaspore, and form triploid endosperm when a second sperm cell fuses with two haploid nuclei in the central cell (or one diploid nucleus, if the two haploid nuclei have already fused) (Friedman et al., 2008). Six other female gametophyte genetic constructs have been described for the angiosperms, each of which can be found among the basal lineages, but are rare among the core eudicots. The specific ontogeny of the female gametophyte determines the number and meiotic descent of the central cell nuclei, affecting endosperm ploidy, heterozygosity, maternal to paternal genome ratio, and potential for genetic conflict. Studies in basal angiosperms reveal that the female gametophyte can be derived from one, two, or four megaspores, and produce diploid (2N) (Nymphaeales, Austrobaileyales),
triploid (3N) (*Amborella, Polygonum*), 5N (*Penea, Plumbago*), 9N or 15N (*Peperomia*) endosperm (Williams and Friedman, 2002; Friedman et al., 2003; Williams and Friedman, 2004; Friedman, 2006; Arias and Williams, 2008).

The basal angiosperm family *Hydatellaceae* (Nymphaeales), a lineage with diploid endosperm, offers a missing link to the gymnosperms (Friedman, 2008). In *Hydatella* and its sister genus, *Trithuria*, megasporangia make a precocious investment to seed nutrition, building up starch in the female gametophyte (perisperm) prior to fertilization. After fertilization, nutrients cease to accumulate in the perisperm, and accumulate, to some extent, in the endosperm. The presence of both endosperm and perisperm nutritive tissues in the seed occurs frequently in basal angiosperms and may represent an intermediate step from gymnosperm lineages, which lacked endosperm (Friedman, 2008; Rudall et al., 2008). In addition, the frequent observation of double megagametophytes in the same ovule in *Trithuria* and other basal angiosperms underscores the developmental lability of basal angiosperm female gametophytes.

Organismal phylogenies can be used to interpret the evolution of female gametophyte development in angiosperms, but still lacking are the developmental genetics to support or test hypotheses of homology between basal and later diverging angiosperms, and between angiosperms and gymnosperms.

**Flower**

Stamens and carpels in the angiosperm flower derive from the leaf-like microsporophylls and megasporophylls found in gymnosperms, and the flower itself may
have originated by union of those reproductive organs to form modules, as is suggested by the close association of the reproductive organs in Gnetales (Endress, 2001b). Stamens in basal angiosperms may have thecae extending the length of the anther and open longitudinally (laminar stamens), as in Amborellaceae, Nymphaeaceae, Magnoliaceae (Judd et al., 2002). In other basal angiosperms, stamens have distinct anthers and a broad, laminar filament, with valvate anther dehiscence (e.g., Laurales). Overall, basal angiosperm stamens do not display pollination mode-specific canalizations of development observed in later diverging lineages. For example, in Poaceae, the anthers dangle from the filament to disperse pollen into the wind (Figure 1-2 C), and in the core eudicot *Kalmia* (Ericaceae), stamens are released from a retracted position when the pollinator lands, slapping the pollinator and releasing pollen onto its body (Figure 1-4 A).

Carpels in basal angiosperms are generally superior and apocarpous, often assembled on an elongated axis, as in Annonaceae and Magnoliaceae (Figure 1-4 B). However, carpels are half-inferior to inferior in *Aristolochiaceae* (Figure 1-4 C), and fully inferior in Monimiaceae (Figure 1-4 D). In *Aristolochia*, the inferior ovary is associated with a specialized structure, the gynostemium, which is composed of the thecae fused to the outer tissues of the style. Development of reproductive organs in *Aristolochia* and in less enigmatic genera in the family (e.g., *Saruma*) is well documented, and suggests transitions in form that may have led to the gynostemium (Gonzalez and Stevenson, 2000b, a; Simpson, 2006). The placentation in basal angiosperm carpels varies, thus the ovaries mature to produce a variety of fruits, including berries (Nymphaeaceae),
Figure 1-4: Floral specializations. A. Core eudicot adaptation: stamens release (black arrows) from retracted position to hit pollinator when disturbed (*Kalmia latifolia*; Ericaceae: Joel McNeal) B. Basal angiosperm flower with superior, apocarpous carpels on elongated axis, laminar stamens (*Magnolia tripetala*, Magnoliaceae: Kurt Stueber) C. Basal angiosperm flower with inferior ovary (*Aristolochia arborea*, Aristolochiaceae: T. Voelker) D. Basal angiosperm with hypanthium (black arrow) (male *Tambourissa*
follicles (Magnoliaceae), capsules (Aristolochiaceae), samaras (Hernandiaceae), drupes (Lauraceae), and aggregate fruits (Annonaceae).

In Tambourissa (Monimiaceae; Laurales), carpels and stamens are attached to a thickened hypanthium (Figure 1-4 D) (Endress and Lorence, 1983). The evolutionary trend from monoecy to dioecy has occurred repeatedly in this genus (Endress, 1992). The female Tambourissa flower tends toward forming a cup- or urn-like structure, rimmed by very small sterile organs (tepals), so that the reproductive organs are enclosed deep within. The hypanthium of the male Tambourissa flower tears as it opens to expose its many stamens. In one species of Tambourissa, a hyperstigma on the female flower blocks access to the carpels entirely, so that pollen germinates on the hyperstigma, and feeding pollinators (beetles) cannot devour the styles of the flower (Endress and Lorence, 1983).

Sepals and petals also first appear in the basal angiosperms, although organ identity and merosity are often indeterminate or variable (Endress, 1994). There may appear to be only one type of perianth organ, occurring in several spirally arranged whorls (tepals), as in Amborella, Nymphaeaceae, Austrobaileya, and several Magnoliales (e.g., Calycanthus, Magnolia) (Figure 1-4 B) (Endress, 2001b). Perianth organs may grade into laminar stamens, and include a whorl or two of organs displaying features of both tepals and stamens (intergrading floral organs). Perianth whorls that do display regular merosity are typically trimerous (e.g., Cabomba, Annonaceae, Lauraceae, and Asaroideae), as is found in the monocots.
Basal angiosperm perianth parts, and sometimes carpels, are colored for pollinator attraction, with pigments limited to reddish browns, white, yellow, cream, and green (Figure 1-4 B-D). Although the perianth organs in basal angiosperms usually lack fusion and are radially symmetric, *Aristolochia* (*Aristolochiaceae*, Piperales) has a monosymmetric unipartite perianth (Figure 1-4 C). The monosymmetric perianth of *Aristolochia* is associated with extensive speciation (Wanke et al., 2006b), as is floral monosymmetry in later diverging lineages (Sargent, 2004).

Some basal angiosperms lack perianth structures altogether, a feature which has evolved repeatedly in *Hydatellaceae*, *Chloranthaceae*, *Piperaceae*, and *Eupomatiaceae* (Endress, 2001b). These families are animal-pollinated nonetheless, as they use scent and heat to effect pollinator attraction. Many Magnoliales produce generally sweet or musky fragrances, and ethereal oil cells are found in some basal angiosperms. Thermogenesis in *Nymphaeaceae*, *Annonaceae*, and *Magnoliaceae* volatilizes fragrances produced by the plant and encouraged pollinators to remain active or breed (Thien et al., 2000; Seymour et al., 2003). Pollinators attracted to the basal angiosperm flower may be rewarded in a variety of ways, with stigmatic secretions, nectar, pollen (typically produced in copious amounts in basal angiosperms), breeding habitat, or additional sterile tissues presented for pollinator consumption, such as the staminodes of *Eupomatia* (Bergstrom et al., 1991) *Dipteran* pollinators attracted to fermenting, fruity, wine-like scents, and rotting meat or dead fish odors (*Austrobaileya, Illicium, Aristolochia, Asarum*), are succumbing to pollination by deceit, or may be responding to pheromones in the fragrance (Rulik et al., 2008).
**Floral developmental genetics**

Functional studies in current flowering plant models have shaped our understanding of floral developmental genetics and led to the discovery of broadly homologous traits, including conservation of floral organ identity genes among core eudicots and monocots (Pnueli et al., 1994; Samach et al., 1997; Kramer and Irish, 1999; Ambrose et al., 2000). These homologies, and others observed in current plant models (Figure 1-3), have suggested hypotheses about the common ancestor of monocots and eudicots. However, current eudicot models are in the highly diverged rosid (*Arabidopsis*) and asterid (e.g., *Antirrhinum, Petunia*) “core” eudicot lineages, which comprise over three quarters of all extant angiosperms (Magallon et al., 1999) and display lineage-specific, highly canalized floral forms (Endress, 1994; Endress, 2001b). Similarly, the monocot models *Zea* and *Oryza* are in Poaceae, a highly derived, species-rich family with a unique inflorescence and floral form. Genetic models such as the ABC model of floral organ identity developed for these later diverging lineages may not apply to basal angiosperms. Expression studies and genomics data from basal species provide insight into the genetics of flower development, showing that floral developmental genetics in basal angiosperms is far more labile than in later diverging lineages.

**Floral organ identity model**

The original ABC model proposed functional and evolutionary homology between genes in *Arabidopsis* and in *Antirrhinum* based on similarity of homeotic mutants, and, when data was available, sequence similarity (Coen and Meyerowitz, 1991;
Ma, 1994). Mutant analysis, overexpression studies, and in situ hybridizations suggested a set of homeotic genes, most belonging to the type II MADS family of transcription factors (Reichmann and Meyerowitz, 1997), were required to establish organ identity in sepal, petal, stamen, and carpel whorls. As originally conceived, A-function alone specified sepal identity, B- and A-functions acting together specified petal identity, B- and C- functions acting together specified stamen identity, C-function alone specified carpel identity, and A- and C-functions negatively regulated each other. Similar findings in the highly modified floral organs of Zea and Oryza demonstrated B function is required to establish lodicules, which are sterile organs positionally equivalent to petals, and C function is required for stamen organ identity (Mena et al., 1996; Ambrose et al., 2000).

In Petunia, another type II MADS family gene was discovered with the capacity to produce ectopic ovules on perianth organs, suggesting an ovule organ identity, or D-function (Colombo et al., 1995), but D function has not been discovered in any other species (Theissen, 2001). Three additional members of the type II MADS family, the SEPALLATA (SEP) genes, were found to physically interact with A,B and C whorl genes to transform vegetative tissues to floral ones in Arabidopsis. These were designated E-function genes (Honma and Goto, 2001). Phylogenetic analyses of plant and animal MADS family genes have indicated the type II A-E genes, or a progenitor, were present in the common ancestor of angiosperms and gymnosperms, and that some type II MADS genes were even present in mosses (Nam et al., 2003).
Genes involved in floral organ identity

The A function genes, *APETAL1 (AP1)* and *APETAL2 (AP2)* in *Arabidopsis*, are both involved in transition from an inflorescence meristem to a floral meristem, as well as in sepal and petal floral organ identity. However, as originally conceived, A-function, has only been clearly demonstrated in *Arabidopsis*. The *AP1* lineage of type II MADS-box genes has duplicated numerous times, including prior to the divergence of the core eudicots (Becker and Theissen, 2003; Nam et al., 2003), and, in the monocots prior to the divergence of the grass family (Preston and Kellogg, 2006). *AP2* is a member of an ethylene response factor (*ERF*) required for transition to a floral meristem in *Arabidopsis*. The euAP2 clade of these genes provides a cadastral function of restricting *AG* expression to the inner two whorls by translational repression attributable to miRNA172 (Chen, 2004; Kim et al., 2006). This cadastral function is not performed by identified A-function genes in *Antirrhinum*. Orthologues of both *AP1* and *AP2* genes have been found in basal angiosperms, but they are expressed broadly across the sterile and reproductive whorls in a manner inconsistent with organ identity genes (Litt and Irish, 2003; Kim et al., 2004; Kim et al., 2005a; Kim et al., 2005b; Shan et al., 2007). Furthermore, floral organ identity can be difficult to interpret in basal angiosperms. In *Amborella*, the sister to all other angiosperms, bracts grade to tepals, then to sporophylls, which are structurally homologous in the unisexual flowers (Buzgo et al., 2004b). There have been many duplications in the *AP1* and *AP2* lineages, and the ancestral role of A-function genes remains difficult to determine (Becker and Theissen, 2003; Nam et al., 2003).
Expression of the B function (petal identity) requires members of a different type II MADS-box gene lineage. These are APETALA3 (AP3) and PISTILLATA (PI) in Arabidopsis; DEFICIENS (DEF) and GLOBULOSA (GLO) in Antirrhinum. These B function genes arose after a duplication event occurring after the divergence of angiosperms from gymnosperms (Kramer and Irish, 2000; Kim et al., 2004; Zahn et al., 2005b). Additional duplication events, particularly just prior to the divergence of the eudicots and another prior to the divergence of the core eudicots, have provided opportunities for neofunctionalization and subfunctionalization (Force et al., 1999). The expression of B-function paralogs in core eudicots suggest the DEFICIENS (DEF) and GLOBULOSA (GLO) lineage has acquired diverse functions in roots, vasculature, leaves, and young seedlings (Zahn et al., 2005b). Expression of a GLO ortholog in the basal eudicot Eschscholzia californica, is consistent with ancestral function of DEF/GLO in petal and stamen organ identity (Zahn et al., 2005b), but in the basal eudicot Sanguinaria and the monocot Sagittaria, expression has been reported to be more broad, sometimes expanding into sepals and carpels (Kramer and Irish, 2000). Expression patterns of B-function homologs are more broad in many basal angiosperms, including Amborella, giving rise to the “fading borders” model of floral organ identity (Buzgo et al., 2004b; Kim et al., 2005b; Chanderbali et al., 2006).

The Petunia ovule identity (D function) genes are a subset of the C function subfamily, which includes AGAMOUS (AG) and SEEDSTICK (STK, formerly AGL11) from Arabidopsis (Zahn et al., 2006). The C (AG) and D (AGL11) clades are derived from a duplication in the type II MADS box lineage prior to the angiosperm radiation (Kramer et al., 2004), which would have allowed conservation of ancestral C function
while allowing neofunctionalization within the \textit{AGL11} lineage. The \textit{AGL11} lineage includes ovule-specific genes expressed in basal eudicots and basal angiosperms and would have afforded the ovule greater independence from the megasporophyte during development (Zahn et al., 2006). The ancestral role of \textit{AG} lineage genes, which are expressed in male and female gymnosperm cones, was likely carpel and stamen identity (Rutledge et al., 1998; Becker et al., 2003; Zhang et al., 2004). Although C-function genes are usually expressed only in the reproductive tissues of basal angiosperms, the presence of C-function genes in \textit{Illicium} and \textit{Persea} suggest the tepals of these taxa may originate from staminate primordia (Kim et al., 2005b; Chanderbali et al., 2006).

The effects of duplication are also evident in the E function (\textit{SEPALLATA}) genes, which were formerly named \textit{AGL2}, -3, -4, and -9, and which are not readily detected due to redundancy. Phylogenetic analysis suggests a duplication prior to the divergence of extant gymnosperms and angiosperms gave rise to the \textit{SEPALLATA (SEP)} genes, which have been found in basal angiosperms and basal eudicots, and are most closely related to the \textit{AGL6} lineage, which has been recovered in gymnosperms (Zahn et al., 2005a). Phylogenetic analysis reveal several duplications within the angiosperm \textit{SEP} lineage, with two copies of \textit{SEP} genes recovered in basal angiosperms, four in the rosids, and up to six in the asterids (Zahn et al., 2005a; Chanderbali et al., 2006). \textit{SEP} expression and function in plant tissues have evolved repeatedly, including roles in fruit maturation and plant architecture, as well as floral development (Malcomber and Kellogg, 2005). \textit{SEP} genes appear to have conserved expression in the floral meristem and ovules, although the conserved floral organ identity function has sometimes been lost, perhaps due to overlapping function with other genes. \textit{SEP} genes display lineage-specific differences in
sequence evolution, expression, and function, particularly within the monocots (Malcomber and Kellogg, 2004), demonstrating the opportunities for neo- and subfunctionalization provided by relaxed constraint on paralogs (Force et al., 1999; Lynch and Conery, 2000).

Rooting the developmental genetics of the floral organ identity model with information from basal angiosperms has revealed elements of the floral organ identity model which are not more broadly applicable. Unlike the four whorled flower of core eudicots, floral form varies widely in basal angiosperms, and is often characterized by gradual shifts from one type of floral organ to the next (Endress, 2001b). Gene expression studies in basal angiosperms have revealed broad patterns of conserved B and C function across the floral meristem, with “fading borders” of influence at the outer and inner edges of their area of expression (Buzgo et al., 2004a; Kim et al., 2004; Kim et al., 2005a). The core eudicot ABC regulatory pattern appears to be a relatively recent innovation, accompanying canalization of floral form and arising subsequent to duplications of ancestral floral organ identity genes (Kramer and Irish, 2000; Kramer and Zimmer, 2006). An earlier model of floral organ identity (“B, C model”), in which only petals and carpels are specified, might better apply to basal angiosperms (Schwarz-Sommer et al., 1990).
Broader applications

"The rapid development as far as we can judge of all the higher plants within recent geological times is an abominable mystery." (Darwin, 1903)

Beyond floral organ identity, existing studies in model systems have generated many hypotheses about ancestral features of flowering plants and the developmental changes that led to early-diverging lineages, as well as to the more canalized eudicots and monocots (Endress, 2001a; Floyd and Friedman, 2001; Soltis et al., 2002; Soltis et al., 2005). Several genome-scale duplications within the angiosperms have provided the raw material for conservative experimentation with new gene functions (Cui et al., 2006). The history of those genes and their roles in evolution of development will provide – at least in part - the solution to Darwin’s abominable mystery.

Floral symmetry, for example, is ancestrally radial, but important lineages in both the monocots and core eudicots display monosymmetry. The first gene to be associated with floral symmetry was CYCLOIDEA (CYC), discovered in the asterid, Antirrhinum (Lamiales) (Luo et al., 1996). Duplication and evolution of CYC genes is implicated in floral form evolution in a later diverging asterid lineage, Dipsacales, and is also required for monosymmetry in the rosids Lotus, Pisum, and Lupinus (Fabales) (Ree et al., 2004; Howarth and Donoghue, 2005; Feng et al., 2006; Wang et al., 2008). CYC genes have also been implicated in the bisymmetric flowers of the basal eudicots (Papaveraceae; Ranunculales) (Damerval et al., 2007; Damerval and Nadot, 2007). The extent to which CYC genes play a role in the monosymmetry of the basal angiosperm, Aristolochia fimbriata, has yet to be determined.
Evolution of development in basal angiosperms can be studied by comparative studies of gene function between systems, as well as by studies of gene family evolution. Functional analyses employ classical and molecular genetics, including mutant analysis, complementation experiments, and over- and under-expression studies to derive an understanding of gene function. Reverse genetics, employing transformation methods, are effective ways to test hypotheses about gene function. In order to test hypotheses about the roles of B and C function genes in basal angiosperms, or hypotheses about the role of CYC genes in basal angiosperms, it will be necessary to have a basal angiosperm experimental model system.

**Basal angiosperm model**

The purposes of a model system are, primarily, to represent the taxa to which it belongs, to provide an easier to use system in which to conduct experimentation, and to facilitate inductive reasoning so conclusions drawn from the model system can be applied more generally (Bolker, 1995). To support those purposes, a model system must be in a phylogenetically informative position, ideally derived from phylogenetically nested taxa so comparisons can be made to closely related species (Baum et al., 2002). To be useful, a “good” model system should have features that support its current and continued use. In biology, this means the model organism must be easily obtained, demanding of few resources, fast growing, productive of ample tissues as needed, and amenable to widely-used technologies.
The highly successful plant model *Arabidopsis thaliana* meets all of these criteria for the core eudicots, particularly the highly successful rosids (Arabidopsis, 2000).

*Arabidopsis* has structures broadly representative of closely related species, and can be readily compared with *Brassica*, a closely related genus in the economically important family, Brassicaceae. *Arabidopsis* is readily obtained, has a short generation time, easy to follow and well-developed protocols for commonly used technologies, and produces ample tissue and seed in a small amount of space. As a result, *Arabidopsis* can be used in forward genetics, in which mutants are analysed to gain information about genetic processes, and it can be exceptionally easily transformed and used in reverse genetics, for functional analyses. It also has a very small genome, which has facilitated genome mapping, sequencing, and studies of gene function (Arabidopsis, 2000). The large community of researchers using *Arabidopsis* has contributed crucial value-added resources (e.g., libraries, genome sequence and expression databases, detailed developmental studies, mutant collections, seed banks), supporting shared problem-solving and leading to rapid advances in understanding. A careful review of basal angiosperms was conducted to identify the best possible candidate for model system development, keeping the *Arabidopsis thaliana* model system in mind as the “gold standard” to be met.

**Transformation systems**

A genetic transformation system is essential for generating transgenic plants in which hypotheses about the evolution of development in basal angiosperms can be tested.
*Agrobacterium*-mediated methods of transformation use the machinery found in the soil bacterium, *Agrobacterium*, to transfer DNA to plant cells and integrate it into the plant genome (Gelvin, 2000). In the commonly used binary vector method, the tumor-inducing (Ti) plasmid in the *Agrobacterium* is disarmed by removing the section of endogenous DNA that is transferred (T-DNA) to the plant cell to induce tumor formation.

The endogenous virulence (*vir*) genes, which are essential for effecting the transfer and subsequent integration of the T-DNA, remain on the disarmed plasmid. The genetically modified *Agrobacterium* are induced to take up a binary plasmid containing origins of replication for *E. coli* and *Agrobacterium*, the endogenous right and left border sequences from the Ti plasmid flanking a T-DNA, as well as a selectable marker and multiple cloning sites. The T-DNA typically includes at least one marker for selection in plant cells, a selectable reporter gene, the exogenous gene of interest to be evaluated in the plant system, the appropriate promoter sequences, and other sequences to improve transformation (e.g., matrix attachment regions) (Maximova et al., 2003).

When plant tissues are infected with *A. tumefaciens*, the *vir* genes are activated to transfer the TDNA in the genetically modified plasmid to the plant chromosome (Gelvin, 2000; Gelvin, 2003). After a period of co-cultivation with *A. tumefaciens*, the plant material is subjected to chemical selection to eliminate non-transformed cells, and counter-selection to kill the *Agrobacterium*. Aminoglycoside antibiotics are toxic to plant cells to some extent, so the neomycin phosphotransferase (*NPTII*) selective marker gene can be constitutively expressed in transformed cells to confer resistance to aminoglycoside antibiotics (Norelli and Aldwinckle, 1993; Petri et al., 2005). The transformed tissue is cleared of *A. tumefaciens* with an antibiotic that is effective against
the Gram-negative cells of *Agrobacterium* but is not toxic to plant cells. The transformed plant cells are regenerated under ongoing conditions of selection to prevent regrowth of bacteria or non-transformed cells.

**Regeneration systems**

Transformation of *Arabidopsis* was originally implemented *in vitro* (An et al., 1986). *In planta* floral dip and spray methods were developed later (Clough and Bent, 1998; Chung et al., 2000). Other plant species still rely on tissue culture-based transformation systems (Ishida et al., 1996; Negrotto et al., 2000), which permit greater control of variables which influence plant growth, including light, water, temperature, nutrients, pH, container, carbohydrate source, and condition of source material (Mohamed et al., 1992; Razdan, 2003; Vega et al., 2008). Plant tissues or callus in culture can often be induced to regenerate in response to treatment with plant growth regulators (PGRs).

PGRs influence development both absolutely and in proportion to one another (Razdan, 2003). Auxins were identified in 1926 by the ability of indole acetic acid, a naturally occurring plant hormone, to promote growth in oat coleoptiles (Went, 1926). Cytokinins were identified in 1948 by their ability to stimulate cell division (Caplin and Steward, 1948). Skoog and Miller (1957) showed that organ formation depended on the balance between cytokinin and auxin. High concentrations of auxin promoted root formation, and high concentrations of cytokinin promoted shoot formation. Theoretically, a 1:1 ratio of cytokinin to auxin could be expected to produce callus (Razdan, 2003).
Plant tissues differ in response to PGRs, which also differ in potency with respect to their ability to induce growth or cell division. The synthetic plant growth regulator thidiazuron (TDZ) has properties of both auxin and cytokinin and is particularly effective at inducing shoot primordia in the first phase of shoot organogenesis (Murthy et al., 1998; Chen and Chang, 2001; Wu et al., 2004). In comparison, somatic embryogenesis protocols begin with a very high auxin level to induce formation of pro-embryos, which are then removed to a more balanced ratio of PGRs for regeneration (Juretic and Jelaska, 1991). The synthetic auxin (2, 4, -D) is particularly effective for inducing somatic embryos (Li et al., 1998; Maximova et al., 2003). The presence of antibiotic compounds in growth media typically alters the response of plant tissue to plant growth regulators, so regeneration protocols must be optimized for transformed cells in the presence of selection agents (de Mayolo et al., 2003).

**Overview of dissertation**

There currently is a need for a basal angiosperm model system amenable to genomics studies and experimentation using methods of developmental biology, molecular biology and classical genetics. Towards this end, I undertook a study to select the best available basal angiosperm with a small genome, self-compatibility, closely related family members with which it could readily be compared, and the ability to be transformed with a genetic marker. Fourteen species were cultivated and evaluated. I developed hand pollination methods and demonstrated self-compatibility for Aristolochia elegans and Aristolochia fimbriata (A. fimbriata), two species with small genome sizes. I
then used those methods to develop in-bred lines to serve as a source of tissue and seed for the NSF supported (DEB 0638595) Ancestral Angiosperm Genome Project (AAGP).

I developed and implemented standardized tissue collection methods for *A. fimbriata* greenhouse and seedling tissues from which AAGP generated expressed sequence tags (ESTs). I used a disarmed strain of *Agrobacterium tumefaciens* to transform *Saruma henryi* and two genotypes of *A. fimbriata* with a plasmid conferring constitutive expression of an *NPTII* antibiotic resistance gene for selection of transformed tissues, and an enhanced green fluorescent protein (*EGFP*) marker gene for identifying transformed cells. Chapter 2 includes the manuscript covering that study, including the survey of basal angiosperms and experimental results that led to the selection of *A. fimbriata*, as well as hand pollination methods and preliminary EST sequencing results.

*In vitro* micropropagation and shoot organogenesis methods were developed to support the transformation protocol, as well as to provide *in vitro* tissue for biochemical analysis of the secondary metabolites produced by *A. fimbriata*. In addition to micropropagation methods, I developed a direct shoot organogenesis method using small pieces of petiole explants, on which shoot primordia were induced in 14-21 days and regenerated within three months. Chapter 3 includes the manuscript of a paper describing the *in vitro* propagation and shoot organogenesis methods for *A. fimbriata*.

Chapter 4 includes preliminary findings from a study of TCP gene family evolution using the available sequences from fully sequenced angiosperm genomes, as well as ESTs from basal angiosperms. This study will be incorporated into a manuscript after further analysis is complete. Chapter 5 reports the conclusions of my work and discusses what should be done next to advance the development of the *A. fimbriata*
model system, including recommendations for early experiments to conduct in a basal angiosperm experimental model. Two appendices follow: Appendix A reviews my work on organelle genome evolution, documenting horizontal transfers from angiosperm chloroplasts to mitochondria. Appendix B is a cultivation supplement, presenting results of seed germination data from several experiments, illustrations of hand pollination methods and descriptions of cultivation methods in greater detail than that provided with the manuscript included in Chapter 2, instructions for using the seed pedigree system I devised, and documentation of two mutant phenotypes observed during the AAGP seedling tissue collection efforts.
Chapter 2

*Aristolochia fimbriata: A proposed experimental model for basal angiosperms*

This chapter contains a manuscript of an article that will be submitted for publication by Barbara Joanne Bliss, Stefan Wanke, Abdelali Barakat, Patrick Kerr Wall, Lena Landherr Sheaffer, Yi Hu, Hong Ma, Christoph Neinhuis, Jim Leebens-Mack, K Arumuganathan, Sandra Clifton, Siela N. Maximova, Claude W. dePamphilis* to Plant Physiology.

This work resulted from collaboration among the labs of Drs. Claude dePamphilis, Hong Ma, and Siela Maximova. Dr. Stefan Wanke produced the phylogeny of the taxa sampled and contributed to polymerase chain reaction (PCR) methods development to verify the green fluorescent protein (GFP) marker gene. Dr. Abdelali Barakat contributed miRNA genomics resources for *A. fimbriata*. Lena Landherr constructed *A. fimbriata* cDNA libraries, conducted transformation experiments and optimized PCR methods to verify transgenic tissues. Dr. Jim Leebens-Mack and Sandra Clifton sequenced libraries and managed EST data. Dr. Kerr Wall conducted bioinformatic analyses of EST data and resulting unigenes. Yi Hu assisted with maintenance of *in vitro* and greenhouse cultivated plants, tissue collection, and RNA extraction. Dr. K Arumuganathan conducted genome size analysis. All other results were generated by Barbara Bliss.
Running head: *Aristolochia fimbriata* as a basal angiosperm model

Corresponding author: Claude W. dePamphilis, Department of Biology, Institute of Molecular Evolutionary Genetics, and the Huck Institutes of the Life Sciences, 201 Life Sciences Building, Pennsylvania State University, University Park, PA 16802, US (814) 863-6412 cwd3@psu.edu

Journal research area: Genetics, Genomics, and Molecular Evolution
Title: *Aristolochia fimbriata*: A proposed experimental model for basal angiosperms

Author full names: Barbara J. Bliss, Stefan Wanke, Abdelali Barakat, P. Kerr Wall, Lena Landherr, Yi Hu, Hong Ma, Christoph Neinhuis, Jim Leebens-Mack, K Arumuganathan, Sandra Clifton, Siela N. Maximova, Claude W. dePamphilis*

Institution addresses: Department of Biology, Institute of Molecular Evolutionary Genetics, and the Huck Institutes of the Life Sciences, 201 Life Sciences Building, Pennsylvania State University, University Park, PA 16802, USA (B.J.B., P.K.W., L.L.S., Y.H., H.M., C.W.D.); Technische Universität Dresden, Institut für Botanik, Plant Phylogenetics and Phylogenomics Group, D-01062 Dresden, Germany (S.W., C.N.); Benaroya Research Institute at Virginia Mason, Flow Cytometry and Imaging Core Laboratory, 1201 Ninth Avenue, Seattle, WA 98101, USA (K.A.); Department of Plant Sciences, University of Georgia, Athens, GA 30602 (J. L-M.); Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA (S.C.); Department of Horticulture, 421 Life Sciences Building, Pennsylvania State University, University Park, PA 16802, USA (S.N.M.)
Footnotes:

Financial sources: This work was supported in part by NSF Plant Genome Research Program, awards DBI-0115684 (The Floral Genome Project) and DEB 0638595 (The Ancestral Angiosperm Genome Project), by the Department of Biology and Huck Institute of Life Sciences of the Pennsylvania State University, and by a postdoctoral grant to SW from the German Academic Exchange Service (DAAD).

Corresponding author; email Claude W. dePamphilis, cwd3@psu.edu
Abstract

“Basal angiosperms” are still lacking a model system, which is crucial for understanding flowering plant evolution via comparative approaches. An evolutionary perspective is essential for designing genetic, developmental, and comparative genomic studies of angiosperms. Using a broad approach among basal angiosperm families, we first reviewed general characteristics critical for model systems. These included availability to a broad scientific community, growth habit, and suitability to represent basal angiosperms displaying a wide spectrum of phenotypic diversity. Most basal angiosperms, predominantly woody or aquatic, were excluded using these criteria. Aristolochiaceae were investigated further for ease of culture, life cycle, genome size, and chromosome number. We demonstrated self-compatibility for Aristolochia elegans and A. fimbriata, and transformation with a GFP reporting construct for Saruma henryi and A. fimbriata. Consequently, we recommend Aristolochia fimbriata as a basal angiosperm model system. Native to Brazil, this rapidly growing perennial is widely cultivated with a seed-to-seed life cycle of just 3 months. Tissue culture regeneration methods are available for this species to support a high-throughput transformation system. Emerging genetic and genomic tools will aid the investigation of floral biology, developmental genetics, biochemical pathways important in plant-insect interactions and human health, and various other features present in early angiosperms.
Introduction

Our present understanding of genetics, genomics, development, evolution, and physiology of living organisms is largely based on work done in model systems. Model systems provide more favorable conditions for observing phenomena and testing hypotheses than other systems afford. Models support inductive reasoning, in which one builds on the understanding of living organisms in general, based on observations made in a specific model organism (Bolker, 1995). Model organisms also provide a focus for researchers to work on a common system, resulting in collaborative and complementary efforts which can yield rapid progress and development of further resources. Models are crucial for understanding basic biological processes.

Models should have several key attributes. At the very least, model organisms should be good representatives of the taxon or process they are chosen to represent (Kellogg and Shaffer, 1993; Bolker, 1995) and they must be accessible (Kellogg and Shaffer, 1993) so that a broad community of scientists can utilize and develop them. Models used for developmental and genetic studies must also offer rapid development, short generation time, be amenable to large scale cultivation, and provide ample tissue for experimentation (Bolker, 1995). Models should also support forward and reverse genetics, as is required for hypothesis testing (Soltis et al., 2002), and have a small genome size, to facilitate molecular genetics and genomics work as well as genome sequencing (Pryer et al., 2002). Finally, for studying the evolution of development, models should support studies among closely related species, so that comparative studies can be interpreted at the molecular level. To the extent that model systems exemplify the
taxa to which they belong, deep studies of the models, along with comparative studies between them and closely related species, can inform and advance theory. The lack of appropriate experimental models limits development of theoretical ones, and such is the case in plant biology today.

Our current understanding of angiosperm evolution has been shaped by multiple phylogenetic studies (Jansen et al., 2007; Moore et al., 2007; Stevens, 2008) which provide the context in which the evolution of any aspect of flowering plants is studied. Of particular interest to both basic and applied plant biology are changes leading to the success and diversification of angiosperm lineages, beginning with the early, mostly species poor lineages of angiosperms (Table 2S-1) previously known as the ANITA grade, followed by the magnoliids (Figure 2-1). The magnoliids contain four major branches and several thousand species (Table 2S-1). Among the species in this lineage are the tree genome model, *Liriodendron tulipfera* (Liang et al., 2007) and the tree fruit model, *Persea americana* (Awad and Young, 1979), as well as the most economically important spice, black pepper (*Piper nigrum*) (Wanke et al., 2007). Taxa included in the magnoliids exhibit many features first appearing in basal angiosperms or angiosperms in general e.g., vessel elements, perianth monosymmetry, alkaloid chemistry, specialized pollination systems, diverse forms of female gametophyte development (Soltis et al., 2002; Soltis et al., 2005; Madrid and Friedman, 2008). The magnoliids are the sister clade and therewith the closest outgroup to the highly diverged monocot and eudicot lineages (Jansen et al., 2007).
Figure 2-1: Angiosperm phylogeny based on Stevens (2007) (http://www.mobot.org/MOBOT/research/APweb/welcome.html). Important model systems and the proposed model, *Aristolochia fimbriata*, are shown along the corresponding clade. *Liriodendron*, *Persea*, *Populus*, and *Carica* are tree species. Species that have been used as flower development models are indicated with an asterisk.
The rich diversity of basal angiosperms provides glimpses into early successful experiments in angiosperm adaptation (Endress and Igersheim, 2000; Floyd and Friedman, 2001; Soltis et al., 2002; Williams and Friedman, 2002; Friedman and Williams, 2004; Endress, 2006). Novel features which are otherwise constrained by function may have evolved more than once, in parallel, such as the structural and developmental similarities of the inflorescence of the monocot *Acorus* with that found in some Piperales (Piperaceae, Saururaceae; members of Magnoliids) (Buzgo and Endress, 2000). Basal lineages retain evidence of “trials” of features that became genetically fixed in later lineages (Endress, 2001b). For example, perianth parts are much more variable in basal angiosperms, as well as basal eudicots and basal monocots, and only became fixed in the core eudicots (Irish, 2006).

Current flowering plant genetic models are all derived from the highly diverged monocot and eudicot lineages (Figure 2-1). Among them, the monocot models *Zea* and *Oryza* occur in Poaceae, which has specialized floral organs and an inflorescence found only in that family. Similarly, the current eudicot models are derived from the highly diverged rosid (*Arabidopsis*) and asterid (*e.g.*, *Antirrhinum*, tomato) “core eudicot” lineages, each of which displays lineage-specific floral forms (Endress, 1994; Endress, 2001b). Studies in current plant models have led to the discovery of broadly homologous traits, including conservation of floral organ identity genes (ABC/quartet models) (Coen and Meyerowitz, 1991; Ma and dePamphilis, 2000; Theissen, 2001). These homologies, and others observed in current plant models (Figure 2-1), have suggested hypotheses about the common ancestor of monocots and eudicots.
An understanding of the evolution of floral development or any fundamental process in flowering plants will require a basal angiosperm experimental model in which to test these hypotheses (Baum et al., 2002; Soltis et al., 2002; Buzgo et al., 2004a; Soltis et al., 2006; Endress, 2007). Although the current model systems represent well the highly successful and derived lineages in which they occur, they do not represent the overall diversity of angiosperms. Information from basal lineages is necessary both to better describe that diversity, to polarize the changes that occurred angiosperm evolution, and to make functional inferences about the common ancestor of early angiosperms.

Current plant models have been selected to address particular questions, but very few are available for use both as genomic and as genetic models. Genomic resources, which emphasize plants with small genomes, have been developed for tree wood and fruit species, including *Populus* (Jansson and Douglas, 2007), *Liriodendron* (Liang et al., 2007), *Persea* (Albert et al., 2005) and *Carica* (Wei and Wing, 2008). However, woody species are too large at maturity and do not have short enough life cycles for general use in genetics experiments. Forward genetics requires a very small organism with a rapid life cycle, and is facilitated by the ability to self-pollinate individuals with desired characteristics. Reverse genetics requires manipulation of DNA or RNA in a targeted manner. Both benefit from a small genome, and transformability is essential for testing hypotheses about gene function. Therefore, we sought to identify a basal angiosperm species having as many important features of a model system as possible to support widespread use as an experimental model in genetics and genomics. We present these essential features in *Aristolochia fimbriata* - small size at maturity, rapid life cycling,
selfcompatibility, small genome size, and transformability - along with relevant findings for other taxa evaluated in our study.

**Results**

**Evaluation of potential models in basal angiosperm orders and families**

We followed a formal selection process to identify a suitable candidate for an experimental model organism among basal angiosperms (Figure 2-2). Many basal angiosperms are uncommon, with limited distribution, and often occur in families with only one genus and few species (Table 2S-1). Taxa were considered readily accessible to a broad scientific community if they could be obtained commercially at a low price and could be readily propagated by seed. Those of limited availability were eliminated, including Amborellales (Amborellaceae), Austrobaileyales (Illiciaceae, Austrobaileyaceae, Trimeniaceae), and Chloranthales (Chloranthaceae) (Figure 2-2). Woody plants requiring a year or more to attain maturity, as well as extensive space for cultivation, were eliminated (Fig 2). These included most Magnoliales (Myristicaceae, Magnoliaceae, Annonaceae, Himantandraceae, Degeneriaceae, Eupomatiaceae), Laurales (Lauraceae, Hernandiaceae, Monimiaceae, Atherospermataceae, Gomortegaceae, Siparunaceae, Calycanthaceae), and Cannellales (Cannellaceae, Winteraceae). The forest tree model species, *Liriodendron* (Magnoliaceae) and the fruit tree model *Persea* (Lauraceae), were among those eliminated due to woody habit and long generation time.
Basal angiosperms

Criteria applied

Accessibility
- Limited
  - Amborellales, Chloranthales, Austrobaileyales

Growth habit
- Woody
  - Magnoliidaes, Laurales, Cannellales
- Aquatic
  - Nymphaeales, Ceratophyllales

Piperales

General angiosperm features
- Parasitic
  - Hydnoraceae
- Highly reduced flowers
  - Piperaceae, Saururaceae

Aristolochiaceae

Rapid life cycle
- Cold period required
  - Saruma, Asarum, A. serpentaria, A. tomentosa, A. macrophylla
  - More than six months to flower
  - A. anguicida, A. clematitis, A. trilobata, A. tagala, A. californica, A. passiflorafolia

Ease of culture
- Large vine
  - A. grandiflora, A. gigantea, A. ringens, A. elegans

Small genome
- Large genome
  - Saruma, Asarum, Pararistolochia

Self-compatibility
Transformability

Aristolochia fimbriata
Many of the families listed above would also be eliminated due to limited commercial availability, as e.g., Trimeniaceae, Himantandraceae, and Gomortegaceae (Table 2S-1). Similarly, many of the families eliminated for limited accessibility have a woody habit (Table 2S-1); these would have been eliminated for that character if they had been more accessible (e.g., Amborella trichopoda, the sister of all other flowering plants). Plants with an aquatic habit, including water lilies (Nymphaeales: Nymphaeaceae, Cabombaceae) and Ceratophyllum (Ceratophyllales: Ceratophyllaceae), were similarly eliminated due to intensive cultivation requirements (space, maintenance) (Figure 2-2). The order Piperales (Aristolochiaceae, Hydnoraceae, Piperaceae, Saururaceae, Lactoridaceae) contains several herbaceous taxa (Table 2S-1). Parasitic plants (Hydnoraceae) and those with highly reduced flowers (Saururaceae, Piperaceae) do not generally represent angiosperms, and so were eliminated (Figure 2-2). Hydnoraceae and Lactoridaceae would also be excluded due to limited availability (Table 2S-1). Among the basal angiosperm families, only Aristolochiaceae contains highly accessible, easily cultivated herbaceous plants with features broadly representative of angiosperms in general.
Aristolochiaceae candidates considered

We surveyed Aristolochiaceae, seeking species with rapid growth, no requirement for vernalization in the life cycle, ease of large scale cultivation, and a small genome size...
to facilitate gene cloning and eventual genome sequencing (Figure 2-2). Members of genus *Aristolochia* have some of the smallest basal angiosperm genome sizes currently known (Figure 2-3). Therefore, we evaluated each genus in this family to identify the best species for model system development. Subfamily Asaroideae genera (*Asarum, Saruma*) (Figure 2-4A,B) require cold treatment to induce flowering, resulting in increased culturing efforts and extended time to flower, so they were eliminated (Figure 2-2). *Thottea*, here represented with *T. siliquosa* (Figure 2-4C), was not possible to obtain for a detailed culture survey (Table 2S-2), as it can only be cultivated with very high maintenance, in a narrow range of conditions. Furthermore, it grows slowly, produces little tissue, few flowers, and few seeds. Of the *Aristolochia* species available for culturing, those requiring cold treatment (*A. serpentaria, A. clematitis* of subg. *Isotrema*, Figure 2-4D) were eliminated from consideration. Members of subgenus *Pararistolochia* were not available for culturing (Figure 2-4E). Those species that did not bloom in six months (*A. californica, A. anguicida, A. macrophylla, A. tomentosa*), produced very few flowers (*A. trilobata, A. passiflorafolia*; Figure 2-4F, G), or formed very large vines (*A. grandiflora, A. ringens, A. labiata, A. gigantea*) (Table 2S-2, Figure 2-2) were also eliminated. The remaining candidates that met our criteria were two smaller members of subgenus *Aristolochia* (*A. elegans and A. fimbriata*), belonging to a group of subtropical and tropical species from South and Middle America (Neinhuis et al., 2005; Wanke et al., 2006a).
Physical and life cycle features of *Aristolochia fimbriata*

We characterized the physical and life cycle features of *Aristolochia fimbriata* (Figure 2-4H) to further assess its potential as a model system. Seeds planted in potting
medium in the greenhouse germinated at rates up to 100%, and flowered in as few as 62 days after planting. Cultivated as a small pot crop in the greenhouse, *A. fimbriata* stock plants occupy minimal bench space (Figure 2-5A), and were not particularly susceptible to any pest or pathogen, though commercial pesticide treatments were applied greenhouse-wide as needed. Vines are supported by a small trellis during periods of flowering or fruit ripening (Figure 2-5B) to prevent mechanical damage and facilitate fruit harvest. Plants flower copiously from axillary nodes on multiple indeterminate stems that arise from the woody root. A new flower opens every two to three days along the stem (Figure 2-5C) facilitating collection of staged tissues. Fruits from open-air pollinations on one-year-old plants averaged over one hundred seeds per capsule.

Although seasonally green in its native habitat, *A. fimbriata* grows year round in greenhouse culture. It forms a large perenniating root that can be divided to produce clones. Stems can be pruned back to the root after fruit collection to support greenhouse sanitation efforts, or “as needed” to stimulate new growth. New stems generated in this way will flower in two weeks. Fruit size, leaf size, and number of seeds increase in older, larger plants. The longevity of individual plants provides an ongoing source of seed from a single experimental subject.
We developed in vitro methods for germinating seeds to support collection of seedling tissues and for comparisons of seed viability. Light during germination is required for true leaves to emerge. Seeds germinated significantly better in wet toweling (mean= 61%, n=11) compared to germination on plates containing solid, sucrose-free media (mean=9%, n=3). Light and age of seed (up to 2.5 years) had no significant effect on germination with 80% germinating by 70 days. More details are available in the Cultivation Supplement.
A phylogenetic perspective of genome sizes

Because a small genome size facilitates cloning genes of interest, detecting regulatory regions, and sequencing the genome of a model organism, our analysis of genome size evolution in Aristolochiaceae focused on species having small genomes, particularly in subgenus Aristolochia, for which we report here the smallest genome size to date from a basal angiosperm (*A. lindneri*) (Figure 2-6). In order to gain insight into the evolution of genome structure as well as genome size in Aristolochiaceae, chromosome numbers from previously published studies (Sugawara, 1987a; Ohi-Toma et al., 2006) were plotted along with genome sizes on a strict consensus tree using TreeGraph (Müller and Müller, 2004) (Figure 2-6).

A direct correlation between chromosome numbers and genome size for the family of Aristolochiaceae was not observed. Aristolochiaceae is subdivided into two subfamilies, Asaroideae and Aristolochioideae. Asaroideae, which has 2n= 26, 52 (*Saruma henryi*) and 2n=26 (*Asarum caudatum*) chromosomes, has about two to ten times the genome size of genus *Aristolochia* (Figure 2-6). *Thottea*, the earliest diverging branch in subfamily Aristolochioideae has the same number of chromosomes (2n=26) as *Asarum* and *Saruma*, but has only 1/9 and 1/5 of the genome size of *Asarum* and *Saruma*, respectively.

In contrast, within genus *Aristolochia*, species in subgenus *Aristolochia* exhibit the smallest genome sizes, but have a wide range of chromosome numbers. Species in subgenus *Isotrema* (Figure 2-4D) are generally characterized by 2n=32 chromosomes
and have small genome sizes (554-774 Mbp mean genome size, 1C). Within subgenus *Pararistolochia* (Figure 2-4E) (1793-4321 Mbp mean genome size, 1C), which is sister

---

**Figure 2-6:** Phylogenetic relationships among sampled Aristolochiaceae, with *Tasmannia lanceolata* (Cannelleales) as outgroup. Maximum parsimony strict consensus tree with genome sizes (Mbp/1C) and chromosome counts indicated. If different genome sizes were obtained from different plants belonging to the same species, the smallest size was plotted on the tree. For range of genome sizes within one species and standard deviation refer to Table 2S-5. Bootstrap values from 1000 replicates are indicated on the branches.
group to subgenus *Aristolochia*, and therefore nested within the clade having species with small genomes, a lineage-specific increase in genome size can be seen. The increase is not associated with an increase in chromosome number, so it can be inferred that there has been an increase in chromosome size. The Australian species of *Pararistolochia* have less than half of the genome size of the African species, but still have the same number of chromosomes (Ohi-Toma et al., 2006). Within the subgenus *Aristolochia* (*Figure 2-4F, G, H*), the different monophyletic groups recovered (*Figure 2-6*) are in accordance with previous studies (Wanke et al., 2006a; Wanke et al., 2007). It is interesting to notice that *Asarum* and subgenus *Pararistolochia* have large chromosomes similar to those of monocots, whereas the remaining clades in Aristolochiaceae have small chromosomes (Ohi-Toma et al., 2006). *Aristolochia fimbriata* is a member of a clade of *Aristolochia* species with 2n=14 chromosomes and genomes roughly the size of *Oryza sativa*.

**Methods for genetics**

We further evaluated selected taxa for self-compatibility and potential for genetic engineering, both of which are critical features of genetic model systems. Self-pollination experiments were conducted with *A. elegans* and *A. fimbriata* because of their small genomes, ease of culture, and prolific flowering (*Figure 2-2, Table 2S-2*). Both species could be hand pollinated to accomplish cross- and self-pollination events using simple methods (*Tables 2S-3, 2S-4*). Morphological changes in the perianth and gynostemium associated with maturation of the anthers and stigmatic surfaces are described for *A. fimbriata* from the day of anthesis (day 1) through day 3 (*Figure 2S-1*). Self-pollination
in *A. fimbriata* was most effective on day 2, both in terms of fruit production and seed viability (*Figure 2S-1, Table 2S-3*). Using *in vitro* germination methods, 59% of seeds produced in open-air pollinations of *A. fimbriata* germinated normally, compared to 50% of seeds produced from self-pollination of day 2 flowers (*Table 2S-3*). Additional details are provided in the Cultivation Supplement.

*Agrobacterium tumefaciens*-mediated genetic transformation experiments were performed with two varieties of *A. fimbriata* (NV, VL) and with *Saruma henryi*. The binary vector utilized (Maximova et al., 2003) contained a neomycin phosphotransferase marker gene (*NPTII*) for antibiotic selection and an enhanced green fluorescent protein reporter gene (*EGFP*); both genes were under the control of the E12-Ω CaMV-35S promoter (Maximova et al., 2003). High frequency transient expression was observed in the leaf explants at seven days after culture initiation (*Figure 2-7A-D*), followed by production of transgenic calli and shoot primordia at 24 days after culture initiation (*Figure 2-7E, F*). Stable transformation was observed in 40% of the *S. henryi* explants (n=25), and in 58% of the *A. fimbriata* explants (n=59). Additionally, transgenic calli, shoot, and root primordia (*Figure 2-7G, H*) were successfully regenerated from stem explants of *A. fimbriata* tissue cultured plants. The integration of *EGFP* in the transgenic calli was confirmed by genomic PCR analysis (*Figure 2-7I*). The PCR reactions including DNA from green fluorescent calli (*Figure 2-7I*, lanes 3-5) resulted in the amplification of one 426bp fragment similar to the control reaction including plasmid pGH00.0131 DNA (*Figure 2-7I*, lane 7). Amplification was not detected in the control reaction containing DNA from non-transgenic *Aristolochia* leaf tissue (*Figure 2-7I*, lane
2. and the control reaction without DNA (Figure 2-71, lane 6). Additional experiments have been initiated for further optimization of the transformation system.

---

**Figure 2-71**

A. Image A

B. Image B

C. Image C

D. Image D

E. Image E

F. Image F

G. Image G

H. Image H

I. Image I

J. Image J
Discussion

*Aristolochia fimbriata* has many characteristics of a valuable experimental model

A basal angiosperm experimental model is needed to explore the diversity of basal angiosperms and to test hypotheses about the evolution of development in flowering plants. *Aristolochia fimbriata* possesses many features desired in a model system. *A. fimbriata* has the physical features for large-scale greenhouse cultivation. It can be made to flower continuously, and provides the essential feature of self-compatibility, which permits the production of large numbers of homozygous individuals required for gene functional analysis in a single life cycle. We have begun to develop inbred lines to facilitate genome mapping and large scale mutagenesis experiments. High-efficiency transformation with a GFP construct allows rapid, nondestructive identification of transformed tissues for subsequent processing, and our *in vitro* micropropagation and regeneration methods (*Figure 2-7J*) can yield greenhouse acclimated plants in three months (Bliss et al., submitted, December 2008). Individual *A. fimbriata* plants survive indefinitely, providing ongoing access to mutant lines that can be cloned for distribution.
to the research community. Currently available for analysis, the VL and NV genotypes possess a number of readily discernible traits including leaf variegation (Figure 2S-3C) and perianth details (Figure 2S-3B, D, E) amenable to investigation at the genetic level. Further inbreeding and crossbreeding these genotypes can be used to dissect the genetic basis for these or other differences, which may yield useful markers, identify linked genes, and ultimately contribute to mapping and assembling the *Aristolochia fimbriata* genome sequence.

*Aristolochia fimbriata* is well positioned for studies of the evolution of development

*Aristolochia fimbriata* is in a strong phylogenetic position to support its use as a model system in evolutionary studies. Aristolochiaceae, with its approximately 550 species in four genera (Wanke et al., 2006a), is one of the most diverse and speciose families among the basal angiosperms. The largest genus in the family, *Aristolochia*, contains approximately 450 species, and has long been of interest to botanists due to its monosymmetric, unipartite fly-trapping perianth (Figure 2S-3E) and unique gynostemium, which is a structure formed by the fusion of the gynoecium with the anthers. *Aristolochia fimbriata* is a typical member of the family, such that its flower was modelled in glass by Leopold and Rudolf Blaschka in the latter half of the nineteenth century (Glass flowers from the Ware Collection in the Botanical Museum of Harvard University, 1940). *A. fimbriata* differs from closely related *Aristolochia* species in several late stage modifications (e.g., fimbriae, papillae, pubescence), as well as in perianth and gynostemium development (Gonzalez and Stevenson, 2000b, a); these features are
potentially suitable for molecular analysis and studies of gene function using an *A. fimbriata* experimental system.

Comparisons between closely related taxa in Aristolochiaceae could support discovery of differences responsible for speciation and may resolve long-standing questions about the origins of floral structures. Unlike *Aristolochia*, the three other genera in the family (*Asarum* (Figure 2-4A), *Saruma* (Figure 2-4B), and *Thottea* (Figure 2-4C)) have radially symmetric perianths, and account for approximately 100 species (Table 2S-1). The bilaterally symmetric perianth of *Aristolochia* represents the most “basal” occurrence of this important floral adaptation in angiosperms. *Aristolochia fimbriata* presents the opportunity to investigate the genetic basis of bilateral symmetry in magnoliids and compare it with that found in the eudicot models *Antirrhinum* (Cubas et al., 2001) and *Lotus* (Feng et al., 2006). Existing microscopic studies of anatomy and development in Aristolochiaceae will facilitate comparative and functional studies, and provide insight into the evolution of development in this family. For *Saruma* the anatomy of stem, leaf, flower, and pollen have been described (Dickison, 1992, 1996), and across the family, ovule and seed development (Gonzales and Rudall, 2003), microsporogenesis (Gonzalez et al., 2001), and inflorescence morphology (Gonzalez, 1999) have been described in detail.

Morphological and gene expression studies argue that Aristolochiaceae offers an excellent system in which to study the role of homologs of B-class MADS-box genes which are required for the organ identities of petal and stamen in higher eudicots and putative homologous organs in grasses (Ma and dePamphilis, 2000; Irish, 2003). The flower of the monotypic genus *Saruma* (Figure 2-4B) resembles that of the typical
magnoliid flower more than any other species in the family, having apparent sepals, petals, stamens, and carpels. A thorough examination of development in the petaloid perianth whorl of *Saruma* indicates it is of staminoid origin (de Craene et al., 2003). In *Aristolochia*, the utricle, tube, and limb (Figure 2S-3E) of the single-whorled perianth originate from the outermost whorl and are interpreted as a calyx (Gonzalez and Stevenson, 2000b). However, Jaramillo and Kramer (2004) reported that putative orthologs of B-class genes (*AP3, PI*) are expressed in the stamens and staminoid “petals” of *Saruma*, as well as in the anthers and inner layer of utricle cells of the *Aristolochia* sepalloid perianth. In *Asarum* and *Thottea*, the appendages opposite the sepals have been interpreted to be petaloid (Leins and Erbar, 1985; Renuka and Swarupanandan, 1986; Sugawara, 1987b; Leins et al., 1988; Gonzalez and Stevenson, 2000a, b), and putative *AP3* orthologs have been found in *Asarum* (Kramer and Irish, 2000; Zhao et al., 2006).

Variations in the expression of putative orthologs of the B-class genes in *Saruma* and *Aristolochia* suggest they might not regulate perianth floral organ identity in the same way as they do in eudicots and grasses, even though expression is found in pollinator-attracting structures. For example, homologs of B-class genes might regulate the identity of the perianth in Aristolochiaceae, even though the perianth does not have the same structure (i.e., petals) found in other taxa. Alternatively, these genes might not regulate the perianth identity in Aristolochiaceae. Experimental evidence from Aristolochiaceae is needed to determine if B-class homologs play a role in organ identity in basal angiosperms, particularly since the expression patterns of these genes in the first perianth whorl of other basal taxa is variable (Kramer and Irish, 1999, 2000; Zahn et al., 2005b).
*Aristolochia* contains highly developed biochemical pathways offering insight into evolution of biochemical synthesis and coevolution with insects

Aristolochiaceae produce a complex mixture of secondary metabolites, as is common in basal angiosperms,. In particular, aristolochic acids and aristolactams are produced in Aristolochiaceae (reviewed in Kumar et al. 2003) found throughout Piperales and the basal eudicots. Compounds produced by alkaloid biosynthesis pathways in the poppy family (Papaveraceae, Ranunculales) are of great pharmacological importance. Compounds from a parallel pathway in Aristolochiaceae (aristolochic acids and aristolactams) are associated with the use of some *Aristolochia* species in traditional medicines (Reddy et al. 1995). The common name “birthwort” attributed to *Aristolochia* refers to traditional use of some species, particularly extracts from root tissues, as abortifacients, emmenagogues, or post-coital antifertility agents (Pakrashi and Pakrasi, 1979). More recently, constituents of primarily root extracts from *Aristolochia* species have been isolated and evaluated for biological activity as antibiotics, antivenoms and tumor-inhibiting agents (Gupta et al., 1996; Broussalis et al., 1999; Otero et al., 2000; Qiu et al., 2000; Gadhi et al., 2001a; Shafi et al., 2002; Elizabeth and Raju, 2006; Meinl et al., 2006), although aristolochic acids become bioactivated and carcinogenic when consumed in the diet (Hwang et al., 2006; Grollman et al., 2007).

In addition to its pharmacological properties, *Aristolochia* provides an opportunity to explore coevolution of secondary metabolites with insects. Dipterans commonly serve as pollinators in *Aristolochia* (Petch, 1924; Lindner, 1928; Lu, 1982; Wolda and Sabrosky, 1986; Hall and Brown, 1993), sometimes having specialized, mutualistic relationships involving egg deposition in the flowers (Vogel, 1978; Disney and Sakai,
Dipteran pollinators are thought to be attracted to secondary metabolites mimicking the aroma of a food source (Petch, 1924; Burgess et al., 2004), or acting as pheromones to attract species-specific, sex-specific pollinators (Rulik et al.), or stimulate oviposition (Sakai, 2002).

*Aristolochia* species are also important host plants for the larval stages of swallowtail butterflies (Papilionidae, Lepidoptera) (Rausher, 1981; Nishida and Fukami, 1989; De Morais and Brown, 1991; Nishida, 1994; Klitzke and Brown, 2000; Fordyce et al., 2005). Secondary metabolites found in particular *Aristolochia* species are critical for the defense and survival of associated species of swallowtail butterfly species during their feeding stage, to the extent that the decline of butterfly populations is attributed to decreased distributions of particular *Aristolochia* species (Sands et al., 1997). Finally, secondary metabolites of *Aristolochia* and related species are of interest for their repellent, insecticidal, and antifeedant activities (Lajide et al., 1993; Alexenizer and Dorn, 2007; Palacios et al., 2007). Biochemistry in *Aristolochia* can be evaluated in greenhouse grown or micropropagated plants, using chemical analyses (e.g., mass spectroscopy, gas chromatography) to further characterize constituents of specific plant tissues in *A. fimbriata*. Biochemical pathways can be further investigated using transformed roots (Figure 2-7G, H) or callus currently available in the transformation-regeneration system, without further optimization.
**Aristolochia** can provide insight into development of woodiness

Woodiness is an important seed plant feature, both for commercial and ecological purposes (Stevens, 2008). Growth forms within the Aristolochiaceae vary widely, presenting an opportunity to investigate growth form traits including flexibility, stiffness, and woodiness of closely related species (Speck et al., 2003). Aristolochiaceae are most commonly perennial, self-supporting herbs (Figure 2-4A, B), procumbent or trailing, non-self-supporting vines (*A. passiflorafolia, A. fimbriata*) (Figure 2-4G, H), and woody lianas (Figure 2-4F). Rarely, they are small woody shrubs (Figure 2-4C); and even more rarely trees (or tree-like forms) (Figure 2-4D). Ancestral members of the family (*Asarum* and *Saruma*) are characterized as perennial rhizomatous herbs, and the sister group to *Aristolochia* is comprised of the woody sub-shrub *Thottea* (Figure 2-4C). Perennial herbs appeared iteratively in the topology (Figure 2-6) of the family phylogeny and are nested within groups of woody vines. Most eudicots and monocots are modular organisms with indeterminate body plans. Shifts in the ontogenetic trajectory may be expected to have a profound effect on the overall size and potential life history of the descendant. This effect might have played a key role in growth form evolution and the development of flexibility, stiffness, and woodiness in Aristolochiaceae (Lahaye et al., 2005; Rowe and Speck, 2005). Differences in these growth form traits can be investigated in *Aristolochia* species of interest, including *A. fimbriata*, beginning with cellular level observations and descriptions of development of secondary growth (“wood”). Molecular and cell biological methods can be used to locate and describe in *Aristolochia* homologs of genes involved in growth form traits in other species (e.g.,
Arabidopsis, Populus, Liriodendron), to further characterize the role of interesting gene products in Aristolochiaceae.

*Aristolochia* can reveal features of the ancestor common to monocots and eudicots

Aristolochiaceae occurs in Piperales, in the magnoliid clade, which is sister to the very large and diverse monocot plus eudicot clades (Jansen et al., 2007). As such, *Aristolochia* can provide a close outgroup for analysis of ancestral traits in the major groups of angiosperms. The ancient features of different lineages in Piperales have long been recognized, earning them classification as “paleoherbs” in early taxonomic works (Taylor and Hickey, 1992). Aristolochiaceae and close relatives show a mixture of features of monocots and eudicots. The more ancestral clade, Asaroideae, is comprised of rhizomatous perennials similar to basal monocots and eudicots. Also found in Piperales are aquatics (*e.g.*, Saururaceae) a common life form in both basal eudicots and monocots (*e.g.*, Acorales, Alismatales). Piperales, and Aristolochiaceae in particular displays other features more commonly associated with monocots, including floral organs in multiples of three, median prophylls (which are shared with nearly all monocots and only few dicot clades), and subtype PII sieve-tube plastids (Huber, 1993; Behnke, 2003). Piperales also share with early diverging monocots (Alismatales, Arales) distichous placement of leaves and palmate leaf venation. Indeed, in some phylogenies, Piperales appear as the closest relative to *Acorus*, the sister of all other monocots (Duvall, 2000). Consequently, many features in monocots and eudicots, both genetic and phenotypic, can be expected to have a homolog in Aristolochiaceae and
its relatives, and would help to characterize the extinct ancestors of the eudicot, monocot and magnoliid clades.

Growing genomic resources in Aristolochiaceae support further development of model system

Genomic resources are growing rapidly for Aristolochiaceae, which will facilitate the identification and study of genes, gene families, and functional studies in basal angiosperms. The Ancestral Angiosperm Genome Project (www.ancestralangiosperm.org) has selected *Aristolochia fimbriata* for deep EST sequencing using a combination of traditional capillary (Albert et al., 2005) and extensive next generation (454, Solexa Illumina) sequencing of libraries constructed from multiple vegetative and reproductive tissues and stages. There are currently 16,451 Sanger EST sequences from *A. fimbriata* in the NCBI EST database (http://www.ncbi.nlm.nih.gov/sites/entrez). In this initial set of *A. fimbriata* ESTs, we used BLAST analyses as described in Albert et al. 2005 to identify putative homologs of many interesting regulatory and signaling genes (see Table 2-1 for examples). Surprisingly, *Aristolochia* cDNA sequences had greater sequence similarity with other monocot or eudicot species than with *Arabidopsis*, highlighting the important role *A. fimbriata* will play in rooting phylogenetic analyses of gene function [results not shown]. Consistent with the composition of tissues included in the libraries, we found putative orthologs of many genes important for development such as auxin efflux carrier PIN1, phytochrome signaling protein GIGANTEA, as well as floral development genes AP2, AP3, AINTEGUMENTA, and SEP3 (Table 2-1). We anticipate these cDNA sequences,
and many more being currently generated, will contribute to new and ongoing studies of
evolution of development in our lab and in others.

Table 2-1: Homologs of genes involved in development, cell wall biosynthesis, and
response to biotic and abiotic stress revealed in preliminary sequencing of two cDNA
libraries constructed from combined *A. fimbriata* tissues. All genes searched (blastx)
against the sequenced plant genomes of *Arabidopsis thaliana*, *Oryza sativa*, *Populus
trichocarpa*, *Vitis vinifera*, *Carica papaya*, *Medicago truncatula*, and *Sorghum bicolor*
(http://www.floralgenome.org/tribedb/index.pl). Values reported are the number of ESTs
in each unigene, the unigene length (bp), percent identity of the unigene to its best blastx
hit, evale, *Arabidopsis AGI*, and the annotation of the best *Arabidopsis* hit as well as the
species with the best overall hit following (in parentheses).

<table>
<thead>
<tr>
<th>ESTs</th>
<th>Length</th>
<th>Identity</th>
<th>Evalue</th>
<th>AGI</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Development</strong></td>
</tr>
<tr>
<td>1</td>
<td>330 bp</td>
<td>82%</td>
<td>6e-49</td>
<td>AT4G36920</td>
<td>AP2/EREBP floral homeotic protein (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>9</td>
<td>755 bp</td>
<td>71%</td>
<td>1e-18</td>
<td>AT1G22770</td>
<td>GIGANTEA (<em>Sorghum bicolor</em>)</td>
</tr>
<tr>
<td>3</td>
<td>1254 bp</td>
<td>84%</td>
<td>2e-73</td>
<td>AT1G73590</td>
<td>PIN1 auxin efflux carrier (<em>Medicago truncatula</em>)</td>
</tr>
<tr>
<td>2</td>
<td>613 bp</td>
<td>88%</td>
<td>3e-44</td>
<td>AT4G37750</td>
<td>AINTEGUMENTA (<em>Carica papaya</em>)</td>
</tr>
<tr>
<td>7</td>
<td>1420 bp</td>
<td>50%</td>
<td>1e-79</td>
<td>AT2G37630</td>
<td>Homologous to Antirrhinum PHANTASTICA (PHAN), maize ROUGH SHEATH2 (RS2) (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>4</td>
<td>747 bp</td>
<td>61%</td>
<td>2e-41</td>
<td>AT1G24260</td>
<td>RNA Slicer that selectively recruits microRNAs and siRNAs (<em>Populus trichocarpa</em>)</td>
</tr>
<tr>
<td>1</td>
<td>736 bp</td>
<td>87%</td>
<td>1e-112</td>
<td>AT1G48410</td>
<td>PHABULOSA contains hd-leucine zipper domains and domain similar to a mammalian sterol binding domain (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>2</td>
<td>956 bp</td>
<td>81%</td>
<td>1e-151</td>
<td>AT2G34710</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>835 bp</td>
<td>33%</td>
<td>6e-40</td>
<td>AT3G54340</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Cell wall biosynthesis</strong></td>
</tr>
<tr>
<td>1</td>
<td>495 bp</td>
<td>56%</td>
<td>6e-14</td>
<td>AT5G42100</td>
<td>Plasmodesmal (Pd)-associated membrane protein involved in Pd callose degradation and gating (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>8</td>
<td>787 bp</td>
<td>55%</td>
<td>4e-28</td>
<td>AT2G38540</td>
<td>Cell wall lipid transfer protein binds calmodulin in a Ca2+-independent manner (<em>Medicago truncatula</em>)</td>
</tr>
<tr>
<td>5</td>
<td>932 bp</td>
<td>70%</td>
<td>2e-65</td>
<td>AT5G08380</td>
<td>ATAGAL1 alpha-galactosidase similar to ATAGAL2 (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>1</td>
<td>737 bp</td>
<td>53%</td>
<td>4e-39</td>
<td>AT4G03210</td>
<td>Xyloglucan endotransglucosylase/hydrolase (XTH) (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>15</td>
<td>986 bp</td>
<td>81%</td>
<td>6e-80</td>
<td>AT1G26770</td>
<td>EXP10: expansin involved in the formation of nematode-induced syncytia in roots (<em>Carica papaya</em>)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Stress response</strong></td>
</tr>
<tr>
<td>3</td>
<td>1725 bp</td>
<td>65%</td>
<td>1e-41</td>
<td>AT3G12500</td>
<td>BASIC CHITINASE in ethylene/jasmonic acid mediated signalling pathway during SAR (<em>Oryza sativa</em>)</td>
</tr>
<tr>
<td>2</td>
<td>972 bp</td>
<td>82%</td>
<td>1e-145</td>
<td>AT5G67030</td>
<td>Zeaxanthin epoxidase gene (<em>Populus trichocarpa</em>)</td>
</tr>
<tr>
<td>3</td>
<td>1372 bp</td>
<td>79%</td>
<td>1e-164</td>
<td>AT3G45640</td>
<td>MAP KINASE 3 (MPK3) upregulated in response to touch, cold, salinity, chitin (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>2</td>
<td>572 bp</td>
<td>71%</td>
<td>1e-59</td>
<td>AT1G05850</td>
<td>POM-POM1; Chitinase-like protein essential for tolerance to heat, salt, drought stresses (<em>Vitis vinifera</em>)</td>
</tr>
<tr>
<td>1</td>
<td>685 bp</td>
<td>55%</td>
<td>2e-45</td>
<td>AT1G59870</td>
<td>ATP binding cassette transporter contributes to SA independent nonhost resistance and heavy metal resistance (<em>Populus trichocarpa</em>)</td>
</tr>
<tr>
<td>6</td>
<td>904 bp</td>
<td>52%</td>
<td>4e-46</td>
<td>AT1G06040</td>
<td>Salt tolerance protein (STO) (<em>Populus trichocarpa</em>)</td>
</tr>
<tr>
<td>1</td>
<td>763 bp</td>
<td>61%</td>
<td>2e-54</td>
<td>AT5G51700</td>
<td>RAR1 disease resistance protein ; Required for R protein accumulation (<em>Carica papaya</em>)</td>
</tr>
</tbody>
</table>
Comparative analyses of cDNA sequences within Aristolochiaceae will be facilitated by 10,274 EST sequences from *Saruma henryi* (Figure 2-4B). *Saruma* was selected by the Floral Genome Project (FGP) to represent Aristolochiaceae for floral transcriptome sequencing ([http://www.floralgenome.org/taxa/](http://www.floralgenome.org/taxa/)) because it appears to display ancestral morphological characters in the family (Dickison, 1992; Leins and Erbar, 1995; Kelly, 1998) and because its flower includes all four (sepal, petal, stamen, and carpel) floral whorls. The ESTs from the *Saruma* premeiotic flower bud cDNA library (Albert et al., 2005) and from *A. fimbriata* form the foundation of the genomic resources available today from this important basal angiosperm family.

In addition to EST sequencing, our group has isolated miRNAs from *A. fimbriata* using cloning and capillary sequencing as well as ultrahigh throughput pyrosequencing (Ali Barakat, P. Kerr Wall, and Claude dePamphilis, in prep.). This study allowed the isolation of hundreds of different miRNA sequences belonging to 32 conserved families as well as several non-conserved families. MicroRNAs are small RNAs (21 nt) that play a major role as regulator of gene expression in various physiological, cellular, but, mainly developmental processes (Jones-Rhoades et al., 2006). Several potential miRNA targets were found in cDNA sequences of *A. fimbriata*. Predicted target genes include transcription factors but also genes implicated in various metabolic processes and in stress defense. The isolation of miRNA from *Aristolochia* presents a good opportunity for analyzing the function of miRNAs in basal angiosperms as well as understanding how miRNA-mediated regulation of gene expression has evolved in land plants by comparing miRNAs in basal angiosperms to those in basal eudicots and lower land plants (Axtell et
al., 2007; Barakat et al., 2007). This core of genetic, genomic, and methodological resources is presently available as a foundation for further development of *Aristolochia fimbriata* as a basal angiosperm model system as well as for immediate use by the scientific community working on various areas of research including evolution of development, plant resistance to biotic and abiotic stresses, gene functional analysis, and comparative genomics.

**Conclusion**

We have used a rigorous process to select and develop resources for the basal angiosperm, *A. fimbriata*. Culturing and hand pollination methods required for rapid generation of homozygous lines needed for genetic experiments are described. The small genome size and immediate availability of sequence data supports ongoing studies of molecular genetics and evolution. Hypotheses about gene evolution and gene function can be tested using a reverse genetic approach, i.e., over and under expression studies in a transformable species suitable for large-scale cultivation. Optimizing the selection phase for transformed *A. fimbriata* explants will yield a high throughput transformation system for investigating evolution of gene function in a basal angiosperm. Along with continued development of genomic resources, *A. fimbriata* promises to be an excellent experimental system to provide further insight into the diversity found among basal angiosperms and to test hypotheses about the evolution of angiosperm genomes, development and biochemistry.
Materials & Methods

Cultivation

We evaluated 24 species of Aristolochiaceae. These were selected to encompass the phenotypic plasticity of the four-whorled, actinomorphic, *Saruma*, the actinomorphic, single-whorled perianths of *Asarum* and *Thottea*, and the monosymmetric, highly modified and diverse flowers of *Aristolochia*. The sampled taxa also reflected the genetic diversity of the whole family recovered by phylogenetic analysis. Plant material was obtained from commercial nurseries, private donations, and academic sources. Vouchers of specimens included in the phylogenetic analysis and sampled for genome sizes have been entered into herbaria as described in Table 2S-5. For these species, we evaluated evolution of genome size and chromosome number in a phylogenetic context. For 14 species of *Aristolochia* we evaluated life cycle and cultivation characteristics. Plants were maintained in the greenhouse at The Pennsylvania State University, University Park, PA. All seeds were germinated in soil-free potting medium (Pro-Mix BX, Premier Horticulture Inc., Quakertown, PA) in shallow germination trays with drainage holes, in the greenhouse at 18-27°C (varying from night to day) and 40-70% humidity. The trays were incubated on heating mats operating at approximately 27°C, as needed. Natural day length was supplemented with high-pressure sodium lamps (1000 watt) October through April to provide twelve-hour days. Plants received regular watering as needed. Depending on the stage of growth, regular fertilizer applications were provided, as a drench, alternating Peter's Professional 15-16-17 Peat Lite Special at 200 PPM nitrogen
(once to twice weekly) with Peters Professional 21-7-7 Acid Special (Scotts Horticulture, Marysville, OH) at 200 PPM nitrogen (approximately every six weeks). The plants were drenched once a month with 100 ppm chelated iron (Sprint 330 10% iron, RoseCare.com, Santa Barbara, CA).

**Genome sizing**

Nuclear genome size estimation were obtained by flow cell cytometry following the protocol described by Arumuganathan and Earle (1991). The mean nuclear DNA content of each plant sample (expressed as pg) was based on 1000 scanned nuclei from sample tissue, compared to a preparation of tissue from the internal standard. Each nuclear preparation was sampled 4 times.

**Phylogenetic analysis**

To clarify the phylogenetic positions of the taxa surveyed and to evaluate genome size in an evolutionary context we constructed a phylogenetic tree based on the plastid *trnK/matK* region. Total DNA was extracted using the CTAB method of (McNeal et al., 2006). Vouchers, DNA, and tissue samples are stored at Dresden, PAC or DD. Amplification and sequencing was performed following methods described in detail by Wanke et al. (2007) using published primers (Wanke et al., 2006a; Wanke et al., 2006b; Wanke et al., 2007). Sequences were manually aligned using PhyDE® (Müller et al., 2007). The alignment is available from TreeBASE (www.treebase.org). Phylogenetic
analysis was performed under maximum parsimony in PAUP* 4.0b10 (Swofford, 2002) using PAUP scripts written by PRAP (Müller, 2004). PRAP was used to implement the Parsimony Ratchet (Nixon, 1999) following procedures described in Wanke et al. (2007) but with 1000 ratchet replicates. Bootstrap values were additionally calculated to infer branch support with 1000 replicates. For an independent evaluation of relationships, a likelihood approach was chosen using the likelihood ratchet described by Morrison (2007) as implemented in PRAP v. 2.0 (Müller, 2004) with default settings.

**Pollination experiments**

Several species of *Aristolochia* had been reported to be self-compatible (*A. fimbriata, A. elegans, A. ridicula, A. ringens*) and generally protogynous, having a receptive stigma before the anthers dehisce. To determine if autogamous pollination could be successful, we attempted hand pollinations of *A. elegans* and *A. fimbriata* on day one (the day of anthesis), day two, day three, day four and day five. Unopened flower buds were covered with pollination bags prior to anthesis and observed daily. Pollinations were accomplished by severing the perianth midway across the utricle, just above the gynostemium (Figure 2S-1 (A)). Pollen was transferred with a toothpick (Figure 2S-1 (B)), and the remnant of the perianth was taped closed to prevent additional pollinator entry. Mature fruits were collected and seeds germinated on wet toweling at 27°C exposed to 16/8 hour day/night cycles. Hand pollination methods are detailed in the Pollination Protocol.
**Genetic transformation**

Protocol for genetic transformation of *A. fimbriata* was developed based on *in vitro* shoot regeneration from leaf and stem explants coupled with *Agrobacterium*-mediated transformation. Leaf and Internodal stem segments (2-3 cm long) from rooted tissue cultured *A. fimbriata* plants were excised and immediately immersed in induction media (Petch, 1924; Maximova et al., 2003; Bliss et al., submitted, December 2008) to keep moist. The explants were inoculated with *Agrobacterium* strain AGL containing plasmid pGH00.0131 as previously described for *Theobroma cacao* L. (Maximova et al., 2003). After the inoculation the explants were blotted on sterile paper towels and co-cultivated on callus initiation medium (SI) (0.5mg/L 6BA, 1mg/L NAA and 1mg/L TDZ) for 64 hours in the dark at 27°C. Following co-cultivation, the explants were transferred to CI supplemented with 50 mg/L Geneticin (G418) (Cellgro, Herndon, VA) and 200 mg/L Claforan (Aventis, New Jersey) and incubated at 27°C in the dark for the remainder of the 14 days. After culture on CI medium, the explants were transferred on shoot regeneration medium (1.75 mg/L 6BA and 1.0 mg/L NAA) with 25 mg/L G418 and 200mg/L Claforan and incubated in the dark for an additional 14 days. After the 28 days, the cultures were incubated under dim light until the development of shoot primordia. Individual glowing shoot primordia were then excised and transferred and maintained for further shoot elongation and rooting on REN2 medium (hormone free) in culture vessels (Sweetheart DSD8X and LDS58) under dim light conditions at 27°C (Bliss et al., submitted, December 2008). EGFP fluorescence was observed and recorded as previously described (Maximova et al., 2003).
Genomic PCR analysis

The incorporation of the transgenes was confirmed by genomic PCR. A set of EGFP specific PCR primers were used for the analysis. The primers amplify a 426bp EGFP fragment (5’-CCA GGA GCG CAC CAT CTT CT-3’ and 5’-CTC GTC CAT GCC GAG AGT GA-3’ (Maximova et al., 2003). Genomic DNA was isolated from non-transgenic Aristolochia leaf tissue and from transgenic calli from independent lines by using a CTAB method of (McNeal et al., 2006). Each PCR reaction (final volume of 20 µl) contained: 5ng/ul DNA (Qiagen purified DNA kit #69104), 10ul JumpStart™ REDTaq® ReadyMix (Sigma #0982), 5uls water, forward and reverse primers at final concentration 0.5 µM. Reactions were prepared on ice. Control PCR reactions were also performed with 1ng/ul plasmid DNA from vector pGH00.0131 and PCR reaction mix without DNA. This represents an equal molar amount of plasmid DNA compared to the GFP DNA contained in 5ng total Aristolochia genomic DNA present in the leaf extract, assuming single copy/insertion of the GFP gene. For the plasmid reactions DNA was isolated using QIAGEN plasmid midi purification kit (QIAGEN Inc., Valencia, CA). PCR conditions for all reactions were: 94ºC for 2 min, then 35 cycles of 94ºC for 45 sec., 62ºC for 45 sec, 72ºC for 1 min. The final cycle was followed by incubation at 72ºC for 7 min. 5 µl of each PCR reaction were loaded onto 1.5% high-resolution agarose gel (Sigma-Aldrich Co., St. Louis, MO, #A-4718) for electrophoresis.
Genetic transformation

Protocol for genetic transformation of *A. fimbriata* was developed based on *in vitro* shoot regeneration from leaf and stem explants coupled with *Agrobacterium*-mediated transformation. Leaf and Internodal stem segments (2-3 cm long) from rooted tissue cultured *A. fimbriata* plants were excised and immediately immersed in induction media (Maximova et al., 2003; Bliss et al., submitted, December 2008) to keep moist. The explants were inoculated with *Agrobacterium* strain AGL containing plasmid pGH00.0131 as previously described for *Theobroma cacao* L. (Maximova et al., 2003). After the inoculation the explants were blotted on sterile paper towels and co-cultivated on callus initiation medium (CI)

Supplemental Material

Table 2S-1: Basal angiosperm characteristics

Table 2S-2: Cultivation features of 25 species of Aristolochiaceae

Cultivation Supplement: Descriptions of perianth maturation, hand pollination methods, and self-compatibility in *A. fimbriata* and *A. elegans*. Includes: Table 2S-3, Table 2S-4, Figure 2S-1, Figure 2S-2

Table 2S-5: Genome sizes, vouchers, sources, and accessions

Figure 2S-3: *Aristolochia fimbriata* genotype and perianth detail

Figure 2S-4: Phylogram of Aristolochiaceae relationships
Acknowledgements

We thank A. Omeis for plant care, Jason Stetson and Yunjiao Joy Wang for DNA isolations; Paula Ralph for plant care and observations, pollination, seed germination, lab assistance and tissue collection; Yi Hu for tissue collection, RNA isolations, and lab assistance; Joel McNeal, Mario Blanco, Al Hill, Larry Rosen, Russ Strover, Victor Wong, the New York Botanical Garden, the Botanical Gardens at the University of Ulm, Dresden, Universidade de Coimbra, and Dawe’s Arboretum for seed and plant material. This research was supported by The Floral Genome Project [], The Ancestral Angiosperm Genome Project [NSF DBI-0638595], the Biology Department of Penn State University, and a postdoctoral fellowship by DAAD for Stefan Wanke is gratefully acknowledged.
Table 2S-1: A summary of relevant basal angiosperm characteristics. Familiar genera include representatives from the family and are not intended to comprise a comprehensive listing. Taxa are considered “commercially available” if one hundred plants or more can be purchased, inexpensively, and can be readily propagated from seed.

1(Stevens, 2008; Bliss et al., submitted, December 2008)
2(Mabberley, 2008)
3Number of species shown for Aristolochiaceae genera from (Wanke et al., 2006a)

<table>
<thead>
<tr>
<th>Order</th>
<th>Family1,2</th>
<th># Genera /species3</th>
<th>Familiar genera3</th>
<th>Growth form1,2</th>
<th>Commercially available</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMBORELLALES</td>
<td>Amborellaceae</td>
<td>1/1</td>
<td>Amborella</td>
<td>shrub, tree</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Nymphaeaceae</td>
<td>5/95</td>
<td>Nymphaea Nuphar</td>
<td>aquatic herb</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Barclaya Victoria</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Euryale</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cabombaceae</td>
<td>2/6</td>
<td>Cabomba Brassenia</td>
<td>aquatic herb</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYMPHAEALES</td>
<td>Hydatellaceae</td>
<td>2/10</td>
<td>Hydatella Trithuria</td>
<td>aquatic herb</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Austrobaileyaceae</td>
<td>1/2</td>
<td>Austrobailey</td>
<td>shrub, tree</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Illiciaceae</td>
<td>3/92</td>
<td>Illicium Schisandra</td>
<td>shrub, tree</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kadsura</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trimeniaceae</td>
<td>1/8</td>
<td>Trimenia</td>
<td>shrub, tree, liana</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERATOPHYLLALES</td>
<td>Ceratophyllaceae</td>
<td>1/2</td>
<td>Ceratophyllum</td>
<td>aquatic herb</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHLORANTHALES</td>
<td>Chloranthaceae</td>
<td>4/75</td>
<td>Chloranthus Ascarina</td>
<td>herb, shrub, tree</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hedyosmum Sarcandra</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Annonaceae</td>
<td>123/2100</td>
<td>Annona Guatteris</td>
<td>shrub, tree, liana</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>Eupomatiaceae</td>
<td>1/3</td>
<td>Eupomatia</td>
<td>shrub, tree</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>Magnoliaceae</td>
<td>2/221</td>
<td>Magnolia Liriodendron</td>
<td>shrub, tree</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>Degeneriaceae</td>
<td>1/2</td>
<td>Degeneria</td>
<td>tree</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Himantandraceae</td>
<td>1/2</td>
<td>Galbulimima</td>
<td>tree</td>
<td>no</td>
</tr>
<tr>
<td>MAGNOLIALES</td>
<td>Myristicaceae</td>
<td>19/520</td>
<td>Myristica Horsfieldia</td>
<td>shrub, tree,</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Virola Knema</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calycanthaceae</td>
<td>3/10</td>
<td>Chimonanthus Calycanthus</td>
<td>shrub, tree</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Idiospermum</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hernandiaceae</td>
<td>4/60</td>
<td>Hernandia Illigera</td>
<td>shrub, tree, liana</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Lauraceae</td>
<td>52/2550</td>
<td>Laurus Liise</td>
<td>shrub, tree, parasitic vine</td>
<td>yes</td>
</tr>
<tr>
<td>Monimiaceae</td>
<td>24/200</td>
<td>Ocotea, Cinnamomum, Cryptocarya, Persea, Lindera, Cassytha, Nectandra, Phoebe, Apollonias, Beilschmiedia, Umbellularia</td>
<td>shrub, liana</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>---------------------------------------------------------------</td>
<td>----------------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Siparunaceae</td>
<td>2/57</td>
<td>Doryphora, Peumus, Ymalos, Mollinedia, Tambourissa, Kihara</td>
<td>shrub, tree</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Gomortegaceae</td>
<td>1/1</td>
<td>Siparuna, Glossocalyx</td>
<td>shrub, tree</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Atherosperma-taceae</td>
<td>7/25</td>
<td>Atherosperma, Daphnandra, Doryphora, Dryadodaphne, Laurelia, Nemuaron</td>
<td>shrub, tree</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Canellaceae</td>
<td>5/13</td>
<td>Canella, Cinnamodendron, Cinnamosma</td>
<td>shrub, tree</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Winteraceae</td>
<td>4/65</td>
<td>Drimys, Zygogynum, Pseudowintera, Takhtajania</td>
<td>shrub, tree</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Hydnoraceae</td>
<td>2/15</td>
<td>Prosopanche, Hydnora</td>
<td>parasitic herb</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Piperaceae</td>
<td>6/2750</td>
<td>Peperomia, Piper, Zippelia, Manekia, Verhuelli</td>
<td>herb</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Saururaceae</td>
<td>4/6</td>
<td>Anemopsis, Houttuynia</td>
<td>herb</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Lactoridaceae</td>
<td>1/1</td>
<td>Lactoris</td>
<td>shrub</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Aristolochiaceae</td>
<td>4/550</td>
<td>Asarum (1 sp.), Thottea (ca. 29 spp.), Aristolochia (ca. 450 spp.)</td>
<td>herb, shrub, liana</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>
Table 2S-2: Cultivation features of 25 Aristolochiaceae taxa evaluated, with findings. Observations based on greenhouse cultivation (at Penn State and/or Botanical Garden, Dresden) except where indicated.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Floral productivity</th>
<th>Maintenance</th>
<th>Vernalization required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asaroideae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Saruma henryii</em> Oliv.</td>
<td>moderate&lt;sup&gt;1&lt;/sup&gt;</td>
<td>low&lt;sup&gt;2,7&lt;/sup&gt;</td>
<td>required</td>
</tr>
<tr>
<td><em>Asarum canadense</em> L.</td>
<td>moderate&lt;sup&gt;1&lt;/sup&gt;</td>
<td>low&lt;sup&gt;2,7&lt;/sup&gt;</td>
<td>required</td>
</tr>
<tr>
<td><strong>Aristolochioideae</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>genus <em>Aristolochia</em> L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>subgenus <em>Isotrema</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. serpentaria</em> L.</td>
<td>ample</td>
<td>low&lt;sup&gt;2,4,7&lt;/sup&gt;</td>
<td>required</td>
</tr>
<tr>
<td><em>A. macrophylla</em> Lam.</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>low&lt;sup&gt;4,7&lt;/sup&gt;</td>
<td>required</td>
</tr>
<tr>
<td><em>A. californica</em> Torr.</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>low&lt;sup&gt;4,7&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. tomentosa</em> Sims</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>low&lt;sup&gt;4,7&lt;/sup&gt;</td>
<td>required</td>
</tr>
<tr>
<td><em>A. holostylis</em> (Duchartre) F. Gonzalez&lt;sup&gt;7&lt;/sup&gt;</td>
<td>sparse</td>
<td>low</td>
<td>no</td>
</tr>
<tr>
<td>subgenus <em>Pararistolochia</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. goldieana</em> (Hook.f.) Hutch. &amp; Dalz&lt;sup&gt;7&lt;/sup&gt;</td>
<td>sparse</td>
<td>moderate</td>
<td>no</td>
</tr>
<tr>
<td><em>A. prevenosa</em> F.Muell.&lt;sup&gt;9&lt;/sup&gt;</td>
<td>sparse</td>
<td>moderate</td>
<td>no</td>
</tr>
<tr>
<td><em>A. promissa</em> (Mast.) Keay&lt;sup&gt;9&lt;/sup&gt;</td>
<td>sparse</td>
<td>moderate</td>
<td>no</td>
</tr>
<tr>
<td><em>A. triactina</em> (Hook. f.) Hutch &amp; Dalz</td>
<td>sparse</td>
<td>moderate</td>
<td>no</td>
</tr>
<tr>
<td>subgenus <em>Aristolochia</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. acuminata</em> Lam.</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>high&lt;sup&gt;5&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. anguicida</em> Jacq.</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>high&lt;sup&gt;4,8&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. clematitis</em> L.</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>high&lt;sup&gt;2,4,7&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><strong>A. elegans</strong> Mast.</td>
<td>ample</td>
<td>moderate&lt;sup&gt;4,8&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. fimbriata</em> Cham.</td>
<td>ample</td>
<td>low&lt;sup&gt;2,4&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. gigantea</em> Mart. &amp; Zucc.</td>
<td>sparse</td>
<td>moderate&lt;sup&gt;5,8&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. grandiflora</em> Sw.</td>
<td>sparse</td>
<td>high&lt;sup&gt;5,8&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. passiflorafolia</em> Rich., A.</td>
<td>sparse</td>
<td>low&lt;sup&gt;4&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. ringens</em> Vahl.</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>high&lt;sup&gt;5,8&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. trilobata</em> L.</td>
<td>sparse&lt;sup&gt;1&lt;/sup&gt;</td>
<td>high&lt;sup&gt;5&lt;/sup&gt;</td>
<td>no</td>
</tr>
<tr>
<td><em>A. lindneri</em> Berg.&lt;sup&gt;9&lt;/sup&gt;</td>
<td>sparse</td>
<td>moderate</td>
<td>no</td>
</tr>
<tr>
<td><em>A. maxima</em> Jacq&lt;sup&gt;9&lt;/sup&gt;</td>
<td>sparse</td>
<td>moderate</td>
<td>no</td>
</tr>
<tr>
<td><em>Aristolochia. sp.</em>&lt;sup&gt;9&lt;/sup&gt;</td>
<td>sparse</td>
<td>moderate</td>
<td>no</td>
</tr>
<tr>
<td>genus <em>Thottea</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>T. siliquosa</em> (Lam.) Hou&lt;sup&gt;9&lt;/sup&gt;</td>
<td>sparse</td>
<td>moderate&lt;sup&gt;6,7&lt;/sup&gt;</td>
<td>no</td>
</tr>
</tbody>
</table>

<sup>1</sup>few or no flowers in first season  
<sup>2</sup>herbaceous perennial  
<sup>3</sup>diminutive vine  
<sup>4</sup>vine  
<sup>5</sup>very large vine  
<sup>6</sup>low tree/shrub  
<sup>7</sup>slow growing  
<sup>8</sup>aggressive  
<sup>9</sup>Botanical Garden, Dresden  
<sup>10</sup>Botanical Garden, Bonn
**Cultivation Supplement**

**Background**

Several species of *Aristolochia* have been reported to be protogynous, and self-compatible (*A. fimbriata, A. elegans, A. ridicula, A. ringens; A. bracteata* (Petch, 1924; Razzak et al., 1992); *A. inflata* (Sakai, 2002). Autonomous self-pollination has been reported in *A. paucinervis*, (Berjano et al., 2006) and cleistogamy has been suspected in *A. serpentaria*. Control of pollination events, including preventing unintended pollen delivery, is essential for reproducible genetic experiments. Therefore, we set out to 1) determine if pollination could be prevented, 2) document post-anthesis changes in the perianth and gynoecium useful for evaluating the receptiveness of the stigma and 3) identify the day on which hand pollinations would be most successful. We evaluated fruit production from autogamous pollination, in which both the pollen and stigmatic surfaces were from the same flower (Faegri and Van Der Pijl, 1979). We also evaluated fruit production from geitogamous pollination, in which pollen was applied to stigmatic surfaces of a different flower on the same plant, selecting older flowers as pollen donor to better represent the presumed protogynous strategy used by *Aristolochia*. Results were compared with “outcrossings,” (xenogamous pollinations) and fruit set on bagged flowers (Faegri and Van Der Pijl, 1979). We subsequently evaluated seed viability from pods produced by self-pollination events, and compared it with the viability of bulk seed collected from open air pollinations. Open air pollinated fruit was collected from the general population of plants in greenhouse rooms housing both VL and NV cultivars, as well as large community of pollinators, which were mostly common greenhouse “fungus
gnats.” We confirmed reports of self-compatibility in *A. elegans* and *A. fimbriata*, and regularly germinated seeds (up to three years old) produced from self-pollinations and open-air pollinations.

**Perianth maturation**

The perianth of *A. fimbriata* opens early on the day of anthesis (Figure 2S-1 A), hence referred to as day one (D1). The lobes of the gynostemium of the D1 flower appear smooth, clean, and curved slightly inward (Figure 2S-1 D). The anthers, located on the outside of the gynostemium are closed, but a portion of an anther locule can be pried loose and placed onto the stigmatic surfaces to effect self-pollination on D1. The fimbriae of the limb of day two (D2) flowers are curved inward, covering the tube and preventing the exit of any trapped pollinators (Figure 2S-1 B). The lobes of the gynostemium are no longer curved inward and pollen is visibly dehiscing from the locules on the outer surfaces of the gynostemium lobes (Figure 2S-1 E). The loose pollen in D2 flowers (and later) can be collected on the tip of the toothpick from the dehiscing anthers on the outside of the gynostemium. The perianth of the day three (D3) flower is beginning to visibly decline (Figure 2S-1 C). The D3 flower releases pollen abundantly, the stigma has lost turgor, and the lobes of the gynostemium are beginning to recurve outward (Figure 2S-1 F).
Hand pollination methods

Selected flowers were enclosed in a pollination bag at an early stage of development to prevent pollinator access to the flower after anthesis. Each pollination bag was constructed of an approximately 12 cm x 30 cm piece of tulle (bridal veil) with an approx. 1.5 mm diameter pore size. The piece of tulle was folded in half, to create the closed end of the bag. The two long sides were sewn shut with a strong, durable, upholstery thread using 2 cm long stitches, leaving two drawstrings at the corners of the

Figure 2S-1: *Aristolochia fimbriata* flower A. Day one perianth at anthesis B. Day two perianth the day after anthesis C. Day three perianth D. Day one gynostemium; anthers have not dehisced E. Day two gynostemium; anthers have dehisced F. Day three gynostemium, pollen abundant.
open, shorter side. Pollination bags were wrapped gently two to three times around a stem containing one or more flower buds, then the drawstring on the bag was tied around the stem below the petiole of the most proximal flower. To conduct hand pollination, the pollination bag was removed from a newly opened flower, and the perianth was removed about halfway through the utricle (Figure 2S-2 A). A toothpick was used to transfer pollen from the anther locules on the outside of the gynostemium of the pollen donor and place it on the stigmatic tissue on the inner surfaces of the gynostemium of the pollen recipient (Figure 2S-2 B). In the self pollinations conducted in these experiments, a single flower was used both as the pollen donor and recipient (autogamous pollination). The mature fruit was harvested approximately 4 weeks after hand pollination, when the capsule dehisced along the six carpel valves, revealing dry seed (Figure 2S-2 C), which was stored in paper envelopes at ambient room conditions. Seeds collected from pods harvested before fully mature germinated poorly, if at all (data not shown).

Figure 2S-2: Hand pollination, A. fimbriata A. Sever the perianth midway through the utricle B. Apply pollen to stigmatic surfaces. (Seal to prevent further pollen delivery) C. Mature fruit dehiscing.
Pollination is prevented with pollination bags

Pollination bags prevented pollination in *A. fimbriata* except on one occasion, representing less than one percent of all occurrences (Table 2S-3). On that occasion, the perianth of the “not pollinated” flower producing fruit was lost, so it could not be inspected for the presence of a pollinator, which was likely responsible for the pollination event. Throughout the experiments, holes in the pollination bag and poorly tied bags were occasionally noted. Observation of an insect in the perianth at any time invalidated hand pollination efforts for that flower.

Table 2S-3: Evaluating self-compatibility in *A. fimbriata*. Day two autogamous pollination yielded greatest percent fruit set and germination success (bold).

<table>
<thead>
<tr>
<th>Self-pollination (autogamous)</th>
<th>Percent fruit set (sample size) (^1)</th>
<th>Percent germination success (sample size) (^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>19.3 (57)</td>
<td>28.0 (164)</td>
</tr>
<tr>
<td>D2</td>
<td>29.2 (154)</td>
<td>49.6 (733)</td>
</tr>
<tr>
<td>D3</td>
<td>7.9 (76)</td>
<td>29.9 (128)</td>
</tr>
<tr>
<td>D4</td>
<td>0 (12)</td>
<td>0</td>
</tr>
<tr>
<td>not pollinated</td>
<td>0.8 (127)(^3)</td>
<td></td>
</tr>
<tr>
<td>not bagged</td>
<td></td>
<td>59.0 (144)</td>
</tr>
</tbody>
</table>
Day two autogamous pollinations are most successful

Hand pollination in *A. fimbriata* was most successful on day two, both in terms of the numbers of seed pods produced from attempted pollinations and the viability of seed produced in those fruits (Table 2S-3). Pods were produced from 29% of the hand pollinations attempted on day two, from 19% of the pollinations attempted on day one, and only 8% of hand pollination attempted on day three. The gynostemia of day four flowers were fragile and often broke, constraining efforts to attempt hand pollinations in flowers more than two days beyond the day of anthesis. Fruits produced by hand pollination contained a far more variable number of seeds than fruit produced by open air, insect pollinator-mediated pollination, which typically contained 10/15 seeds per carpel, six carpels per fruit. The reduced seed production in fruits from self (hand-) pollination events may be due to limited pollen transfer with hand pollination, or to mechanical damage.
Self-compatibility in *Aristolochia elegans*

Autogamous, geitonogamous, and xenogamous pollinations were evaluated using the hand pollination methods described for *A. fimbriata*, using very small sample sizes. Self pollinations of day one and day two flowers were very successful (Table 2S-4). Viability of seed produced by hand-pollinations was not evaluated for *A. elegans*.

Table 2S-4: Percent fruit set (sample size) in *A. elegans* resulting from hand pollinations of flowers on days 1-4 with autogamous, geitonogamous and xenogamous pollen from flowers on days 1-3. Compare to 0% fruit set on 19 flowers, bagged before anthesis to prevent pollinator entry. Day 1 and 2 autogamous pollinations worked well, as did xenogamous pollinations (bold).

<table>
<thead>
<tr>
<th>Age of stigmatic surfaces ↓</th>
<th>Pollination type</th>
<th>Age of pollen source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Autogamous</td>
<td>Geitonogamous</td>
</tr>
<tr>
<td>D1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2</td>
<td>90 (10)</td>
<td></td>
</tr>
<tr>
<td>D3</td>
<td>33 (3)</td>
<td></td>
</tr>
<tr>
<td>D4</td>
<td>20 (5)</td>
<td></td>
</tr>
<tr>
<td>D5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Seed germination methods

We developed *in vitro* methods for germinating seeds to support collection of seedling tissues and for comparisons of seed viability. Seeds collected from fruits
produced by open-air pollination of mixed VL and NV greenhouse stocks germinated significantly better in wet toweling packets in a growth chamber providing 16h days, 8h nights, 30°C days, 16°C nights than in sucrose-free media. Seeds incubated in 16h light/8 h did not germinate significantly better than seeds incubated in total darkness, but the first set of true leaves never developed in seedlings emerging in total darkness, thus the etiolated seedlings died (data not shown). However, seeds germinated most quickly in potting medium, with bottom heat, in the greenhouse (up to 95% in two weeks) than in wet toweling in a growth chamber. Furthermore, seedlings transferred from wet toweling to potting medium in the greenhouse after the first set of true leaves appeared typically bloomed two to three weeks later (at an older age) than did seeds germinated in potting medium in the greenhouse. This delay reflects the additional resources available to seeds germinating in potting medium, evident as a larger rooted seedling at four weeks than observed on seedlings germinated in toweling (data not shown).

**Effect of seed age on germination**

Using *in vitro* (in wet toweling, 25 seeds per replicate) germination, we evaluated the effect of age on seed viability and found no significant difference in total percent germination of seeds planted at seven day intervals, from 7 days after collection through 147 days after collection, and mean germination was 90% (n=15) by day 63 (data not shown). Furthermore, germination in 2.5 year old seeds was not significantly lower than in 2 year old or 1.5 year old seeds, and mean germination was 80% (n=12) by day 77.
Table 2S-5: Genome sizes, vouchers, sources and accessions for sequence data used. Abbreviations: R, Rice (*Oryza sativa* ssp. *japonica* cv. ‘Nipponbare’ 0.9 pg/2C); S, Soybean (*Glycine max* cv. ‘Dunbar’ 2.35 pg/2C), T, Tobacco (*Nicotiana tabacum* cv. ‘SR-1’ 9.15 pg/2C); BG, Botanical Garden; PAC, Pennsylvania State University; DR, Dresden. Mbp DNA for plant species is based on the assumption 1pg=980 Mbp according to (Pfeifer, 1966; Wang et al., 2005). (BJB=collection number of the first author). XXX Genbank accession numbers -- need to submit them when reviews come back.

![Table](image)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>genome size (pg/2C) +/- std. dev. (n=4)</th>
<th>Standard DNA content of sample species (Mbp/1C)</th>
<th>Voucher, herbarium</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aristolochia acuminata</em> Lam.</td>
<td>1.15 +/-0.007</td>
<td>S 564</td>
<td>BJB06.06A, PAC</td>
<td>Victor Wong (private coll.)</td>
</tr>
<tr>
<td><em>Aristolochia anguicida</em> Jacq.</td>
<td>0.81 +/-0.005</td>
<td>S 397</td>
<td>BJB06.03A, PAC</td>
<td>Mario Blanco (private coll.)</td>
</tr>
<tr>
<td><em>Aristolochia californica</em> Torr.</td>
<td>1.58 +/-0.006</td>
<td>S 774</td>
<td>BG Dresden</td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia clematitis</em> L.</td>
<td>0.96 +/-0.001</td>
<td>S 470</td>
<td>BJB03.07A, PAC</td>
<td>Seneca Hill Perennials, New York</td>
</tr>
<tr>
<td><em>Aristolochia elegans</em> Mast.</td>
<td>0.81 +/-0.007</td>
<td>S 397</td>
<td>BJB03.02A, PAC</td>
<td>Park Seed Company, cat. #0179-7</td>
</tr>
<tr>
<td><em>Aristolochia fimbriata</em> Cham.</td>
<td>0.91 +/-0.004</td>
<td>S 446</td>
<td>BJB03.04A, PAC</td>
<td>Larry D. Rosen (“VL”) (private coll.)</td>
</tr>
<tr>
<td><em>Aristolochia gigantea</em> Mart. &amp; Zuec.</td>
<td>0.89 +/-0.004</td>
<td>S 470</td>
<td>BJB03.05A, PAC</td>
<td>Russ Stover (“VL”) (private coll.)</td>
</tr>
<tr>
<td><em>Aristolochia grandiflora</em> Sw.</td>
<td>1.19 +/-0.009</td>
<td>S 583</td>
<td>BJB03.02A, PAC</td>
<td>Kartuz Greenhouses, California</td>
</tr>
<tr>
<td></td>
<td>1.20 +/-0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.17 +/-0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.21 +/-0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.13 +/-0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taxon</td>
<td>Size</td>
<td>Country</td>
<td>Collection Details</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia goldieana</em> (Hook.f.) Hutch. &amp; Dalz.</td>
<td>1.26 +/- 0.006</td>
<td>617</td>
<td>Wanke &amp; Neinhuis s.n., DR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>BG Dresden</td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia holostylis</em> (Duchartre) F. Gonzalez</td>
<td>4.28 +/- 0.113</td>
<td>S 2097</td>
<td>Neinhuis 117, DR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG Dresden</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia lindneri</em> Berg.</td>
<td>0.67 +/- 0.018</td>
<td>S 328</td>
<td>BG Dresden, Bolivia, San Jose de Chiquitos</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG Munich</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia macrophylla</em> Lam.</td>
<td>1.54 +/- 0.011</td>
<td>S 755</td>
<td>Neinhuis s.n., DR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.53 +/- 0.033</td>
<td>750</td>
<td>BG Munich</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.52 +/- 0.025</td>
<td>745</td>
<td>Dawes Arboretum, Ohio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.52 +/- 0.018</td>
<td>745 BJBO4.07A, PAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia maxima</em> Jacq</td>
<td>0.77 +/- 0.009</td>
<td>S 755</td>
<td>N.Pabon-Mora &amp; F. Gonzalez, NY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.78 +/- 0.017</td>
<td>S 764</td>
<td>NYBG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gonzalez 4018, COL</td>
<td>Panama, Parnama,</td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia. sp.</em></td>
<td>0.74 +/- 0.017</td>
<td>S 363</td>
<td>Wanke &amp; Neinhuis s.n., DR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG Munich</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia passiflorafolia</em> Rich.</td>
<td>0.74 +/- 0.006</td>
<td>S 363</td>
<td>Mario Blanco (private coll.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BJBO6.05A, PAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia prevenosa</em> F.Muell.</td>
<td>1.83 +/- 0.006</td>
<td>S 1793</td>
<td>Neinhuis &amp; Wanke s.n., DR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG Dresden, Queensland, Australia, BG Dresden</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia promissa</em> (Mast.) Keay</td>
<td>4.41 +/- 0.013</td>
<td>S 4321</td>
<td>Neinhuis 118, DR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG Dresden</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia ringens</em> Vahl.</td>
<td>0.93 +/- 0.004</td>
<td>S 456</td>
<td>Mario Blanco (private coll.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.91 +/- 0.003</td>
<td>446 BJBO6.07A, PAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia serpentaria</em> L.</td>
<td>1.69 +/- 0.024</td>
<td>S 828</td>
<td>Larry D. Rosen (private coll.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.67 +/- 0.019</td>
<td>818 BJBO3.03A, PAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.57 +/- 0.018</td>
<td>769</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.56 +/- 0.013</td>
<td>764</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.61 +/- 0.035</td>
<td>789</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.58 +/- 0.008</td>
<td>774 BJBO5.01A, PAC</td>
<td>B&amp;T World Seeds</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Priv. coll. B. Westlund</td>
<td>USA, Texas, Travis Co.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>784</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia tomentosa</em> Sims</td>
<td>1.13 +/- 0.008</td>
<td>S 554</td>
<td>Seneca Hill Perennials, New York</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.39 +/- 0.010</td>
<td>681 BJBO3.06A, PAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.40 +/- 0.022</td>
<td>686</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.44 +/- 0.011</td>
<td>706 BJBO6.01A, PAC</td>
<td>B&amp;T World Seeds</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Priv. coll. B. Westlund</td>
<td>USA, Texas, Travis Co.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>774</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neinhuis 113, DR</td>
<td>BG Dresden</td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia triactina</em> (Hook, f.) Hutch &amp; Dalz.</td>
<td>4.39 +/- 0.059</td>
<td>S 4302</td>
<td>Neinhuis 119, DR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG Dresden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Sample</td>
<td>Value</td>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------</td>
<td>-----------</td>
<td>-------------------------------</td>
<td></td>
</tr>
<tr>
<td><em>Aristolochia trilobata</em> L.</td>
<td>S</td>
<td>0.91 +/- 0.002</td>
<td>446</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.99 +/- 0.010</td>
<td>485</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.01 +/- 0.002</td>
<td>495</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>1.02 +/- 0.004</td>
<td>500 BJB04.04A, PAC Kartuz Greenhouses, California</td>
<td></td>
</tr>
<tr>
<td><em>Asarum canadense</em> L.</td>
<td>S</td>
<td>10.19 +/- 0.040</td>
<td>4993</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.97 +/- 0.083</td>
<td>4885</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>11.04 +/- 0.307</td>
<td>5410 BJB04.03A, PAC Joel McNeal, (private coll.)</td>
<td></td>
</tr>
<tr>
<td><em>Saruma henryii</em> Oliv.</td>
<td>T</td>
<td>6.12 +/- 0.044</td>
<td>2999 BJB06.08A, PAC Heronswood Nursery, Washington</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neinhuis 120, DR</td>
<td>BG Dresden</td>
<td></td>
</tr>
<tr>
<td><em>Thottea siliquosa</em> (Lam.) Hou</td>
<td>S</td>
<td>1.25 +/- 0.018</td>
<td>613 Neinhuis 121, DR India, Kerala, BG Dresden</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2S-3: *Aristolochia fimbriata* genotype and perianth detail A, B VL genotype  C, D NV genotype A, C Presence, absence of leaf variegation B, D Perianth varies in shape and color E. Perianth is highly modified for insect pollination. Modifications include limb (li), tube (tu), syrinx (sy), utricle (ut) and gynostemium (gy), which has stamen locules on the outside and interior stigmatic surfaces. Glass model by Leopold and Rudolph Blatschka illustrated by Fritz Kredel (reproduced with permission).
Figure 2S-4: Phylogram of Aristolochiaceae relationships, showing minimal variation in branch lengths within Aristolochiaceae. Only one tree was found in search with maximum likelihood methods.
Chapter 3

*In vitro* propagation, shoot regeneration and rooting protocols for *Aristolochia fimbriata*, potential model plant for basal angiosperms

This chapter is a manuscript for submission by Barbara J. Bliss, Hong Ma, Claude dePamphilis, Lena Landherr, Yi Hu, Siela N. Maximova to *Plant Cell, Tissue and Organ Culture*.

This work resulted from collaboration among the labs of Drs. Claude dePamphilis, Hong Ma, and Siela Maximova. Lena Landherr performed experiments optimizing reduced light/enriched nutrient micropropagation methods (REN2, Table 3-1), and initial stem/petiole segment regeneration methods (Table 3-6). Yi Hu assisted with micropropagation and regeneration experiments (Tables 3-3, 3-4). All other results were generated by Barbara Bliss.
Short original article for submission to Plant Cell. Tissue and Organ Culture

Title: Regeneration and plantlet development from somatic tissues of *Aristolochia fimbriata*

Authors: Barbara J. Bliss\(^1\), Lena Landherr\(^1\), Claude W. dePamphilis\(^1\), Hong Ma\(^1\), Yi Hu\(^1\), Siela N. Maximova\(^2\)

\(^1\)Department of Biology, Institute of Molecular Evolutionary Genetics, and the Huck Institutes of the Life Sciences, 201 Life Sciences Building, The Pennsylvania State University, University Park, PA 16802, USA

\(^2\)The Department of Horticulture, 421 Life Sciences Building, The Pennsylvania State University, University Park, PA 16802, USA

S. N. Maximova (author for correspondence);
The Department of Horticulture, 421 Life Sciences Building, The Pennsylvania State University, University Park, PA 16802, USA

e-mail: snm104@psu.edu

Tel: 814-863-7286

Fax: 814-865-1357

Suggested running head: Regeneration and plantlet development from somatic tissues of *Aristolochia fimbriata*
Abstract

*Aristolochia fimbriata* is a small herbaceous perennial in the basal angiosperm family *Aristolochiaceae*. The family contains diverse floral forms ranging from radial to monosymmetric flowers with a wide variety of insect pollinators. Additionally, *Aristolochia* species contain secondary metabolites that are important natural toxins and traditional medicines, and are critical to the reproduction of swallowtail butterflies. These characteristics, in combination with the small genome size and short life cycle of *Aristolochia fimbriata*, have prompted further development of this species as a model system to study the evolution of basal angiosperms. As a prerequisite for developing a genetic transformation procedure for *Aristolochia*, we have developed protocols for *in vitro* plant multiplication, shoot organogenesis, rooting, and acclimation of tissue culture-derived *Aristolochia* plants. Nodal cuttings from two different *Aristolochia* genotypes were sterilized and cultured for *in vitro* propagation. Explants were further multiplied in large numbers and rooted with up to 100% success. Shoot regeneration was achieved from whole leaf, internodal stem, and petiole explants. Regeneration was observed starting at 28 days after culture initiation on MS basal medium supplemented with 0.5 mg/L 6-benzoamino-purine (6BA), 1.0 mg/L α-naphthalene acetic acid (NAA), and 1.0 mg/L thidiazuron (TDZ), followed by regeneration on MS basal medium supplemented with 1.75 mg/L 6BA and 1.0 mg/L NAA. Maximal shoot primordia formation occurred at frequencies up to 97%, with stem explants. Large numbers of regenerated shoots were rooted and acclimated to greenhouse conditions and developed flowers after 2-4 weeks.
Keywords: *Aristolochia fimbriata*, basal angiosperm, micropropagation, shoot organogenesis
**Introduction**

*Aristolochia fimbriata* is a flowering perennial plant cultivated for its interesting flowers, attraction of butterflies, and traditional medicinal properties. The flowers of *Aristolochia* species have a unipartite, monosymmetric perianth adapted for insect pollination, while other genera in *Aristolochiaceae* have radially symmetric flowers (e.g., *Saruma, Asarum* (Gonzalez and Stevenson, 2000b). *Aristolochia* species have been studied for their pollination strategies associated with floral aroma and a fly-trapping perianth (Petch, 1924; Hall and Brown, 1993; Sakai, 2002; Banziger et al., 2006; Murugan et al., 2006; Trujillo and Sersic, 2006), as well as for species-specific host plant relationships with swallowtail butterfly larvae. Secondary metabolites produced by *Aristolochia* plants are critical for the defense and survival of the butterflies during their larval feeding stage (Rausher, 1981; Klitzke and Brown, 2000), such that the decline of particular swallowtail butterfly populations is attributed to the declining distribution of particular *Aristolochia* species (Sands et al., 1997). These secondary metabolites include aristolochic acids and aristolactams produced via alkaloid biosynthesis pathways (Kumar et al., 2003).

*Aristolochia* species, particularly root tissues, have been used as traditional medicines (Reddy et al., 1995) and have been subsequently studied to reveal the molecular mechanisms underlying their observed human health effects (Lemos et al., 1993; Levi et al., 1998; Nortier et al., 2000; Qiu et al., 2000; Hranjec et al., 2005; Hwang et al., 2006; Meinl et al., 2006). The potent toxins found in *Aristolochia* offer promise in specific applications, including antivenom (Otero et al., 2000; Jimenez-Ferrer et al.,
2005; Abubakar et al., 2006), antibacterial (Gadhi et al., 1999; Gadhi et al., 2001a; Gadhi et al., 2001b), antifertility (Pakrashi and Chakrabarty, 1978; Pakrashi and Pakrasi, 1979; Gupta et al., 1996), cytotoxic (Kupchan and Doskotch, 1962; Hinou et al., 1990), antimicrobial or trypanocidal (Abe et al., 2002; Elizabeth and Raju, 2006; Kumar et al., 2006), and insecticidal (Lajide et al., 1993; Broussalis et al., 1999; Nascimento et al., 2004; Jbilou et al., 2006).

In addition to its medicinal value and importance for butterfly reproduction, *Aristolochia* occupies an important phylogenetic position. As a basal angiosperm from the magnoliid clade (Jansen et al., 2007), it offers uncommon opportunities for studying the evolution of development in flowering plants. Evolutionarily conserved developmental pathways, such as those controlling floral organ identities in angiosperms, have been identified via functional studies in model plants including *Arabidopsis*, *Petunia*, *Zea*, and *Oryza* (Kramer et al., 1998; Ma and dePamphilis, 2000; Whipple et al., 2004; Agrawal et al., 2005; Zahn et al., 2005a) Comparative studies of these plant species have led to hypotheses about the evolution of gene function in flowering plants (Jaramillo and Kramer, 2004; Kim et al., 2006). However, the monocot and eudicot lineages from which current model systems arose diverged only about 113-133 million years ago (mya), some 28-48 million years after the angiosperm divergence (approximately 161 mya) (Bell et al., 2005; Leebens-Mack et al., 2005). If pathways found in monocots and/or core eudicots are also found to be present in a basal angiosperm such as *Aristolochia*, this would imply that the conserved components were present in angiosperms before the origin of these two major lineages of angiosperms. Comparative gene-functional analysis between basal angiosperms and derived model
plants is a powerful approach to test this hypothesis. Therefore, the development of a basal angiosperm species as an experimental model plant will be of great benefit to the study of plant evolutionary biology.

Plant genetic manipulation is essential for testing hypotheses about gene function, and Agrobacterium-mediated transformation offers a convenient and efficient tool for this purpose. Transformation can be used to generate both loss-of-function and gain-of-function transgenic plants. The easy-to-use and efficient in planta “dip” transformation method currently used with Arabidopsis thaliana is rare among plant systems (Bent, 2000). Traditional tissue culture based transformation systems are used in other model plant species, including Nicotiana tabacum (An et al., 1986), Medicago truncatula (McKersie et al., 1997), Lotus japonica (Handberg and Stougaard, 1992), Petunia hybrida (Napoli et al., 1990), Lycopersicon esculentum (Dan et al., 2006), Brachypodium distachyon (Draper et al., 2001), Oryza (Sallaud et al., 2003) and Zea (Frame et al., 2002).

Efficient propagation and regeneration systems for Aristolochia fimbriata are important prerequisites for the development of Agrobacterium tumefaciens-mediated transformation. Prior to this study, in vitro methods for Aristolochia indica, a species valued for medicinal use (Shafi et al., 2002), were developed using axillary shoot multiplication and shoot organogenesis to provide source tissue for the purification of secondary metabolites (Remashree et al., 1997); (Soniya and Sujitha, 2006). Here we report the development of methods for micropropagation and shoot regeneration of two genotypes of A. fimbriata, a species with a small genome size (Bharathan et al., 1994) and promising characteristics for use as a basal angiosperm model system, such as short
generation time and small size. These methods will provide tools for fast propagation of large number of clonal plants and will support the development of genetic transformation methods.

Materials and Methods

Plant material

Seeds of *Aristolochia fimbriata* were provided by Larry D. Rosen, Florida, USA and by Jardim Botanico, Departamento de Botanica, Universidade de Coimbra, Portugal. Plants were grown from the seeds and maintained in a greenhouse at The Pennsylvania State University, University Park, PA. All seeds were germinated in soil-free potting medium (Pro-Mix BX, Premier Horticulture Inc., Quakertown, PA) in shallow germination trays with drainage holes, in the greenhouse at 18-27°C (varying from night to day) and 40-70% humidity. The trays were incubated on heating mats operating at approximately 27°C. Natural day length was supplemented with high-pressure sodium lamps (1000 watt) from October through April to provide twelve-hour days. Plants received regular watering as needed. Depending on the stage of growth, regular fertilizer applications were provided, as a drench, alternating Peter's Professional 15-16-17 Peat Lite Special at 200 PPM nitrogen (once to twice weekly) with Peters Professional 21-7-7 Acid Special (Scotts Horticulture, Marysville, OH) at 200 PPM nitrogen (approximately every six weeks). The plants were drenched once a month with 100 ppm chelated iron (Sprint 330 10% iron, RoseCare.com, Santa Barbara, CA). Plants grown from Florida
seeds had variegated leaves (VL), while those from Coimbra were not variegated (NV) (Figure 3-1 A and B).

**Micropropagation initiation and multiplication**

Six month old, healthy, flowering plants of both NV and VL genotypes grown in greenhouse were selected for *in vitro* propagation. Green stems were collected from newly developed branch terminals (NV: n=30, VL: n=35, VL: n=60). Only the three most distal (youngest) nodes per stem were used, such that the excised explants were 1 - 2 cm long, 1.5 – 2.0 mm in diameter, and contained one terminal or axillary bud. Fully expanded leaves were removed and the resulting green nodal cuttings/explants were sterilized for 20 minutes in 1% Chlorox solution (Commercial Solutions®; 6.15% sodium hypochlorite) with 1% Tween 20 detergent (polyoxyethylene sorbitan monolaurate; Sigma Chemical Co., St. Louis, MO, USA). Explants were rinsed three times in sterile de-ionized water and stems freshly trimmed to a 45° angle before being inserted into 60-75 mL micropropagation initiation medium (MI, Table 1) in disposable plastic food cups (Sweetheart, Owings Mills, MD, #DSD8X and #LDS58). Five explants were placed in each cup. Cultures were incubated in a Conviron growth chamber (Winnipeg, Canada) at 25°C with 300 micromoles m⁻² sec⁻¹ light (16 hours a day) for approximately 20 days. For further *in vitro* multiplication, single node stem explants 1-2 cm long with a terminal or axillary shoot were excised from the newly developed tissue cultured *A. fimbriata* plants and cultured on different micropropagation media (MP). The MP media evaluated
contained 6-benzoaminopurine (6BA) and α-naphthalene acetic acid (NAA) in eight combinations (Table 3-1). The number of shoots propagated from each greenhouse cutting was counted after 21 days and significant differences among the media were determined by the nonparametric Mann-Whitney test, using Minitab WINSV12.11 software (State College, Pennsylvania, U.S.A.).

Table 3-1: Media formulations for micropropagation and shoot organogenesis.

<table>
<thead>
<tr>
<th>Media Components</th>
<th>MI</th>
<th>MP</th>
<th>RI</th>
<th>REN1</th>
<th>REN2</th>
<th>SI</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS basal medium (g/L)</td>
<td>4.4</td>
<td>4.4</td>
<td>2.2</td>
<td>2.2</td>
<td>3.3</td>
<td>4.4</td>
<td>4.4</td>
</tr>
<tr>
<td>6BA (mg/L)</td>
<td>2.5</td>
<td>-</td>
<td>1.0</td>
<td>5.0</td>
<td>-</td>
<td>-</td>
<td>0.0 - 0.5</td>
</tr>
<tr>
<td>NAA (mg/L)</td>
<td>-</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
<td>0.50 - 1.75</td>
<td>1.0</td>
</tr>
<tr>
<td>TDZ (mg/L)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0 - 1.1</td>
<td>-</td>
</tr>
<tr>
<td>IBA (mg/L)</td>
<td>0.25</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sucrose (g/L)</td>
<td>30.0</td>
<td>30.0</td>
<td>20.0</td>
<td>20.0</td>
<td>20.0</td>
<td>30.0</td>
<td>30.0</td>
</tr>
<tr>
<td>PhytaGel (g/L)</td>
<td>2.0</td>
<td>2.5</td>
<td>2.7</td>
<td>2.8</td>
<td>2.7</td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td>pH</td>
<td>5.5</td>
<td>5.7</td>
<td>5.5</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Abbreviations: MI - micropropagation initiation; MP – micropropagation medium; RI - root initiation; REN1 - root elongation #1; REN2 - root elongation #2; SI - shoot induction; SR – shoot regeneration; MS - Murashige and Skoog basal medium; 6BA - 6-benzoaminopurine; NAA - α-naphthalene acetic acid; TDZ - thidiazuron; IBA - indole-3-butyric acid.
In vitro rooting

*In vitro* propagated apical or axillary shoots with two nodes and two buds were excised from propagules in MP medium. Fully expanded leaves were removed, the stems freshly cut at a 45 degree angle and inserted into hormone-free modified root elongation medium #1 or #2 (REN1 or REN2) in which the agar in the original RE medium (Maximova et al., 1998), was substituted with PhytaGel (REN1, REN2, Table 3-1). Additionally, REN2 medium contained 50% more MS salts and vitamins than REN1. Explants were placed in the medium straight up so that the lowest node and stem below were fully immersed. Explants were incubated in the dark for 3 days and then transferred to low-light conditions (20-100 micromoles m$^{-2}$ sec$^{-1}$, 16 hours a day) at 25°C. When roots were visible, each plant was transferred to an individual container with fresh REN1 (cultured under high-light conditions: 300 micromoles m$^{-2}$ sec$^{-1}$ light, 16 hours a day) or REN2 medium (cultured under medium-light conditions: 150 micromoles m$^{-2}$ sec$^{-1}$) at 25°C. Rooted plants were maintained on REN1 or REN2 media and every 21-28 days, plants were trimmed and transferred to fresh medium. Plant/shoot clusters of more than three shoots were divided at the crown so that individual rooted plants and non-rooted shoots were separated and transferred to fresh medium. Only rooted plants were used as stock plants in further rooting and regeneration experiments.

The effect of incubation with different concentrations of indole-3-butyric acid (IBA) on rooting efficiency was evaluated. For this experiment 8-15 stem explants from NV and VL genotypes were excised from rooted plantlets in REN1 medium and inserted into solid root initiation (RI) medium supplemented with 0, 0.5, or 1.0 mg/L IBA (RI,

Table 3-1). Explants contained two nodes and fully expanded leaves. Explants in RI medium were incubated in the dark for 3 days at 25°C, then transferred to low light conditions until roots appeared, at which time the newly rooted plantlets were returned to culture on fresh hormone-free REN1 medium in a Conviron growth chamber under high-light conditions. The number of rooted plants was recorded after 30 days. To determine if any factors or interactions among factors had a significant effect on the number of roots produced or average root length the percentage rooting data were analyzed with the two-way general linear ANOVA model of Minitab WINSV12.11.

**Shoot regeneration from whole leaf explants**

All regeneration media evaluated contained 4.4 g/L Murashige and Skoog (MS) basal medium salts (Murashige and Skoog 1962) with Gamborg vitamins (Gamborg et al. 1968) (Sigma M5519), 30 g/l sucrose, solidified with PhytaGel (Sigma P8169) (Table 3-1). Leaf explants were cultured first on shoot induction media (SI) with different concentrations of 6BA, NAA and thiadiazuron (TDZ) (Table 3-1). SI media (50 ml aliquots) were poured in 100 x 20 mm Petri dishes (VWR cat. 25382-166, Becton Dickinson Falcon, Franklin Lakes, NJ, USA). The top two fully expanded, dark green apical leaves of primary or axillary shoots from *in vitro* rooted NV plants and VL plants on REN1 medium were cut into approximately 1.5 cm² explants. The explants from the base of the leaves included 0.5-1 cm (more than one half) of the petiole. After loosely chopping up the medium with sterile forceps, the explants were gently pressed, abaxial
surface up, into the medium to assure good contact. A total of 25 explants per genotype (5 explants per Petri dish) were cultured for the individual treatments evaluated. Cultures were incubated in the dark at 25°C for 14 days on SI medium. The explants were then transferred to shoot regeneration (SR) media with different concentrations of 6BA (Table 3-1), and incubated for an additional 14 days in the dark, at 25°C. During the course of the experiment the leaf explants expanded and were transferred from Petri dishes to disposable plastic food cups as explant size increased. After a total of 28 days in the dark, cultures on SR media were transferred to low-light for shoot elongation. The explants were inspected and transferred to fresh SR every 14 days, 4 more times. Explants producing shoot primordia were subdivided into clusters of 1-3 shoots as needed, to allow elongation and additional shoot proliferation. After 21-42 days of elongation, individual shoots with two extended nodes and leaves were transferred from SR media to rooting medium (REN1, Table 3-1) as described above. A binary logistic regression model was applied to all the media treatments producing shoots using Minitab WINSV12.11 to determine if plant genotype, level of NAA in the SI media, level of 6BA in the SR media, or interactions had a significant effect on the percent of explants regenerating.

**Shoot regeneration from petiole and stem sections**

For the first two sets of experiments, explants were generated by excising 2 - 3 mm long stem and petiole segments from the upper three nodes of dark green rooted *A. fimbriata* VL and NV plants maintained in REN1. Explants were placed horizontally on SI medium (Table 3-1) containing 0.5 mg/L 6BA, 1 mg/L NAA, and 1 mg/L TDZ, and
pressed lightly into the media to ensure adequate contact (up to 60 explants per plate). Cultures were incubated in the dark at 25-27°C for 14 days, and then transferred to low-light. Cultures were separated in three different groups and each group was incubated on SI for a different period of time including: 14, 21 or 40 days. After that, the cultures from each group were transferred to SR medium (Table 3-1) containing 1.75 mg/L 6BA and 1.0 mg/L NAA. All media were refreshed every 14 days. Number of explants regenerating was evaluated at 40 days. Using Minitab WINSV12.11, a binary logistic regression model was applied to determine if genotype, tissue type, length of time on SI media, or interactions significantly affected regeneration.

In a third set of experiments, stem and petiole explants (2-3 mm long) were selected from rooted VL plants maintained on REN2 medium (Table 3-1). The explants were cultured on SI medium (20 to 30 explants per plate) for 14 days as described above then transferred to fresh SR medium every 14 days. At each transfer, cultures were evaluated and explants with visible shoot primordia were moved to low-light. The rest of the explants were incubated in the dark until primordia were developed. When distinct leaves were visible, the base and sides of the newly regenerated shoots were trimmed to remove remaining callus tissue and the shoots were transferred to REN2 medium in sterile cups under low-light to promote shoot elongation. Explants on REN2 were transferred to fresh medium every 4-5 weeks until roots and shoots with fully expanded leaves were developed. Plants with two or more roots and minimum 3 cm total root length were transplanted to into multi-cell plant trays with 4-cell packs (Kord, Canada) filled with water-saturated soil-free potting mix consisting of one part Metro-Mix 200 (Sun Gro Horticulture. Ltd., Vancouver, B. C., Canada) and four parts Miracle-Gro.
Potting Mix, which contains Miracle-Gro slow-release plant fertilizer (Scotts Miracle-Gro, Scotts Horticulture, Marysville, OH). Potted plants were covered with clear plastic lids and placed in the growth chamber under medium-light, shaded with one layer of white paper towels for one day to reduce transplant shock. Plants were acclimated for one more week by gradually opening the plastic lids to reduce the humidity. The lids were completely removed by the end of day three. The acclimated plants already in soil were then transferred to the greenhouse (conditions previously described) where they were monitored weekly and the number of healthy plants was recorded.

Results

Micropropagation and in vitro rooting

To generate and maintain a large number of aseptic plants as a source of explant tissue for regeneration and transformation experiments, we developed protocols for *A. fimbriata* micropropagation and subsequent *in vitro* rooting (Fig. 3-2). Single node green cuttings from two different genotypes of *A. fimbriata* were obtained from greenhouse stock plants and introduced into tissue culture on MI medium (Table 3-1). All cuttings of both genotypes developed at least two new axillary shoots, doubling the number of shoots in 21 days. To optimize micropropagation, we evaluated the effect of NAA and 6BA concentrations in MP medium. At 21 days after culture initiation the mean multiplication rate for treatments without NAA was significantly greater (P<0.05) than for treatments containing NAA (Table 3-2). Statistical analysis also indicated that this response was not
a function of 6BA level (Table 3-2). However, our results demonstrated that explants in MP containing 1 mg/mL 6BA and no NAA had the highest mean multiplication rate of 3.06 fold (P<0.05) (Table 3-2).

Shoots obtained by multiplication or de novo organogenesis were further cultured for rooting. To optimize the rooting protocol, we experimented with RI medium supplemented with different concentrations of IBA (Table 3-1), providing a brief period of root induction, followed by transfer to hormone free medium (Fig. 3-2). The number of roots and average root length were recorded for each of the concentrations after 28 days of culture. The data met the normal distribution assumption of the two-way ANOVA model, but no significant effects of genotype, IBA concentration, or interaction were detected. Our results indicated that all explants placed in hormone-free medium

Figure 3-2: Schematic representation of micropropagation and regeneration protocols of A. fimbriata. The number of days at each step is indicated.
Table 3-2: The effect of plant growth regulators on axillary shoot proliferation at 21 days after culture initiation.

<table>
<thead>
<tr>
<th>Combination of plant growth regulators</th>
<th># New shoots harvested</th>
<th># Nodal cuttings initiated</th>
<th>Multiplication factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAA mg/L</td>
<td>6BA mg/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>3</td>
<td>36</td>
<td>25</td>
</tr>
<tr>
<td>0.5</td>
<td>3</td>
<td>36</td>
<td>20</td>
</tr>
<tr>
<td>0.1</td>
<td>5</td>
<td>50</td>
<td>22</td>
</tr>
<tr>
<td>0.1</td>
<td>4</td>
<td>48</td>
<td>25</td>
</tr>
<tr>
<td>0.1</td>
<td>3</td>
<td>46</td>
<td>24</td>
</tr>
<tr>
<td>0.1</td>
<td>2</td>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td><strong>0.0</strong></td>
<td>3</td>
<td><strong>47</strong></td>
<td><strong>20</strong></td>
</tr>
<tr>
<td><strong>0.0</strong></td>
<td>1</td>
<td><strong>52</strong></td>
<td><strong>17</strong></td>
</tr>
</tbody>
</table>

*Mean multiplication for treatments without NAA was significantly greater (P<0.05).

(REN1, Table 3-1) with the lowest node submerged, followed by culturing in low-light, developed roots in 21-30 days (Table 3-3). Therefore we discontinued use of RI medium and continued to introduce single shoots directly into REN1 or REN2 medium for rooting and further maintenance. The tissue culture initiation, multiplication, and rooting protocols reported here (Fig. 3-2) yielded approximately six times more plants than the original number of explants introduced into culture and have been successfully implemented with both genotypes (NV, VL) of greenhouse-cultivated plants ranging from three months to four years in age.

Rooted plants in the initial REN1 media developed large, deep green leaves, abundant roots, and multiple stems from a single crown, similar to plants grown in the greenhouse. Compact, rooted stock plants in REN1 media provided an ample supply of fresh leaves for shoot regeneration experiments. The stock plants were trimmed and
transferred to fresh REN1 medium every 3-5 weeks. To reduce the need for frequent transfer and to improve the overall health of the stock plants, during the later stages of the experiment, the rooting medium was modified by increasing the MS basal media concentration by 50% (REN2) and the light intensity was decreased by about 50% to 150 micromoles m\(^{-2}\) sec\(^{-1}\). Plants under these conditions produced longer internodes and reduced number of roots compared to the REN1 plants. The plants on REN2 also required media replacement less frequently (every 4-5 weeks).

Table 3-3: Effect of indole-3-butyric acid (IBA) on rooting of shoots multiplied in vitro. Two-way ANOVA indicated no significant differences (\(\alpha = 0.05\)) in average number of roots/shoot or average root length among IBA levels.

<table>
<thead>
<tr>
<th>IBA (mg/L)</th>
<th>% Rooted shoots (n)</th>
<th>Average number of roots/shoot ± SD (^a)</th>
<th>Average total root length/plant (mm) ± SD (^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>100 (17)</td>
<td>2.6 ± 1.1</td>
<td>32.8 ± 14.6</td>
</tr>
<tr>
<td>0.5</td>
<td>100 (29)</td>
<td>2.6 ± 1.6</td>
<td>29.2 ± 15.8</td>
</tr>
<tr>
<td>1.0</td>
<td>100 (27)</td>
<td>2.7 ± 1.6</td>
<td>27.2 ± 14.2</td>
</tr>
</tbody>
</table>

\(^a\) - Only the number of the primary adventitious roots was counted.
\(^b\) - All primary and secondary adventitious roots were measured.

**Shoot regeneration**

To develop shoot organogenesis methods for *A. fimbriata*, we subjected leaf explants to different concentrations of auxins and cytokinins in a two-step protocol consisting of shoot induction (SI) and shoot regeneration (SR) steps. Results from our preliminary experiments, titrating TDZ and 6BA in different combinations, indicated that shoot primordia were initiated only on SI media supplemented with 0.5 mg/L 6BA and 1.0 mg/L TDZ (data not shown). Thus we further optimized the protocol by evaluating
four different concentrations of NAA in the SI medium and three different concentrations of 6BA in the SR medium. Each NAA treatment in the SI medium was followed by each 6BA treatment in the SR medium, for a total of 24 separate combinations of SI/SR medium (Table 3-4). Shoot regeneration from whole leaf explants was observed on all SI media containing 1.0 mg/L NAA (Table 3-4). Regeneration from explants induced with SI media containing 1.0 mg/L NAA was not significantly different than media containing 0.5 or 1.5 mg/L NAA. However, 1.75 mg/L NAA in the SI media produced significantly fewer regenerating explants (P<0.05). The SR media supplemented with 1.5 or 2.0 mg/L

Table 3-4: Direct organogenesis of *Aristolochia fimbriata* from whole leaf explants. The shoot induction medium (SI) contained 0.5 mg/L 6BA and 1.0 mg/L TDZ in addition to the NAA concentrations evaluated. The shoot regeneration medium (SR) contained 1.0 mg/L NAA in addition to the 6BA concentrations evaluated.

<table>
<thead>
<tr>
<th>Induction medium</th>
<th>Regeneration medium</th>
<th>% Explants regenerating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>NAA (mg/L)</em></td>
<td><em>6BA (mg/L)</em></td>
</tr>
<tr>
<td>0.50</td>
<td>1.50</td>
<td>20 4 12</td>
</tr>
<tr>
<td>0.50</td>
<td>1.75</td>
<td>0 0 0</td>
</tr>
<tr>
<td>0.50</td>
<td>2.00</td>
<td>16 4 10</td>
</tr>
<tr>
<td>1.00</td>
<td>1.50</td>
<td>20 8 14</td>
</tr>
<tr>
<td>1.00</td>
<td>1.75</td>
<td>36 36 36</td>
</tr>
<tr>
<td>1.00</td>
<td>2.00</td>
<td>12 4 8</td>
</tr>
<tr>
<td>1.50</td>
<td>1.50</td>
<td>0 0 0</td>
</tr>
<tr>
<td>1.50</td>
<td>1.75</td>
<td>8 20 14</td>
</tr>
<tr>
<td>1.50</td>
<td>2.00</td>
<td>4 16 10</td>
</tr>
<tr>
<td>1.75</td>
<td>1.50</td>
<td>0 0 0</td>
</tr>
<tr>
<td>1.75</td>
<td>1.75</td>
<td>8 0 4</td>
</tr>
<tr>
<td>1.75</td>
<td>2.00</td>
<td>0 0 0</td>
</tr>
</tbody>
</table>

*Binary logistic regression factors with a significant effect on regeneration (P<0.05). The overall best combination of media is highlighted in bold.
6BA produced lower number of regenerating explants than 1.75 mg/L 6BA (P<0.05). Overall, the greatest percentage of regeneration (36%) occurred in explants initiated on SI medium containing 1.0 mg/L NAA, followed by regeneration on SR media containing 1.75 mg/L 6BA (Table 3-4). The results of binary logistic regression analysis indicated that the amounts of NAA in the SI media and 6BA in the SR media each independently had a significant effect on regeneration. Genotype and interactions were not significant factors in the statistical model. Close observation of whole leaf explants on SR medium indicated that shoot primordia originated only from the petiole region (Fig. 3-1C). Shoots elongated and developed 2-3 leaves between 28 and 56 days on SR medium (Fig. 3-1D).

To maximize the efficiency of the protocol, we evaluated the regeneration potential of small sections of different vegetative organs. Root and leaf blade sections, and sections containing the base of the leaf without the petiole did not regenerate. Regeneration from leaf bases with the petiole attached (data not shown) was comparable to that of the whole leaf explant (36%, Table 3-4). Stem and petiole sections regenerated more successfully (up to 67.5 %), with minimal callus formation (Table 3-5). Additionally, stem and petiole explants offered the substantial advantage of reducing the space required by the regeneration system by 80%.

The culture time of stem and petiole explants on SI medium was studied in two separate experiments where incubation periods of 14, 21 and 40 days were evaluated. For this experiment explants were collected from plants cultured on REN1 under high-light conditions. The results indicated that the percentage of regenerating explants was the
highest for explants cultured on SI medium for 14 days only (Table 3-5). The results of the first experiment demonstrated that explants cultured on SI medium for 21 days produced a significantly lower percentage of regenerating explants compared to the explants cultured on SI medium for 14 days (P<0.05). However, there was no significant difference observed between the 14-day and 21-day treatments in the second experiment. Explants cultured on SI medium for 40 days in both experiments produced regenerating explants with significantly lower success (Table 3-5). Statistical analysis of the data
indicated that stem explants were significantly less likely (P < 0.05) to regenerate than petiole explants. The tissue type and number of days on SI (14, 21, 40) were significant factors (P < 0.05) in the statistical model. Interactions and genotype were not significant factors. Overall, the best regeneration occurred from petiole explants induced on SI medium for 14 days (Table 3-5).

The regeneration of stem and petiole explants was further evaluated with explants collected from plants maintained on REN2 medium and under medium-light conditions. In this experiment, 60% of petiole explants regenerated shoots, which was less than the 80 - 97% regeneration recorded for stem explants (Table 3-6). Both petiole and stem explants regenerated at higher frequency than the maximum regeneration recorded for whole leaf, stem and petiole explants collected from plants maintained in vitro in high-light conditions (Tables 3-4, 3-5). Furthermore, on average, each stem or petiole explant in the last experiment produced 3 shoots per explant (Fig. 3-1E and F),

Table 3-6: Shoot regeneration, rooting and acclimation of plants regenerated from petiole and stem explants from plants maintained in REN2 medium.

<table>
<thead>
<tr>
<th>Experiment number</th>
<th>Tissue type</th>
<th>Total # explants</th>
<th>Explants regenerating (%)</th>
<th>Total # shoots</th>
<th>Rooted shoots (%)</th>
<th>Acclimated plant (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>petiole</td>
<td>10</td>
<td>60</td>
<td>21</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>stem</td>
<td>10</td>
<td>80</td>
<td>69</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>stem</td>
<td>30</td>
<td>97</td>
<td>90</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>stem</td>
<td>18</td>
<td>94</td>
<td>46</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

with the highest number of shoots per explant (7) recorded for stem explants (Table 3-6).

All of the newly regenerated shoots uniformly elongated and developed roots on REN2 medium approximately four weeks after transfer (Table 3-6, Fig. 3-1G). Plants with at
least two roots and minimum 3 cm total root length were transferred to soil and acclimated to greenhouse conditions (Fig. 3-1H). Acclimation to greenhouse conditions was 100% successful (Table 3-6).

Discussion

Aristolochiaceae occupies an important phylogenetic position for research on evolution of plant development. Aristolochia species also produce many secondary metabolites with valuable medicinal properties and have interesting morphological, and biochemical features important for different insects (Rulik et al., 2008). This manuscript describes the first successful protocol for micropropagation, regeneration, in vitro rooting and acclimation of two genotypes of Aristolochia fimbriata (Fig. 3-2).

During our initial attempts to regenerate shoots from Aristolochia fimbriata leaf explants, we evaluated shoot organogenesis methods reported for Aristolochia indica (Manjula et al., 1997; Remashree et al., 1997; Soniya and Sujitha, 2006), tobacco, and apple (Fisher and Guiltinan, 1995; Maximova et al., 1998). The higher levels of cytokinin applied under these protocols caused the tissue to blacken and die in 2-3 weeks. Thus, the concentration of 6BA was reduced in the regeneration media, while the concentration of TDZ (a synthetic plant growth regulator that has both auxin and cytokinin effects) remained at 1 mg/l as applied for regeneration from apple leaf tissue (Maximova et al., 1998; Murthy et al., 1998). Additionally, the results from our study indicated that the light intensity applied during the micropropagation of A. fimbriata had an influence on the shoot and plant growth during micropropagation, and also on shoot organogenesis.
from stem and petiole explants. The initial high-light micropropagation protocol produced plants closely resembling greenhouse plants with respect to compact form, leaf size, and root development. The plants propagated under the medium-light developed slightly elongated stems with smaller, but darker-green leaves compared to the high-light plants. The cultures under high-light needed to be transferred to fresh medium every 3 weeks due to leaf yellowing, while the plants under medium-light and increased MS concentration required transfers every 4-5 weeks. Moreover stem and petiole explants from plants under medium-light regenerated more uniformly and with greater success than the explants from the high-light plants. The variation observed in the shoot regeneration response could be explained with different “preconditioning” of the tissues under the different light condition and increased basal medium concentration (Mohamed et al., 1992). Plant tissues could be developing different cell sizes or have different photosynthesis rates, or the effects could be due to changes in the nutrient metabolism or endogenous hormone production (Saebo et al., 1995; Husaini and Abdin, 2007; Molina et al., 2007; Tabatabaei et al., 2008). Further analysis of the tissues from plants propagated under the different micropropagation regimes is necessary to reveal the underlying physiological differences influencing organogenesis.

Conclusions

The high frequency protocols reported here for in vitro propagation of *A. fimbriata* could provide a large quantity of greenhouse ready clonal material in a period as short as three months, with minimal space requirements. Furthermore, these protocols
are currently used for the development of a genetic transformation system for *A. fimbriata* VL genotype. The ability to study gene function *in vivo* in this species will complete an important requirement for the establishment of *A. fimbriata* as a model system for basal angiosperms. To further support these efforts, our group has generated over 30,000 EST sequences from traditional cDNA libraries constructed of whole plant tissues and pre-meiotic flower buds of *A. fimbriata* (B. J. Bliss, L. Landherr, Y. Hu, S. Clifton, J.H. Leebens-Mack, H. Ma, and C. W. dePamphilis, unpublished data) and greatly expanded transcriptome sequencing is underway as part of the Ancestral Angiosperm Genome Project (NSF DEB-0638595). The addition of *A. fimbriata* to the increasing selection of plant model systems will provide insight into the immediate progenitor of monocots and eudicots, the two largest and most successful major angiosperms clades.

**Acknowledgements**

This work was supported by National Science Foundation grants to C. dePamphilis and H. Ma (DBI-0115684 and DBI-0638595) and to M. Guiltinan (NSF 430-47/60A), a DOE grant to H. Ma (DE-FG02-02ER15332), and by the Department of Biology and Huck Institute of Life Sciences of the Pennsylvania State University. We thank M. Guiltinan for providing the tissue culture lab and growth facility space, and for editing this manuscript. We also thank L. Rosen and Jardín Botánico, Universidade de Coimbra for providing seeds; Anthony Omeis for plant care; Brett Shook, Laura Warg, and Paula Ralph for assistance with tissue culture experiments; Guanfang Wang, Zhe
Chen, and Yan Zhang for statistical support, and Dr. Stefan Wanke for valuable discussion. Our initial efforts in developing a micropropagation system for *Aristolochia fimbriata* were aided by the unpublished findings of C. Bravo, G. Yormann, and B. Llorente, recorded in Acta Horticulturae conference proceedings (1999).
Chapter 4

TCP gene family evolution

This chapter contains data, analyses, and plans for research that may be submitted for publication by Barbara Bliss.¹

¹Department of Biology, Institute of Molecular Evolutionary Genetics, and the Huck Institutes of the Life Sciences, 201 Life Sciences Building, Pennsylvania State University, University Park, PA 16802, USA
Founding members: *TB1, CYC, and PCF*

The TCP family of transcription factors was named for the initially recognized members, *TEOSINTE-BRANCHED1* (*TB1*, from *Zea mays*), *CYCLOIDEA* (*CYC* from *Antirrhinum majus*) and *PROLIFERATING CELL FACTOR* genes (*PCFs* from *Oryza sativa ssp. japonica*) (Cubas et al., 1999a). TCP genes participate in cell cycle regulation and modulating organ outgrowth, affecting the development of plant architecture, floral form, leaf morphology, and other aspects of plant development. Existing studies have examined TCP gene family evolution using sequences from the fully sequenced genomes of *Arabidopsis* and *Oryza* (Howarth and Donoghue, 2006; Yao et al., 2007), and traced TCP genes back to the earliest land plant lineage (Navaud et al., 2007). Recent work on the role of TCP genes in the development of form in monocots and core eudicots suggests they play a general role in the evolution of form in flowering plants. An understanding of TCP sequence evolution in the context of recently available basal angiosperm sequence data will support future studies of the evolution of function of this important gene family.

A review of the published literature concerning TCP gene function in flowering plants is followed by my preliminary work to understand how molecular evolution has led to lineage-specific traits under the influence of TCP family genes.

The *TEOSINTE-BRANCHED1* (*TB1*) was identified as the locus responsible for the existing *tb1-ref* maize loss of function mutant (Burnham, 1959). The *tb1-ref* mutant produced long lateral branches with tassels (male inflorescences) at some upper nodes, and tillers (axillary branches) at the basal nodes; a plant architecture resembling that of teosinte, the wild Mexican grass (Doebley and Stec, 1991). Modern *Zea mays* produces
reduced lateral branches that terminate without branching in ears (female inflorescences). Quantitative trait loci (QTL) analysis identified the TB1 locus as one of five loci contributing to the remarkable change in plant and inflorescence architecture between Zea mays ssp. mays and its teosinte (Zea mays ssp. parviglumis) progenitor (Doebley et al., 1995) that facilitated the transition to modern agriculture.

CYCLOIDEA (CYC) was found in a semi-peloric mutant of Antirrhinum majus and was the first gene described affecting flower symmetry (Luo et al., 1996). The Antirrhinum flower has a single dorsoventral plane of symmetry, with one distinct ventral (abaxial) petal, two lateral petals, and two larger dorsal (adaxial) petals. The adaxial stamen is aborted, leaving only four stamens at maturity. Complete peloric mutants have radially symmetric flowers and can occur in natural populations of zygomorphic species, or can be induced. Semi-peloric Antirrhinum mutants lacking a functional CYCLOIDEA (CYC) gene exhibited increased radial symmetry because the lateral, and, to some extent, dorsal organs, acquired a ventralized form (Luo et al., 1996). CYC mutants also produced an extra floral organ in the sepal, petal, and stamen whorls (Luo et al., 1996).

The third members of the TCP family, PROLIFERATING CELL FACTOR1 (PCF1) and PCF2 are required for activating proliferating cell nuclear antigen (PCNA) in Oryza sativa ssp. japonica (Kosugi and Ohashi, 1997). PCNA is essential for regulating proliferation and cell cycle progression during the transition from G1-to-S phase, exclusively in meristematic tissues (Waga et al., 1994; Kosugi and Ohashi, 1997; Warbrick, 2000). PCF1 is expressed in all tissues examined, but PCF2 was found only in meristematic tissue where PCNA was also expressed (Kosugi and Ohashi, 1997). PCF1
and PCF2 may act by formation of homo- or heterodimers, activating PCNA expression only in tissues where all are expressed (Kosugi and Ohashi, 2002).

**TCP gene functions**

The TCP family transcription factors act generally to promote elaboration of form (PCF1 and PCF2) or to repress it (CYC and TB1) by regulating the cell cycle. These genes have been collected into two broad classes of TCP proteins based on amino acid similarity in the TCP transcription factor domain, which is composed of approximately 221 bases (Cubas et al., 1999a). The TCP domain contains a unique basic helix loop helix (bHLH) DNA binding motif, which has a different structure and DNA-binding specificity than does the bHLH domain found in MyoD (Murre et al., 1989a; Murre et al., 1989b; Weintraub et al., 1991; Cubas et al., 1999a). The class I proteins contain the proliferating cell factor proteins. The class II proteins have been subdivided into class II and class III, with class II remaining the designation for the clade containing CYC and TB1, and referred to by some as the ECE clade (Howarth and Donoghue, 2006; Koyama et al., 2007; Navaud et al., 2007; Yao et al., 2007).

**Class I TCP proteins**

In *Arabidopsis*, the class I protein TCP20 interacts with the *Arabidopsis PCNA2* promoter in a manner similar to that of PCF2 with *Oryza PCNA*. Both PCNA2 and PCNA contain a telo box element, and a cis-acting site II motif, although the bases differ slightly
TCP20 binds the $PCNA2$ site II motif, which is required and sufficient to activate $PCNA2$ expression in cycling cells in meristematic tissue (Tremousaygue et al., 1999; Tremousaygue et al., 2003). TCP20 also interacts with AtPura, a protein that binds the telo box on $PCNA2$, and may stimulate transcription by opening chromatin and interacting with transactivators, recruiting them for further control of cell cycle activities (Tremousaygue et al., 1999; Tremousaygue et al., 2003). Along with $PCNA$ and AtPura, Arabidopsis mitotic cyclin, CYCB1:1 binds TCP20 at the GCCCR motif, upregulating $CYCB1:1$ at the G2/M transition (Li et al., 2005). Many promoters of ribosomal proteins and other proteins required for cell cycling contain the site II and telo box elements (Tremousaygue et al., 2003), as do nuclear genes encoding components of the mitochondrial oxidative phosphorylation machinery which is essential for mitochondrial biogenesis prior to cell division (Welchen and Gonzalez, 2006).

**Class II TCP proteins**

Class II TCP gene expression generally represses outgrowth. $TB1$ was shown to repress branch and female inflorescence outgrowth from secondary axillary meristems where $TB1$ mRNA accumulates (Doebley et al., 1997; Hubbard et al., 2002; Howarth and Donoghue, 2006). Two homologs of $TB1$ in Arabidopsis, $BRC1/TBL1$ and $BRC2$ (formerly $TCP18$ and $TCP12$) also repress axillary outgrowth with high accumulations of mRNA in unelongated buds and low levels in elongating buds (Aguilar-Martinez et al., 2007; Finlayson, 2007). Genetic analyses indicate they respond to and integrate auxin and
the MORE AXILLARY GROWTH (MAX) signals to control bud outgrowth (Aguilar-Martinez et al., 2007; Finlayson, 2007). In Sorghum, increased exposure to far red light induces a shade avoidance syndrome, accompanied by reduced levels of phytochrome B, reduced axillary outgrowth, and increased expression of TB1, suggesting TCP genes play a role in producing architectural responses of plants to light in the environment (Kebrom et al., 2006)

CYC, another class II TCP gene, has been shown to regulate cyclins D1, D3a, and D3b (Gaudin et al., 2000). The D-cyclins are induced by cytokinins, and act to induce cell division. CYC specifically represses cyclin D3b in the dorsal stamen primordia, which are aborted in the mature wild-type flower, consistent with termination of cell division activity (Riou-Khamlichi et al., 1999; Gaudin et al., 2000). TB1 and related genes may produce their effects by similar interactions with cyclins.

In Arabidopsis, TCP1, the most likely CYC ortholog, is expressed early in the adaxial (dorsal) part of the floral meristem (as CYC is in Antirrhinum), but expression in Arabidopsis is not maintained later in development. The difference in timing may account for the polysymmetry of the mature Arabidopsis flower (Cubas et al., 2001). In Brassicaceae, the CYC homolog laTCP1 from Iberis amara is expressed differentially across the corolla when the greatest petal growth occurs, and produces a wild-type monosymmetric perianth (Busch and Zachgo, 2007). TCP1 (Arabidopsis) and CYC (Antirrhinum) are also expressed in the adaxial regions of all axillary shoot meristems in Arabidopsis, and in axillary shoot meristems immediately below the oldest floral meristem in Antirrhinum (Cubas et al., 2001; Clark and Coen, 2002).
In *Solanum tuberosum*, transcripts of the class II protein, *SSTCP1*, are concentrated in dormant buds, which are typically arrested in G1. In sprouting buds, which consist of cycling cells, the transcripts of *SSTCP1* are not detectable (Faivre-Rampant et al., 2004).

*CYC* genes have been implicated in floral form in all lineages investigated, including the bisymmetric flowers of the basal eudicots (Papaveraceae; Ranunculales) (Damerval et al., 2007; Damerval and Nadot, 2007). Duplications in class II TCP family genes have provided raw materials for neo- and sub-functionalization, supporting floral form divergence (Ohno, 1970). The extent to which class II TCP genes have been involved in independent origins and losses of monosymmetry has been investigated in large families with highly derived and canalized floral form (Orchidaceae, Fabaceae, Asteridae, Dipsacaceae) (Cubas, 2004) as well as in several species in the same family and order as *Antirrhinum* (Plantaginaceae, Lamiales) (Vieira et al., 1999; Gubitz et al., 2003; Hileman and Baum, 2003). These analyses indicate a duplication in the *CYC* lineage occurred prior to the divergence of the Antirrhinae to generate the *CYC/DICH* paralogs, which are now under purifying selection (Hileman and Baum, 2003).

In *Antirrhinum*, a second locus *DICHOTOMA (DICH)*, is responsible for a semi-peloric mutant in which organ number is not affected but dorsal petal morphology is ventralized (Luo et al., 1996). The sequence of *DICHOTOMA* is 66% identical to that of *CYCLOIDEA*. Mutations in both *CYC* and *DICH* were required to produce a true peloric phenotype, with radially symmetric, fully ventralized petals (Luo et al., 1999). In *Mojavea confertifolia*, a species closely related to *Antirrhinum*, expression of *CYC* and *DICH* orthologs differed from those of *Antirrhinum majus* in a manner correlated with
differences in floral form between the two species, particularly with respect to additional stamen abortion (Hileman and Baum, 2003). DICH orthologs have not been found outside the Antirrhinaceae.

In the composite inflorescence of *Gerbera*, the expression gradient of a CYC homolog across the inflorescence establishes differences in symmetry and petal fusion associated with flower type (ray, trans, disc) (Broholm et al., 2008).

In Adoxaceae and Caprifoliaceae, in the Dipscales (Campanulids clade of asterids), within the class II (ECE genes), three clades of CYC-like genes were identified (CYC1, CYC2, CYC3), suggesting two duplications in lineages leading to the Caprifoliaceae (Howarth and Donoghue, 2005). More CYCLOIDEA orthologs have been found in other orders and families within Asteridae (Calceolariaceae, Scrophulariaceae, and Oleaceae in Lamiales; Rubiaceae in Gentianales; Solanaceae and Boraginaceae in Solanales), consistent with evidence of genome duplications early in the asterid lineage (Reeves and Olmstead, 2003).

However, in Gesneriaceae, another predominantly zygomorphic family in Lamiales, several independent reversals to radial symmetry have taken place, generating actinomorphic species. Amplification of genomic DNA revealed only one clear CYC–like gene for most members of the family (Citerne et al., 2000; Smith et al., 2004). The protein-coding regions of the CYC orthologs in the naturally occurring gesneriadiactinomorphs did not suggest that loss of function in the protein-coding sequence was responsible for reversions to actinomorphy (Citerne et al., 2000; Smith et al., 2004). Furthermore, individuals within a population displayed no visible phenotype associated with the numerous single nucleotide polymorphisms in the population, supporting the
neutral mutation hypothesis (Kimura, 1968). In addition, there was no evidence of directional selection on lineages leading to radially symmetric taxa (Smith et al., 2004). Changes in protein-protein interactions, or in gene regulation (i.e., methylation, expression in different tissues, loss of cis-regulatory elements, miRNA posttranslational inactivation) could account for such differences.

**Class III TCP proteins**

The involvement of TCP proteins in establishing lateral organ boundaries was first reported for the *Antirrhinum cupuliformis* (cup) mutant, based on interactions of CUP with a TCP domain-containing transcript. In *Arabidopsis*, the class III TCP proteins TCP3, and to a lesser extent TCP2, TCP4, TCP5, TCP10, TCP13, TCP17, TCP24 were associated with restricted expression of *CUP-SHAPED COTYLEDON* (*CUC*) genes to boundaries of organ primordia (Weir et al., 2004; Koyama et al., 2007). Repression of TCP3 induced shoot formation on cotyledons and produced many organ defects attributable to improperly specified organ boundaries, while overexpression was mitigated by the endogenous activity of miR319/JAW on TCP 3, TCP2, TCP4, TCP10, and TCP24. When the miR319/JAW target sequence on these class III transcripts was eliminated, overexpression of TCP3 or related genes produced fused cotyledons and shoot developmental defects (Koyama et al., 2007).

Class III TCP genes also play a role in leaf form. Mutations the *Antirrhinum* gene, *CINCINNATA* (*CIN*), produce reduced petal lobes and leaves with a crinkly phenotype. Cell size, and expression patterns of *CYCLIN D3B* and *HISTONE4*, indicated the cell-
cycle arrest front in the mutant leaf is delayed as it moves from leaf tip to base (Nath et al., 2003; Crawford et al., 2004). Growth of medial leaf cells is arrested in the mutant, while the marginal cells continue to proliferate (McConnell and Barton, 2003; Nath et al., 2003).

The homologs of CIN in Arabidopsis are TCP2, TCP3, TCP4, TCP5, and TCP24, each of which is regulated by miR319 (Palatnik et al., 2003). TCP4 negatively regulates leaf growth and positively regulates leaf senescence in Arabidopsis, through increasing JA biosynthesis, particularly at the oxylipin biosynthesis step of the pathway (Schommer et al., 2008). In Solanum lycopersicon, miR319 regulation of a presumably class III TCP gene is also required for development of compound leaf (Ori et al., 2007).

**Evolution by changing gene interactions**

TCP genes appear to be able to interact with a variety of other genes. In Antirrhinum, at the periphery of the floral meristem, and especially at the base of the organ primordia and in the ventral petals, CYC and DICH impose a dorsalizing effect by negatively regulating DIVARICATA (DIV), a myb family gene that confers ventral identity on wild type petals in a dosage dependent manner (Almeida et al., 1997). This effect is mediated by RADIALIS (RAD) a myb-like gene. CYC and DICH activate RAD in the early dorsal domain of the floral meristem, which antagonizes the ventralizing effect of DIV in those dorsal regions. The activation of RAD by CYC does not occur in Arabidopsis, indicating the regulatory network in which these genes are involved is lineage-specific. Although CYC does participate in the development of Arabidopsis floral
organ form, it does not utilize the same network as in *Antirrhinum* (Corley et al., 2005; Costa et al., 2005).

In the rosids, *CYCLOIDEA* orthologs have been collected from Brassicaceae (*Arabidopsis*) and from Fabaceae, another large family containing zygomorphic flowers (Citerne et al., 2000; Fukuda et al., 2003; Ree et al., 2004). In an investigation of the interactions of CYC-like genes Most recently, over- and underexpression studies in the model plant *Lotus japonicus* have demonstrated the requirement for the CYC ortholog LjCYC2 in zygomorphy. Overexpression of LjCYC2 generated adaxial cell types in the lateral and ventral petals, and dorsalized (adaxialized) form in the lateral petals. Suppression of LjCYC2 conferred abaxial shape and cell type on the lateral petals, and affected inflorescence morphology with some shorter internodes and frequent fusion of flower pairs (Feng et al., 2006). A lateralizing counterpart of RAD (*Antirrhinum*), *Keeled wing in Lotus* (*Kew*1), was also identified in *Lotus*. *Kew*1 shares a similar phenotype to the *KEELED WING* (*K*) mutant in *Pisum*, and macrosynteny with a region containing the locus responsible for the *keeled wing* (*K*) mutant in *Pisum* (Wang et al., 2008). Though the mechanism of CYC regulation is conserved (dorsalizing) between the asterid and rosid lineages, the interactions of CYC with other members of the pathway have diverged.

**Regulation by methylation**

In addition to translational control and interactions with other genes, TCP genes are regulated by transcriptional control. In maize, comparisons of the sequence diversity of the protein-coding region of the *TB1* gene and the non-transcribed regulatory region
upstream of the *TB1* gene to the same regions in teosinte indicated the regulatory region, not the protein coding region, was under selection (Wang et al., 1999). Some members of the TCP family are regulated by methylation (Cubas et al., 1999b; Feng et al., 2006), and microRNA (miRNA) (Ori et al., 2007; Palatnik et al., 2007).

In the naturally occurring peloric mutant of *Linaria vulgaris* (Plantaginaceae) radial symmetry is maintained by hypermethylation of the genomic region containing the *CYC* ortholog alone: demethylation can restore bilateral symmetry in this species (Cubas et al., 1999b). Also, in the naturally occurring allopolyploid, *Arabidopsis suecica*, TCP family genes were among the nearly 23% genes exhibiting differential expression of parental genes (Lee and Chen, 2001). As a result of the polyploidy, only the *TCP3* allele from the *A. thaliana* parent was expressed. The *TCP3* allele from the *Cardaminopsis arenosa* parent was silenced by methylation, and could be reactivated by demethylation (Lee and Chen, 2001). *TCP3* is a class III TCP gene (Navaud et al., 2007; Yao et al., 2007).

The dePamphilis lab has an ongoing commitment to provide informatics tools to facilitate the use and analysis of millions of ESTs generated by genomics projects. My contribution has primarily been to test and provide feedback regarding features under development, initially with curation of TCP and auxin response factor gene family members, using the LIMS and Phylomine web interfaces for FGP under the direction of Jim Leebens-Mack. Tools we initially conceived of for gene family curation and analysis are still under development, but have ample functionality to apply to the evolution of the TCP gene family. Here, I have evaluated the utility of Tribes and an automated gene family analysis pipeline developed as a component of the FGP/AAGP bioinformatics
resource, to collect sequences, produce an alignment, and to generate phylogenetic trees using maximum likelihood methods.
Materials and Methods

Automated pipeline

All sequences from the fully sequenced genomes of seven species (*Populus trichocarpa* (Poptr), *Vitis vinifera* (Vitvi), *Sorghum bicolor* (Sorbi), *Carica papaya* (Carpa), *Medicago truncatula* (Medtr), *Arabidopsis thaliana* (Arath), *Oryza sativa* ssp. *japonica* (Orysa), *Physcomitrella patens* (Phypa), and *Selaginella moellendorfii* (Selmo) were collected into Plant Tribes (Wall et al., 2008). The sequences in the Supertribes were collected and aligned with MUSCLE (Edgar, 2004) and evaluated using RAxML (Random accelerated maximum likelihood) (Stamatakis, 2006).

Verification of sequences

The Tribe collection and alignment were evaluated to verify that all known TCP domain-containing sequences from *Arabidopsis* and *Oryza* were included (Table 4-1)) and that all amino acid sequences from other taxa also contained the TCP domain (PFAM domain 03634, http://www.sanger.ac.uk/Software/Pfam/search.shtml). Sequences having fewer than 50 characters of the TCP domain were eliminated from the alignment.
Table 4-1: *Arabidopsis thaliana* (At) and *Oryza sativa* ssp. *japonica* (Os) TCP genes from Yao et al. (2007) and Navaud et al. (2007).

<table>
<thead>
<tr>
<th></th>
<th>Gene</th>
<th>LOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>At4g18390</td>
<td>LOC_Os01g11550</td>
</tr>
<tr>
<td>2</td>
<td>At1g53230</td>
<td>LOC_Os01g55750</td>
</tr>
<tr>
<td>3</td>
<td>At3g15030</td>
<td>LOC_Os01g69980</td>
</tr>
<tr>
<td>4</td>
<td>At5g60970</td>
<td>LOC_Os02g42380</td>
</tr>
<tr>
<td>5</td>
<td>At5g41030</td>
<td>LOC_Os02g51280</td>
</tr>
<tr>
<td>6</td>
<td>At5g23280</td>
<td>LOC_Os02g51310</td>
</tr>
<tr>
<td>7</td>
<td>At1g58100</td>
<td>LOC_Os03g49880</td>
</tr>
<tr>
<td>8</td>
<td>At2g45680</td>
<td>LOC_Os03g57190</td>
</tr>
<tr>
<td>9</td>
<td>At2g31070</td>
<td>LOC_Os04g11830</td>
</tr>
<tr>
<td>10</td>
<td>At2g37000</td>
<td>LOC_Os04g44440</td>
</tr>
<tr>
<td>11</td>
<td>At1g68800</td>
<td>LOC_Os05g43760</td>
</tr>
<tr>
<td>12</td>
<td>At3g02150</td>
<td>LOC_Os06g12230</td>
</tr>
<tr>
<td>13</td>
<td>At3g47620</td>
<td>LOC_Os07g04510</td>
</tr>
<tr>
<td>14</td>
<td>At1g69690</td>
<td>LOC_Os07g05720</td>
</tr>
<tr>
<td>15</td>
<td>At3g45150</td>
<td>LOC_Os08g33530</td>
</tr>
<tr>
<td>16</td>
<td>At5g08070</td>
<td>LOC_Os08g43160</td>
</tr>
<tr>
<td>17</td>
<td>At3g18550</td>
<td>LOC_Os09g24480</td>
</tr>
<tr>
<td>18</td>
<td>At5g51910</td>
<td>LOC_Os09g34950</td>
</tr>
<tr>
<td>19</td>
<td>At3g27010</td>
<td>LOC_Os11g07460</td>
</tr>
<tr>
<td>20</td>
<td>At5g08330</td>
<td>LOC_Os12g02090</td>
</tr>
<tr>
<td>21</td>
<td>At1g72010</td>
<td>LOC_Os12g07480</td>
</tr>
<tr>
<td>22</td>
<td>At1g35560</td>
<td>LOC_Os12g42190</td>
</tr>
<tr>
<td>23</td>
<td>At1g30210</td>
<td></td>
</tr>
</tbody>
</table>
Results

Sequence collection

Tribes returned five supertribes, of which two had no members with the TCP domain. The remaining three supertribes, each contained 4-7 tribes (Table 4-2). A subset of these contained the PFAM domain of interest (Table 4-3).

Table 4-2: Three supertribes of genomic sequences comprise Arabidopsis and rice genes of interest.

<table>
<thead>
<tr>
<th>SUP</th>
<th>TRIBE</th>
<th>Phypa</th>
<th>Selmo</th>
<th>Orysa</th>
<th>Sorbi</th>
<th>Vitvi</th>
<th>Poptr</th>
<th>Medtr</th>
<th>Carpa</th>
<th>Arath</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>777</td>
<td>2034</td>
<td></td>
<td></td>
<td>19</td>
<td></td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>777</td>
<td>8345</td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>777</td>
<td>9164</td>
<td>1</td>
<td></td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>777</td>
<td>20097</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>777</td>
<td>20098</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>777</td>
<td>30671</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>777</td>
<td>79451</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1148</td>
<td>767</td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>47</td>
</tr>
<tr>
<td>1148</td>
<td>3418</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1148</td>
<td>20099</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1148</td>
<td>20826</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1148</td>
<td>29237</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1148</td>
<td>71899</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2066</td>
<td>282</td>
<td>4</td>
<td>3</td>
<td>11</td>
<td>18</td>
<td>11</td>
<td>20</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>98</td>
</tr>
<tr>
<td>2066</td>
<td>14235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2066</td>
<td>29203</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2066</td>
<td>59121</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>6</td>
<td>6</td>
<td>24</td>
<td>29</td>
<td>44</td>
<td>35</td>
<td>7</td>
<td>28</td>
<td>24</td>
<td>203</td>
</tr>
</tbody>
</table>

In supertribe 777, only one of 19 Vitis sequences in tribe 2034 (Table 4-2) contained the TCP domain (Table 4-3). Neither of the Vitis sequences in tribe 8345 or
79451 contained the TCP domain, and the one *Vitis* sequence in tribe 9164 contained a TCP domain. The remaining sequences in tribe 9164 all contained TCP domains (Table 4-3). Besides the sequences in tribe 9164 the *Carica* sequences in tribe 20098 had a TCP domain. None of the *Carica* sequences in tribe 2034 or 30671 contained the TCP domain.

In supertribe 1148, all members of tribes 767 and 3418 contained TCP domains, but tribes 20826, 29237, and 29237 had none (Figure 4-2). In supertribe 2066, all members of tribe 282 contained TCP domains, but tribes 14235, 29203, and 59121 had none (Figure 4-3).

Table 4-3: Genomic sequences containing TCP domain (PFAM domain 03634).

<table>
<thead>
<tr>
<th>SUP</th>
<th>TRIBE</th>
<th>Phypa</th>
<th>Selmo</th>
<th>Orysa</th>
<th>Sorbi</th>
<th>Vitvi</th>
<th>Poptr</th>
<th>Medtr</th>
<th>Carpa</th>
<th>Arath</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>777</td>
<td>2034</td>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>777</td>
<td>8345</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>777</td>
<td>9164</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>777</td>
<td>20097</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>777</td>
<td>20098</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>777</td>
<td>30671</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>777</td>
<td>79451</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1148</td>
<td>767</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>46</td>
</tr>
<tr>
<td>1148</td>
<td>3418</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>1148</td>
<td>20099</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>1148</td>
<td>20826</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1148</td>
<td>29237</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1148</td>
<td>71899</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2066</td>
<td>282</td>
<td>3</td>
<td>3</td>
<td>11</td>
<td>18</td>
<td>11</td>
<td>11</td>
<td>20</td>
<td>5</td>
<td>13</td>
<td>21</td>
</tr>
<tr>
<td>2066</td>
<td>14235</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2066</td>
<td>29203</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2066</td>
<td>59121</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Totals</td>
<td>4</td>
<td>6</td>
<td>22</td>
<td>29</td>
<td>19</td>
<td>34</td>
<td>7</td>
<td>21</td>
<td>24</td>
<td>166</td>
<td></td>
</tr>
</tbody>
</table>
An additional *Arabidopsis* gene, At1g67620 was returned in the tribes. This gene had been omitted from recently published studies (Navaud et al., 2007; Yao et al., 2007) and had once been identified as a pseudogene in Genbank. However, functional data for At1g67620 have been available for some time. Altogether, the 24 *Arabidopsis* genes matched those originally described (Cubas et al., 1999a; Cubas et al., 2001).

**Alignments**

The alignment produced by the automated pipeline used default values of MUSCLE and located the TCP domain in those sequences having a TCP domain, but many sequences lacked it (Figure 4-1, 4-2, 4-3).

**Phylogenetic analysis**

The phylogenetic tree generated by RAxML (Figure 4-4) included only sequences containing a TCP domain. Clades corresponding to classes 1-3 as found by Yao (2007) and Navaud (2007) (Figure 4-5, 4-6) could be identified.
Figure 4-1: Supertribe 777, MUSCLE alignment of amino acids.
Figure 4-2: Supertribe 1148, MUSCLE alignment of amino acids.
Figure 4-3: Supertribe 2066, MUSCLE alignment of amino acids.
Figure 4-4: RaxML phylogenetic analysis of TCP domain-containing sequences from automated pipeline. Pink, blue, and green indicate class 1, 2, and 3 genes, respectively.
**Discussion**

An automated pipeline using Tribes was evaluated for its ability to replicate or improve upon the findings of Yao et al. and Navaud et al. (2007; Yao et al., 2007) ([Figure 4-5, 4-6]). A number of sequences, particularly from *Vitis*, were automatically included in the alignment although they had no TCP domain. *Arabidopsis* class 2 and 3 genes were distributed in supertribes 777 and 1148, class 1 genes were in supertribe 2066. Supertribe 777 was dominated by *Vitis* sequences. Some similarity, other than that of a TCP domain, caused additional sequences to be collected by Tribes. I suspect that that similarity was a consequence of the disproportional representation of *Vitis* sequences in the database.
Figure 4-5: Neighbor joining tree and expression pattern of TCP genes in rice and Arabidopsis from Yao et al. (2007).
The alignment of the TCP domain distinguished between classes I and II but did not produce monophyletic lineages as per the reference phylogenies (Figure 4-5, 4-6). The HMM model for the TCP gene family reveals several regions of more or less conservation (Figure 4-8). TCP genes contain many informative motifs, as reported by previous researchers. With the abundance of sequence information that is available, it should be possible to align and investigate the molecular evolution of motifs including, the miR319 target sequence in the class III genes. It may be possible to identify selection, or perhaps loss of constraint in a lineage in which the functionality of the TCP gene has been lost (e.g., class II genes in Gesneriaceae). A complete sequence analysis may suggest new gene interactions. Functional analyses of TCP genes in the basal angiosperm, *A. fimbriata* will provide important insight into the early role of this gene family in flowering plants, and advance our understanding of the evolution of symmetry and form in angiosperms.
Figure 4-7: Classes 1, 2, 3 TCP sequences. Pink indicates Class 1 (PCF-like) sequences, blue indicates Class 2 sequences, green indicates Class 3 sequences.
Figure 4-8: Class 1 (PCF-like) TCP sequences. Pink indicates Class 1 (PCF-like) sequences, blue indicates Class 3 sequences.
Figure 4-9: HMM logo PFAM03634 (http://pfam.sanger.ac.uk/family?entry=TCP).
Chapter 5
Retrospective and future direction for studies of evolution of development in basal angiosperms

Basal angiosperm lineages contain relatively few extant species, but they contain more diversity than in all other angiosperm lineages combined. Basal angiosperms offer a direct view of the common ancestor of monocots and eudicots, as well as unique insights into the ancestor of all angiosperms. The magnoliids contain the most species and much of the diversity found in basal angiosperm lineages, hence three species from that lineage were selected by the Floral Genome Project (FGP) for floral transcriptome sequencing aimed at identifying components of genetic diversity responsible for variations in floral diversity (Soltis et al., 2002). The three magnoliid species included two tree species (Liriodendron tulipfera and Persea americana), and an herbaceous perennial (Saruma henryi). Liriodendron and Persea have two floral whorls of three tepals, and the flower of Saruma has two well-differentiated outer whorls that appear to be sepals and petals. However, the petal whorl of Saruma is known to arise from tissue that also gives rise to stamens (de Craene et al., 2003). Thus, the Saruma flower suggests an alternate origin of petals from those found in core eudicots. Because of this interesting origin of the petal whorl, and because Saruma is herbaceous, it seemed like a potential candidate for a basal angiosperm model system needed to address hypotheses of evolution of development (Baum et al., 2002).
I joined the FGP and applied my horticultural training and experience to obtaining and cultivate over one hundred *Saruma henryi* stock plants needed for generating sufficient tissue for the flower bud cDNA library. I worked with the PSU Plant Pathology Dept. to identify and address existing problems of crop failure attributed to fungal pathogens, then with Tony Omeis to develop and implement prophylactic crop care protocols. I consulted with Andy Stephenson and the PSU Horticulture Dept. and developed vernalization methods to insure continuous availability of source tissue for the FGP consortium. I collected the tissue needed for the flower bud cDNA library, and for subsequent FGP needs. Plants were obtained from the National Arboretum, Heronswood Nurseries, Seneca Hill Perennials, and Sunshine Farms, with the greatest contribution of tissue from the Heronswood plants. I sequenced the internal transcribed spacer region from each population and all were essentially identical (Soltis and Kuzoff, 1993). FGP generated 10,274 EST’s from *Saruma*, the most deeply sequenced magnoliid species.

Altogether, FGP generated over 100,000 floral bud cDNA sequences from fifteen angiosperms and gymnosperms, filling critical gaps in sequence and expression data needed to provide perspective on gene family evolution in the MADS-box family, as well as others. FGP provided many first glimpses into basal angiosperm sequence and expression data that continue to generate hypotheses about evolution of angiosperm development, which can only be tested in a basal angiosperm plant model. Our cultivation experience with *Saruma*, including its extended time to flower, subsequent requirement for vernalization, difficulty propagating, and genome size evidence pointing to polyploidy (Bharathan et al., 1994) demonstrated its limited utility as a plant experimental model. Therefore, I surveyed all *Aristolochia* species available from
commercial seed catalogs, nurseries, public botanical gardens, private collectors, and
index semina, using a suite of criteria to select the best candidate for development as a
basal angiosperm experimental model (Bliss et al., in preparation, 2008). Leaf explants
from A. fimbriata and Saruma were easily transformed, but the Saruma tissues evidenced
no capacity for regeneration. Micropropagation and regeneration protocols for A.
fimbriata were developed to support the transformation system (Bliss et al., submitted,
December 2008).

**Improving the Aristolochia fimbriata transformation system**

Explant type, concentration of basal media salts, co-cultivation period, antibiotics,
and pre-condition of stock material have all been modified since my initial work with the
Aristolochia fimbriata transformation system. Although more work has been done with
the shoot organogenesis system, experiments with somatic embryogenesis were also
promising. Production of transgenic lines via somatic embryogenesis would offer the
advantage of generating non-chimeric secondary somatic embryos which can be
germinated to produce transgenic plants (Dandekar et al., 1989). The shoot organogenesis
system yielded results more quickly, but has the inherent tendency to produce chimeras,
which can be eliminated by improved selection to eliminate non-transformed cells.
However, at the present time the main bottleneck in the shoot organogenesis system is
not the production of chimeras, but in getting the transformed tissues to regenerate.
Problems in the water supply are suspected.
In the shoot organogenesis system, reduced light in micropropagation of source tissue generated significant differences in regeneration of stem and petiole tissues, the physiological basis for which is not understood, but is not unique to *A. fimbriata* (for discussion, see Bliss et al. (submitted, December 2008)). Characterizing the endogenous hormone production, including describing the molecular structures and production levels in plant tissues, would inform efforts to optimize the *A. fimbriata* transformation system and contribute to the basic science of plant growth regulation in lineages preceding the divergence of monocots from eudicots. The highest level of regeneration (indicated by primordia formation) occurred in petiole explants sliced open lengthwise (Bliss et al., submitted, December 2008), so a transformation protocol using petiole explants may also display better regeneration. Ongoing problems achieving elongation of primordia on the elongation medium have been relieved by decreasing the length of time on the second, TDZ-free, shoot regeneration medium. Reducing NAA level in that medium may also improve primordia regeneration. The increase in MS salts that lowered the requirements for stock plant maintenance may be inhibiting production of additional buds (shoot primordia) in regenerating explants, so perhaps should be reduced in the third step of regeneration, before roots have formed (Das et al., 1998).

At present, regeneration problems in the transformation system should be addressed in several ways. The populations of transformed cells that currently survive selection are not able to regenerate in the subsequent counter selection regime. This may be attributable to low survival of transformed cells, excess survival of untransformed cells, inability of the transformed cells to regenerate under the antibiotic regime, or inability of the transformed cells to regenerate at all. Optimal selection and counter-
selection antibiotics should be identified which can quickly and completely eradicate non-transformed tissues, yet allow the transformed tissues to regenerate. Increasing the ratio of surface area to explant size will increase exposure to the selection agent, and support rapid selection. Application of a liquid medium containing high levels of antibiotics for a short period may also support rapid elimination of both *A. tumefaciens* and non-transformed cells. Reducing the co-cultivation time to 48-72 hours may also improve overall health of the tissues, and/or allow a different population of cells to survive in the explant. Another way to alter the population of growing, transformed, but non-regenerating cells is to alter the exposed cells of the explant so that a different population of cells – hopefully ones with better ability to regenerate – will be exposed and transformed. Stem and petiole explants (particularly petiole explants) that were split lengthwise regenerated more frequently than did less dissected explants (Expt. B., Table 3-5), therefore it may be beneficial to initiate transformation using explants split lengthwise.

A systematic assessment of selection antibiotics can be conducted to identify the antibiotic having the best ability to select for transformed cells. In the experiment shown (Figure 5-1), the selection agent kanamycin was not sufficient for seedling selection (Figure 5-1 B) at the concentration tested (100 mg/L). It also was not adequate for selection of transformed cells in tissue culture (data not shown). However, subtle growth effects from the selection agent were evident after several weeks (Figure 5-1 C) and the tips of seedling roots exposed to the kanamycin appeared yellow (Figure 5-1 D).
An optimal counter-selection agent should be identified through similar experiments. It should permit direct organogenesis in untransformed tissues when used at levels effective for the control of Agrobacterium. The antibiotic properties of A. fimbriata may contribute to counter selection, reducing the required level of the optimal counter-
selection agent. Rinsing explants with distilled water after co-cultivation and after selection in liquid medium may further reduce dependency on counter-selection antibiotics.

A protocol will also be needed for selecting potentially transformed seeds collected from a hemizygous parent produced from successful transformation. GFP expression in the seedling may not provide an ideal system for selection, as *A. fimbriata* seedlings require light for emergence of true leaves, and light promotes chlorophyll accumulation, which interferes with visualization of GFP. Furthermore, seedlings are not tolerant of the desiccating conditions of the microscope stage, so use of GFP to identify transformed seedlings puts at risk the survival of the transformed plant. Antibiotic selection can be used if seeds are germinated in moist paper toweling, sterilized with 10% sodium hypochlorite (bleach) solution, rinsed and inserted into freshly made medium containing antibiotics for seedling selection. Seedling germination on media was not reliable (see “No Kan” Figure 5-1 B, and Figures B1-6), and the extended time required (when germination did occur) may have allowed heat or light to degrade the antibiotics.

A different approach would be to take advantage of the ease with which the plant can be regenerated from root tissue in the greenhouse. Large roots divided into 5-9 cm sections and half covered with soil produce new shoots and generate new plants. Many large, healthy transformed roots were generated in initial transformation experiments (Figure 5-2). Transformed roots can be grown in vitro until large, then acclimated and planted in soil for shoot regeneration, in which transformation could be confirmed by PCR.
In *vivo* methods for transformation are desirable, but promising images from early experiments could not be confirmed. The PCR protocol available now to confirm transformation of *A. fimbriata* tissues will support renewed attempts at *in vivo* transformation. Efforts to improve any transformation method (but *in vivo* methods in particular) should take into account the effect of temperature on transformation (Dillen et al., 1997): co-cultivation *in planta* should be done in an appropriately cool location. In
addition to temperature, optimizing the acetylsyringone and surfactant concentrations may further improve transformation results (Sunilkumar and Rathore, 2001).

**Evolution of gene function in basal angiosperms**

The *A. fimbriata* experimental model offers opportunities for investigating features of physiology and secondary metabolite production in basal angiosperms (see (Bliss et al., in preparation, 2008). In *Arabidopsis*, the 35S-ATG-1 and 35S-ATG-34 *AG* overexpression constructs produced the *ap2* phenotype, in which sepals and petals were absent and the floral whorls displayed carpel-stamen-stamen-carpel organ identity from the outermost to inner whorls (Mizukami and Ma, 1992). Loss of *AG* function in *Arabidopsis* generates loss of meristem determinacy (more than 15 whorls of floral organs) and loss of floral organ identity (absence of stamens and carpels) in both the *ag-l* mutant and with an antisense ATG-1 construct (Mizukami and Ma, 1995).

In *A. fimbriata*, over expression of *AG* orthologs from *Saruma* or *Aristolochia* (Kramer et al., 2004; Zahn et al., 2006) would be expected to replace the unipartite perianth with ectopic stamens and carpels, but it is not clear how that would be arranged given the wild-type inferior ovary and stamen adnation with the style in the gynostemium structure. Loss of *AG* function would be expected to eliminate the gynostemium, but it is not clear what perianth-like structures would take its place, or to what extent meristem identity is a C-function role in *A. fimbriata*. Jaramillo and Kramer (2004) reported orthologs of B-class genes (AP3, PI) expressed in the inner layer of utricle cells of the *Aristolochia* sepalloid perianth (as well as in the anthers), and in the staminoid “petals” of
Saruma (as well as in the stamens). It seems likely some sort of perianth tissue would replace the reproductive organs, but it not clear whether one or two different whorls of perianth tissue would appear. Although no AP1 ortholog has been reported, an AP2 ortholog was found in the unigenes from A. fimbriata (Luo et al., 1996; Feng et al., 2006). A transgenic plant over-expressing AP2 would be expected to resemble the A. fimbriata equivalent of the AG loss of function transgenic plant, described above, and loss of AP2 function in a transgenic A. fimbriata plant would be expected to display a phenotype of AG over-expression.

The TCP family has also played a significant role in the evolution of form in angiosperms. Repression of the TCP CYC subfamily of genes is associated with inhibition of outgrowth in the perianth of Antirrhinum and Lotus (Luo et al., 1996), in the potato “eye” (Faivre-Rampant et al., 2004), and in the lateral branches of Zea, Oryza, and Arabidopsis (Doebly et al., 1997; Takeda et al., 2003; Li et al., 2005; Finlayson, 2007). Overexpression of a CYC subfamily gene in A. fimbriata under a CaMV (constitutive) promoter might produce an embryonic lethal phenotype, but otherwise would be expected to inhibit growth in many locations. Expression could be targeted to particular tissues with different promoters. As some CYC subfamily members are regulated by miRNA (Busov et al., 2008), constitutive expression of miRNA or repression (by RNA interference) of a miRNA would also be expected alter morphology. Studies of molecular evolution of this important gene family will suggest additional hypotheses to test in the A. fimbriata experimental model.
Appendix A

Whole genome comparisons reveal intergenomic transfers

This appendix contains material that may be submitted for publication by Barbara Bliss¹.

¹Department of Biology, Institute of Molecular Evolutionary Genetics, and the Huck Institutes of the Life Sciences, 201 Life Sciences Building, Pennsylvania State University, University Park, PA 16802, USA
Abstract

Whole genome comparison performed between fully sequenced chloroplast and mitochondrial genomes revealed numerous regions of high identity. In comparisons between the *Arabidopsis* chloroplast and mitochondrial genome, and between the *Spinacia* chloroplast and *Beta* mitochondrial genome, over 75 gap-free alignments of at least 100 bases with at least 70% identity were identified. A computer simulation was devised to evaluate whether these regions of high identity had occurred by chance, and found they had not. Five of the high-identity regions, *psbA*, *rbcL*, *ycf1*, *petG*, and *psaA*, were selected for phylogenetic analysis to test the hypotheses that the significant similarities in sequence were due to horizontal transfer from chloroplast to mitochondrion. Three of these, *psbA*, *rbcL*, *ycf1*, were verified using PCR. My results confirm the utility of whole genome comparisons for revealing horizontal transfers, confirm previously reported horizontal transfers of chloroplast to mitochondrial sequences, and report evidence for horizontal transfer in the asterid lineage.
Introduction

Evolutionary relationships can be inferred from sequence similarity by applying the principle of descent with modification (Darwin, 1958). Sequence data from any of the three (mitochondrial, chloroplast, and nuclear) genomes may be selected for use in plant phylogenetic analysis, depending on the potential of the sequence to distinguish the species of interest (Soltis and Soltis, 1998; Karol et al., 2001). Horizontal gene transfer (HGT) contradicts patterns of vertical genome transmission, as has been described in the community of prokaryotic ancestors at the base of the tree of life, in which evolutionary relationships more resemble a net than a hierarchy (Doolittle, 1999; Jansen et al., 2007). The earliest eukaryotic cells also used HGT as endosymbiotic relationships were established with the prokaryotic precursors of the chloroplast and mitochondrion (Baldauf and Palmer, 1990; Woese, 1998; Woese, 2000; Dyall et al., 2004). Not to be relegated to the past, HGT continues as an ongoing process (Bergthorsson et al., 2003; Stegemann et al., 2003) and may account for 1/16,000 tobacco plants carrying a new section of chloroplast DNA in its nuclear genome (Huang et al., 2003). HGT provides an additional source of variability in the genome, one that is particularly important to understand when tracing phylogenetic relationships.

Historically, horizontal transfers between plant organelle genomes have been revealed one at a time, using careful hybridization (Stern and Palmer, 1984; Sederoff et al., 1986; Baldauf and Palmer, 1990) together with phylogenetic analysis of sequence data (Delwiche and Palmer, 1996). Present-day availability of whole, well-annotated genomes and the tools to manipulate them allow rapid identification and analysis of
HGTs, offering new insights into poorly understood events and screening organelle genomes for horizontally transferred sequence data, which could be misleading in reconstructions of organismal phylogenies.

In this study, I used a bioinformatics approach to develop a whole genome comparison of multiple organelles, revealing numerous putative horizontal transfers, large and small. Experimental evidence from PCR and sequencing reactions confirmed the existence of previously reported sequences, as well as new ones. Pairwise comparisons and phylogenetic analyses of alignments from fully sequenced genomes support the hypothesis that horizontal gene transfer has extensively invaded organelle genomes.

**Methods**

**Sequence selection, alignment, and phylogenetic analysis**

PipMaker (Adams et al., 2000; Schwartz et al., 2000) was used to generate pairwise alignments of closely related mitochondrial (mt) and chloroplast (cp) genomes. The *strong-hits* program of PipTools ([http://bio.cse.psu.edu/pipmaker](http://bio.cse.psu.edu/pipmaker)) was used to identify alignments having at least 70% identity for 100 bases or more, and interrupted by four or fewer gaps. MultiPipMaker (Schwartz et al., 2003) was used to align and produce stacked sets of percent identity plots (pips) comparing chloroplast and mitochondrial genome to each other (*Table A-1*). Annotations in GenBank
(http://www.ncbi.nlm.nih.gov/Genbank/) were used to derive labels for exons in the

*Arabidopsis* and *Spinacia* chloroplast

**Table A-1**: List of species and GenBank accessions for 25 organelle genome sequences.

<table>
<thead>
<tr>
<th>species</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chloroplast genomes</strong></td>
<td></td>
</tr>
<tr>
<td>1 Adiantum capillus-veneris</td>
<td>NC004766</td>
</tr>
<tr>
<td>2 Amborella tricopoda</td>
<td>NC005086</td>
</tr>
<tr>
<td>3 Arabidopsis thaliana</td>
<td>NC000932</td>
</tr>
<tr>
<td>4 Calycanthus fertilis</td>
<td>NC004993</td>
</tr>
<tr>
<td>5 Chaetospheridium globosum</td>
<td>NC004115</td>
</tr>
<tr>
<td>6 Chlamydomonas reinhardtii</td>
<td>NC005353</td>
</tr>
<tr>
<td>7 Lotus japonicus</td>
<td>NC002694</td>
</tr>
<tr>
<td>8 Marchantia polymorpha</td>
<td>NC001319</td>
</tr>
<tr>
<td>9 Medicago truncatula</td>
<td>NC003119</td>
</tr>
<tr>
<td>10 Mesostigma viride</td>
<td>NC002186</td>
</tr>
<tr>
<td>11 Nicotiana tabacum</td>
<td>NC001879</td>
</tr>
<tr>
<td>12 Oenothera elata</td>
<td>NC002693</td>
</tr>
<tr>
<td>13 Oryza sativa ssp. japonica</td>
<td>NC001320</td>
</tr>
<tr>
<td>14 Pinus thunbergii</td>
<td>NC001631</td>
</tr>
<tr>
<td>15 Populus trichocarpa</td>
<td>NC009143</td>
</tr>
<tr>
<td>16 Psilotum nudum</td>
<td>NC003386</td>
</tr>
<tr>
<td>17 Spinacia oleracea</td>
<td>NC002202</td>
</tr>
<tr>
<td>18 Triticum aestivum</td>
<td>NC002762</td>
</tr>
<tr>
<td>19 Zea mays ssp. mays</td>
<td>NC001666</td>
</tr>
<tr>
<td><strong>Mitochondrial genomes</strong></td>
<td></td>
</tr>
<tr>
<td>1 Arabidopsis thaliana</td>
<td>NC001284</td>
</tr>
<tr>
<td>2 Beta vulgaris</td>
<td>NC002511</td>
</tr>
<tr>
<td>3 Chaetospheridium globosum</td>
<td>NC004118</td>
</tr>
<tr>
<td>4 Marchantia polymorpha</td>
<td>NC001660</td>
</tr>
<tr>
<td>5 Oryza sativa ssp. japonica</td>
<td>NC011033</td>
</tr>
<tr>
<td>6 Zea mays ssp. mays</td>
<td>NC007982</td>
</tr>
</tbody>
</table>

genomes. RNA-coding genes, which are used by both genomes, were excluded and only chloroplast-specific sequences (Shimko et al., 2001) were further evaluated. Selected regions of the multiple pairwise DNA alignments produced by MultiPipMaker were excised with the **slice** program of PipTools (http://bio.cse.psu.edu/pipmaker), examined,
and manually adjusted with the GENEDOC sequence analysis program (Nicholas et al., 1997) and with ClustalX (Thompson et al., 1997). Alignments were analyzed with PAUP* version 4.10b (Swofford, 2002). Nucleotide sequences were analyzed using Neighbor-Joining/UPGMA method (version 3.573) with negative branch lengths allowed, and amino acid sequence alignments were analyzed with parsimony.

**Comparing distributions of frequencies of similarity scores**

The C programming language was used in the Metrowerks CodeWarrior Integrated Development Environment (IDE) release 2 to develop a computer simulation (Bliss, B. J. dissertation 2008 enclosure\Cprogramming\overlap\Console App Release.exe: contact author for access to program and related files) to compare the distribution of frequency of similarity scores in alignments of actual organelle genomes with the distribution of frequency of similarity scores in alignments of a simulated genome sequence with an actual “target” genome sequence. The simulation reads an alignment file from MultiPipMaker containing multiple pairwise alignments of one “target” sequence with two or more “query” sequences. Along the alignment, windows of a user-specified length are evaluated and scored for percent similarity (0.00-1.00). The frequency with which each similarity score occurs in the alignment is tallied. For each “query” sequence in the input alignment, the program calculates the base composition over the length of the sequence and generates a specified number of replicate, simulated “query” sequences. Each replicate is composed of randomly specified nucleotides, such that each simulated “query” sequence has the same base composition as the original
“query” genome sequence. The frequency distributions for the alignments of the actual and simulated genome sequences are output in a text file suitable for use with a spreadsheet program.

**Amplification and sequencing**

Total DNA was extracted using the CTAB method of McNeal et al. (2006). After DNA was rinsed with 70% ethanol, samples were re-suspended and stored in .01M Tris. MacVector (Rastogi, 1999) was used to develop primers for the regions flanking the chloroplast genes *psbA*, *rbcL*, and *ycf1* found in both the *Arabidopsis* mitochondrial and chloroplast genomes (Table A-2).

**Table A-2**: Primers to amplify chloroplast genes from *Arabidopsis* cp and mt genomes.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Organelle genome</th>
<th>Forward primer (5'-3')</th>
<th>Reverse primer (5'-3')</th>
<th>Annealing temp. °C</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>psbA</em></td>
<td>cp</td>
<td>AAT CCA CTT GGC TAC ATC CG</td>
<td>TTT CCG TCT GGG TAT GCG TC</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>mt</td>
<td>CCA ATG GTT CCT TCT CTG AGC GT</td>
<td>GCC AGT AGA AGA CGA CAA TAG GTG AGG T</td>
<td>54</td>
</tr>
<tr>
<td><em>rbcL</em></td>
<td>cp</td>
<td>CTT TTT GAA GGA GGT TCG GTT AC</td>
<td>TGA ATA CCC CCT GAA GCC AC</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>mt</td>
<td>ATA GCC TTC CTG CGG GTA GGG T</td>
<td>CTT GTG AAC TCA GAT AGC CTT GTG G</td>
<td>54</td>
</tr>
<tr>
<td><em>ycf1</em></td>
<td>cp</td>
<td>CTT GGT CCT GTT TAG TCC CAC</td>
<td>AAA GCC AAG GGG TCA ACA GG</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>mt</td>
<td>AAG AAG TTG TCC CCT CTT CTC</td>
<td>GTC TTC CCA TTC ATT CCC AG</td>
<td>51</td>
</tr>
</tbody>
</table>

Polymerase chain reaction (PCR) total volume 50 μL contained 5.0 μL Mg-free Taq buffer (Stratagene: 10X) and 5.0 μL MgCl₂ (25 mM) for *psbA* and *rbcL*, or 5.0 μL Taq extender buffer with Mg (Stratagene: 10X) and 0.5 μL Taq X enzyme (Stratagene) for *ycf1*; 8.0 μL dNTPs (Pharmacia:dATP, dCTP, dGTP, dTTP each at 1.25 mM); 1 μL
each of left and right primers (MWG Biotech: 20 μM); 0.5 μL Taq (Stratagene: 5U/μL); 1 μL DNA template. PCR conditions were 94°C for 2 minutes initially, followed by 30 cycles of denaturing at 94°C for 30 seconds, annealing at primer-specific temperatures (Table A-2) for 60 seconds (touch-down 1°C cycle), extension at 72°C for 2 minutes, and an additional 5 minutes of extension at 72°C in the last cycle. Five μls of each PCR reaction were loaded onto 2.5% high-resolution agarose gel (Sigma-Aldrich Co., St. Louis, MO, #A-4718) for electrophoresis. Selected PCR products were excised from the gel, followed by gel extraction (QiaQuick, Qiagen). The chloroplast and mitochondrial ycf1 sequences were determined by automated DNA sequencing on a Beckman-Coulter CEQ2000 genetic analyzer according to manufacturer’s instructions.

**Results**

In pairwise comparisons of the Arabidopsis and Spinacia chloroplast genomes with 17 other organelle genomes, MultiPipMaker found numerous regions of high quality alignment occurring with P<0.01 (red) and P<0.05 (green) (Figure A-1 A, B).
Figure A-1: MultiPipMaker plots of organelle genome sequences aligned to chloroplast target genome sequences. **A. Arabidopsis** chloroplast genome compared to 17 other organelle genomes. **B. Spinacia** chloroplast genome compared to 17 other organelle genomes. Exon names on upper edge indicate coding sequences recorded in GenBank for the indicated chloroplast genome. High quality alignments are indicated in red (P<0.01) and green (P<0.05).
The shape of curves produced by distributions of similarity scores from pairwise alignments of closely related mt v. cp genome sequences differed from the shape of curves produced by distributions of similarity scores from pairwise alignments constructed with simulated mt sequences (Figure A-2). Alignments were evaluated in windows of 50, 100, 150, and 300 bases. Regardless of the window size, species or genome used, alignments containing simulated sequences produced a bell-shaped curve (random distribution) and alignments using actual sequences produced an L-shaped curve (Figure A-2).

In alignments containing simulated sequences, as window size increased, the range of high frequency similarity scores narrowed as the amplitude increased, such that the center of the distribution remained the same. For example, in the alignment of the simulated Arabidopsis mt genome v. the Arabidopsis cp genome, when the alignment was evaluated with short (50 bp) windows, over 5% of the alignments had approximately 5-35% identity (a range of 30 percentage points), but when evaluated with long (300 bp) windows, the same proportion had approximately 10-30% identity (a range of only 20 percentage points) (Figure A-2 A). Correspondingly, there were fewer short windows with precisely 20% identity than there were long windows. Regardless of window size, the greatest proportion of the alignment with the simulated sequences displayed 20% sequence identity (Figure A-2 A).

In alignments with actual genomes, alignments with over 50% identity occurred consistently, although less than 5% of the time for any particular value. In comparison, in alignments containing simulated genomes, windows with over 50% identity never occurred (Figure A-2 A). In addition, there were a greater proportion of windows having
more than 50% identity in the mitochondrion v. chloroplast genome alignments of Beta mt v. Spinacia cp, Oryza mt v. Oryza cp, and Zea mt v. Zea cp than in the Arabidopsis mt v. Arabidopsis cp genome alignments (Figure A-2 A-D). In Zea, over 10% of the aligned
organelle genomes shared 90% identity, and in *Oryza* and *Beta/Spinacia*, over 5% of the aligned organelle genomes shared 90% identity (*Figure A-2 B-D*).

The shape of distribution of similarity scores in pairwise alignments of cp v. cp genome sequences differed from the shape of the distributions of similarity scores in pairwise alignments of mt v. cp genomes (*Figure A-3*). As in *Figure A-2*, distributions from the simulated organelle genome sequence alignments (series with broken lines, all colors), all displayed bell shaped curves while the distributions from actual organelle genome sequence alignments (series with solid lines, all colors), displayed L-shaped curves (*Figure A-3*). However, distributions of similarity scores in pairwise alignments of 11 chloroplast genomes (series with solid blue–green lines) to the target *Arabidopsis* (*Figure A-3 A*) and *Spinacia* (*Figure A-3 B*) chloroplast genomes all took approximately the same shape, while distributions of similarity scores in the alignments of six mitochondrial genomes (series with solid yellow-red lines) to the target cp genomes overall took a different shape (*Figure A-3 A, B*). The distributions of scores in the pairwise alignments of cp genomes, actual or simulated, v. target cp genomes (series with solid or broken blue–green lines) produced a higher frequency of 100 bp, high identity alignments (*Figure A-3 A, B*).
Figure A-3: Distribution of similarity scores for 100 bp windows from 17 pairwise alignments of mt genome sequence or simulated mt genome sequence to cp genome sequence for A. *Arabidopsis* and B. *Spinacia* cp genomes.
Fully sequenced mt genomes were also aligned with closely related cp genome sequences using PipMaker (Schwartz et al., 2000). The strong-hits program of PipTools indicated the pairwise alignments of the Arabidopsis and Beta mt genomes to their closest cp genome sequences (Arabidopsis or Spinacia) had more than 75 gap-free, 100 bp alignments with at least 70% identity in regions annotated with chloroplast-specific genes (data not shown). These alignments included coding sequences for the photosystem II reaction center polypeptide D1 (psbA); the large subunit of ribulose bisphosphate carboxylase (rbcL); the cytochrome b6/f complex subunit 5 (petG); the photosystem II reaction center protein D2 (psbD); two conserved, large open reading frames frequently found in angiosperms (ycf1, ycf2); photosystem I P700 apoprotein A1 (psaA); and an approximately 500 bp open reading frame (orf) (Table A-3). Three of these sequences, rbcL, psbA and ycf1 were excised from MultiPipMaker multiple sequence alignments for input to PAUP* for phylogenetic analysis and verification by amplification and sequencing.
Bootstrap (500 replicates) consensus trees generated from analyses of the MultiPipMaker alignments of the *psbA*, *rbcL*, and *ycf1* sequences indicated the *Arabidopsis* chloroplast and mitochondrial sequences were most closely related to each other (*Figure A*–*4*) than to other chloroplast or mitochondrial sequences.

Table A-3: Locations of chloroplast-specific sequences in pairwise alignments of mt genome sequences with cp genome sequences. Base location references the pairwise alignment of *Beta* mt v. *Spinacia* cp genome sequences or *Arabidopsis* mt v. *Arabidopsis* cp genome sequences.

<table>
<thead>
<tr>
<th>exon (pipname)</th>
<th>Organelle genome base location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Arabidopsis</em> chloroplast</td>
</tr>
<tr>
<td><em>psbA</em></td>
<td>218-750</td>
</tr>
<tr>
<td><em>rbcL</em></td>
<td>55493-56072</td>
</tr>
<tr>
<td><em>petG-1</em></td>
<td>64599-64791</td>
</tr>
<tr>
<td><em>petG-2</em></td>
<td>64860-64976</td>
</tr>
<tr>
<td><em>psbD</em></td>
<td>33567-33674</td>
</tr>
<tr>
<td><em>ycf2</em></td>
<td>86178-86608</td>
</tr>
<tr>
<td></td>
<td>127374-127487</td>
</tr>
<tr>
<td><em>ycf1</em></td>
<td>1228712-129641</td>
</tr>
<tr>
<td></td>
<td>109008-109937</td>
</tr>
<tr>
<td></td>
<td>126568-127489</td>
</tr>
<tr>
<td></td>
<td>105956-106911</td>
</tr>
<tr>
<td><em>psaA</em></td>
<td>39012-39162</td>
</tr>
<tr>
<td><em>orf</em></td>
<td>105187-105698</td>
</tr>
</tbody>
</table>
Figure A-4: Strict consensus trees from NJ analysis with PAUP* (500 bootstrap replicates) using aligned coding sequences from cp and mt genomes for A. *rbcL* (13 taxa) B. *psbA* (13 taxa) C. *ycf1* (15 taxa).
To verify that the three *Arabidopsis* sequences analyzed in Figure A-4 were not misannotations, I used primers designed from the pairwise alignments to separately amplify the *psbA*, *rbcL*, and *ycf1* coding sequences from the *Arabidopsis* chloroplast and mitochondrial genomes (Figure A-5). The *ycf1* chloroplast and mitochondrial genome products were sequenced and had less than 1% difference from the corresponding accessions in GenBank (data not shown).

---

**Figure A-5:** Successful amplification of *psbA* (**A**, **C**), *rbcL* (**B**, **D**), and *ycf1* (**E**, **F**), from the *Arabidopsis thaliana* (**A**, **B**, **E**) chloroplast and (**C**, **D**, **F**) mitochondrial genomes.
Bootstrap (500 replicates) consensus trees generated from analyses of the MultiPipMaker alignments of petG and psaA sequences revealed the Spinacia chloroplast and Beta mitochondrial sequences are more closely related to each other (Figure A-6) than to other chloroplast or mitochondrial sequences.

Figure A-6: Strict consensus trees from PAUP* (500 bootstrap replicates) analysis of aligned coding sequences from cp and mt genomes using A. NJ analysis of petG nucleotide sequence (15 taxa) B. Parsimony analysis of psaA amino acid sequences (13 taxa).
chloroplast or mitochondrial sequences, despite the genome-wide greater similarity of mitochondrial genomes to other mitochondrial genomes (Figure A-3).

**Discussion**

Genomic sequences displaying similarity to those of the chloroplast or mitochondrial genomes may be similar due to horizontal rather than vertical transmission. In-depth studies have revealed repeated transfers of conserved sequences from the plant chloroplast and mitochondrial genomes to the plant nucleus (Baldauf and Palmer, 1990; Adams et al., 2001; Millen et al., 2001; Adams et al., 2002), Bioinformatic methods have revealed horizontal transfers to the nucleus on a larger scale (Lin et al., 1999; Stupar et al., 2001; Rice Chromosome 10 Sequencing Consortium, 2003). I sought to determine if comparisons of fully sequenced organelle genomes could yield additional insight into horizontal transfers.

In this study, whole genome sequence alignments of multiple chloroplast and mitochondria genomes revealed numerous regions of high-identity. Those regions attributable to common, conserved function (e.g., ribosomal RNA genes) were excluded from the analysis. I hypothesized that the remaining high-identity alignments arose either by horizontal transfer, or else by chance, due to the base composition of the mt genome sequence. In simulations conducted to evaluate the latter alternative, gap-free alignments of 100 bp or more and greater than 50% identity were not observed. Gap-free alignments of over 100 bp with 70% or greater sequence identity were subjected to phylogenetic analyses, which demonstrated that the *psbA, rbcL, ycf1, petG*, and *psaA* sequences from
the fully sequenced mt genomes were most closely related to the corresponding sequences in the cp genome, reflecting horizontal transfer. Additional analyses could query nuclear genome sequence data to determine if horizontally transferred sequences were transferred directly from the chloroplast to the mitochondrion, or if the nuclear genome served as an intermediate.

The greater proportion of high-identity alignments between the Zea/Zea, Oryza/Oryza, and Beta/Spinacia organelle genomes than between the Arabidopsis/Arabidopsis ones suggests transfers from the Arabidopsis chloroplast to mitochondrion may have happened longer ago than the others, or may otherwise be diverging more rapidly. Cummings et al. (2003) found that, among the transfers of rbcL to the Arabidopsis, Brassica, Zea, Oryza, and Ipomoea mt genomes, those in the crucifer lineage were most ancient. Comparative, whole genome analysis can be used to describe the timing of large-scale genomic events (Cui et al., 2006). Including additional species in whole genome analysis of organelle sequence data can help to better estimate the frequency and timing of large scale horizontal transfers.

Acknowledgements

Xiaomu Wei and Joel McNeal provided technical assistance. Dr. Webb Miller provided extended access to the PipMaker server http://bio.cse.psu.edu/pipmaker/. The lab of Dr. Hong Ma provided A. thaliana, Lansberg ecotype. The lab of Dr. Ramesh Raina provided A. thaliana, Columbia ecotype DNA. Primers and materials were funded from the Chloroplast Genome grant awarded to Dr. Claude dePamphilis.
Appendix B

Cultivation Supplement

This appendix contains material that may be submitted for publication by Barbara Bliss. This Cultivation Supplement describes seed germination experiments and cultivation methods. Included are detailed, illustrated instructions for hand pollinating Aristolochia fimbriata and related species (e.g., A. elegans). The system used for recording plant pedigree is described herein, as are abnormal seedlings initially observed during AAGP seedling tissue collection efforts.

1Department of Biology, Institute of Molecular Evolutionary Genetics, and the Huck Institutes of the Life Sciences, 201 Life Sciences Building, Pennsylvania State University, University Park, PA  16802, USA
Seed germination experiments

Reliable seed germination methods are prerequisite for developing inbred lines, for collecting tissues for seedling expression libraries (etiolated and non-etiolated tissues used by AAGP), for developing seedling selection protocols for antibiotic selection of potentially transgenic individuals, and for evaluating seed storage methods. *Aristolochia* germinates best in potting medium in the greenhouse, but, some circumstances require soil-free germination to allow continual access to the germinating seed. In order to identify an efficient method of germinating seeds, I initially investigated several methods of germination, using 25 seeds each, placed in varying conditions of light, temperature, and medium for 42 days (Figure B-1). Seeds germinated with varying success in different conditions. I experimented with total darkness (D), and with 16 hour days with 8 hour nights (16/8). Light and temperature conditions were provided by varying means. Agrowth chamber provided 30°C days, 16°C nights (GrCham). A Conviron room programmed for maize provided 25°C days, 25°C nights (cornEnviron). An incubator set at 30°C, provided no light for seeds incubated on petri dishes (PD) containing sucrose-free, half-strength MS media (PDmedia), or wet filter paper (PDfp). Seeds were from open-air, insect-mediated pollinations of the S1 and S2 lines combined (OAL) or from self-pollinations of individual S1 and S2 plants (S1-30, S2-18). Data collected from the initial experiment were not appropriate for statistical analysis, but the three samples with the highest percent germination (mean= 0.96, SD 0.04) were those germinated in petri dishes containing filter paper (PDfp; Figure B-1). Lack of humidity in the incubator and in the cornEnviron was problematic, and seeds often dried out.
Figure B-1: Seeds planted in varying conditions on 9/28/2006. Conditions of light, medium, humidity and temperature were varied. Seeds germinated irregularly. Greatest percent germination results were obtained using petri dishes (PD) containing wet filter paper (fp) instead of no sucrose media (media).

In another experiment to determine if seeds would germinate in the dark as well as in a light/dark regime (Figure B-2), seeds from open air pollinations of S1-27 flowers
were planted in petri dishes containing sucrose-free, half-strength MS media in an attempt to provide continuous moisture. The plates were incubated in a growth chamber operating at 30°C days, 16°C nights, either in total darkness (D) or 16 hour days/8 hour nights (16/8). Overall, the percent germination was low (mean 0.30, SD 0.1354) with no significant difference (two sample T test) between mean percent germinated in 16/8 and D treatments.

Figure B-2: Comparing effects of light on germination of seeds planted on ½ MS media. Seeds from open air pollinations of S1-27 flowers were planted on 10/6/2006 on ½ MS medium in petri dishes and incubated at 30°C days, 16°C nights, either in total darkness (D) or 16 hour days/8 hour nights (16/8). There was no significant difference (two sample T test) in mean percent germination in 16/8 and D treatments.

To directly evaluate germination on media compared with wet paper (Figure B-3), seeds from open air pollinations of S1-35 or S2-01 plants were planted in petri dishes
(PD; 25 seeds each) containing sucrose-free, half-strength MS (media) or wet filter paper (fp) and incubated in total darkness in a growth chamber at 30°C days, 16°C nights.

![Graph comparing seed germination on wet filter paper or medium. Seeds from open air pollinations of flowers on S1-35 or S2-01 plants were planted on 10/13/2006 and incubated in total darkness in a growth chamber at 30°C days, 16°C nights. Seeds germinated significantly better on wet filter paper (fp) than in petri dishes (PD) containing sucrose-free ½ MS media.](image)

Figure B-3: Comparing seed germination on wet filter paper or medium. Seeds from open air pollinations of flowers on S1-35 or S2-01 plants were planted on 10/13/2006 and incubated in total darkness in a growth chamber at 30°C days, 16°C nights. Seeds germinated significantly better on wet filter paper (fp) than in petri dishes (PD) containing sucrose-free ½ MS media.

Mean percent germination at 42 days was significantly better (Student’s two sample t-test, P < 0.05) in wet filter paper (n=11, mean=0.6145, SD 0.1572) than in media (n=3, mean=0.0933, SD 0.0611) (Figure B-3).

To evaluate the effects of light on seeds germinating in wet toweling (Figure B-4), seeds from open air pollinations of flowers on five different S2 plants were planted in packets of wet towelling, 25 seeds each, and incubated in total darkness (D, n=10) or in 16 hour days/8 hour nights (16/8, n=10) in a growth chamber at 30°C days, 16°C nights. There was no significant difference between the two treatments at 42, 70, or 91 days.
Mean percent germination at 91 days was 0.7940 (SD 0.1103). At 42 days, mean overall percent germination was 0.4480, SD 0.2131 (Figure B-4).

Results from Figure B-4 were lower than the mean percent germination at 42 days in a more brief germination exercise shown in Figure B-4, conducted with the same methods using seeds from different open-air pollinated S2 plants. In Figure B-4, the 42 day mean percent germination rate was 0.6171 (SD 0.2416), and percent germination in 16/8 (n=14, mean=0.723, SD 0.179) was significantly better than in D (n=14, mean=0.511, SD 0.255) (Student’s t test, P<0.05) (Figure B-5).
There was considerable variability of the effect of light on seed germination (Figure B-2, B-4, B-5, B-6). It may be that seeds from S2-16 and S2-17 generally germinate better than other seeds, but it does not appear that light itself is an important factor for seed germination. Paula Ralph helped with some of these experiments. We noted that seeds germinating in total darkness died, and light was required for the true leaves to emerge.

Figure B-5: Effects of light on germination of seeds planted in wet toweling. Seeds from open air pollinations of S2 plants planted 12/15/2006 in wet toweling, incubated in a growth chamber with 30°C days, 16°C nights had significantly higher percent germination after 42 days in 16/8 (n=14) than did seeds incubated in D (n=14).
Figure B-6: Comparing effect of light on germination of seeds planted on soft media. Seeds from open air pollinations of flowers on S1-44 plants planted 11/13/2006 on PD containing soft media and incubated in D (n=5) or in 16/8 (n=5) in a growth chamber at 30°C days, 16°C nights. Mean percent germination at 70 days was significantly greater in D.
Germinating in toweling

Although greenhouse germination in soil, on a heating mat generally produced the best results (Bliss et al., in preparation, 2008), seeds are germinated in toweling packets when seedlings must be handled. No other method was found to be superior (see Figures B-1 through B-6, also Figure 5-1). The toweling method was used exclusively to germinate seeds for AAGP seeding tissue collection (Figure B-7).

Figure B-7: Toweling germination method. A. Seeds arranged in center third of wet towel. B. Fold over one side C. Fold over the other, apply more water D. Insert toweling packet into plastic bag (plastic pedigree label is included in this bag), fold over and incubate under light.
Up to 80 (or more) seeds were arranged in the center third of a wet paper towel (Scott Hi Wet Strength), each side of which was then folded back over the seed array (Figure B-7 A, B). The packet was squirited with distilled water to saturate all layers of the toweling thoroughly. The tri-folded wet (but not dripping) towel packet was put into a polyethylene plastic storage bag, folded to loosely close it over and around the towel packet. The outside of the plastic bag was labeled to indicate the contents and placed under lights in a growth chamber operating at 35°C for 16 hour days, 25°C for 8 hour nights. Germination was recorded with the appearance of the hypocotyl emerging from the seed coat. Seeds also have been germinated on shelves in a room maintained at 70°F (21°C), approximately 18 inches (50 cm) below fluorescent lights fixtures containing 40W cool-white bulbs operating 16 hours/day.
Seed storage experiments

Seeds contain the germline of the individual, so the seeds produced by a transformed plant or naturally occurring mutant represent an important resource for studying the genetic basis of the phenotype. For annual plants, including Arabidopsis and Zea, seed storage is critical because the individual dies after producing seed. For A. fimbriata, a long-lived herbaceous perennial that can also be vegetatively propagated for ongoing use as a stock plant, seed longevity is less crucial, but it remains important for planning future experiments. In order to determine if seeds would remain viable after a period of time in storage, I looked for evidence of any initial decline in seed viability after short term (7-150 days; Figure B-8) and after longer-term (522-942 days; Figure B-9) storage in the workplace, at room temperature.

There were no significant differences in germination among seeds planted in wet toweling weekly, for 22 weeks (Figure B-8). These findings were consistent with observations from the greenhouse. After 30 days, 50% of the seeds planted at seven day intervals each group had germinated, and 90% germination was obtained in 63 days. This was considerably worse than in the greenhouse, where 90% germination was often observed in 30 days.
In the older seeds (Figure B-9), seeds in the middle group (761 days old), were significantly more likely to likely to germinate than the other groups, and the youngest group of seeds (577 days old) was slightly more likely to germinate than the oldest group of seeds (942 days old) (Figure B-9). The middle group of seeds, which had the highest percent germination, were collected in November, and the newest and oldest seeds were collected in May. These results suggest something other than seed age was affecting

Figure B-8: Short term effects of seed age on germination. Successive planting of seeds from open-air pollinations were germinated in toweling packets at weekly intervals, seven to 150 days after collection, with no significant differences in germination. Most packets reached maximum germination by 60 days, with over 50% germination obtained by 30 days. Mean germination at day 35 (n=10) was 0.7695 (SD 0.0199), at day 63 (n=15) was 0.8987 (SD 0.0537).
germination, perhaps the time of year when pollination and maturation took place, or conditions in the greenhouse when the capsules were maturing.

Figure B-9: Longer-term effects of seed age on germination. Seeds of different ages from open air pollinations of flowers on mixed plants (VL, and NV) were planted 6/15/2007 in wet toweling and incubated in 16 hour days/8 hour nights (16/8, n=14) in a growth chamber with 30°C days, 16°C nights. The midrange seed age (761 days old) was significantly ($P < 0.05$ in binary regression analysis) more likely (3.83 times more likely) to germinate than the newest seeds (577 days old). The oldest seeds 942 days old were 0.90 times as likely to germinate as the newest seeds, not significantly different.
Inducing flowering

*A. fimbriata* has a vining habit, and produces numerous indeterminate shoots, 1-3 m long, depending on conditions. Blooming can be induced in the same manner as is done for many herbaceous perennials, by pruning and feeding, as needed, to stimulate a flush of new growth. Cut plants back to the rhizome approximately two weeks before pollination efforts are scheduled to begin (*Figure B-10 A*). After shoots have begun to proliferate, set pots containing new shoots into larger pots, and trellis them (*Figure B-10 B*). Older, larger plants may be potted directly into large pots outfit with large trellises. Arrange trellised plants for pollination, fruit production, and harvest (*Figure B-10 C*).

Flower and secondary shoot meristems arise from axils (*Figure B-11*), so a new flower will open every day, or nearly so. Shoots will produce flowers indefinitely (*Figure B-12*), as they elongate, but productivity declines as shoots lengthen and age. Pruning can improve productivity. Heading (removal of terminal meristems) will induce axillary meristem outgrowth. Thinning (removal of shoots at the crown) will induce production of new shoots from the crown. Selective pruning can be used to maintain induction of flowers on fresh, lower shoots while fruit matures. Collect fruit after the capsule opens.
Figure B-10: Inducing blooms

A. Prune back *A. fimbriata* to induce new shoot proliferation, which will be visible in 3-5 days. This plant is approximately two years old.

B. Set pot containing plant into larger pot with trellis, in preparation for hand pollinations and seed collection

C. Arrange all plants for pollination, fruit maturation, fruit harvest.
Plants sharing a desirable pedigree may be isolated in a room apart from *A. fimbriata* plants of other parentage in order to generate closely related seeds by open air pollination, effected by insects naturally occurring in the greenhouse (e.g., fungus gnats and other dipterans). For AAGP, first and second generation inbred plants (S1 and S2 descended from LR08) were isolated and used as the source of tissues for transcriptome sequencing. Seeds from fruits produced by open-air pollinations in that room were germinated to yield seedling tissues used by AAGP.

Figure **B-11**: *A. fimbriata* produces successive axillary blooms along the shoot.
Figure B-12: *A. fimbriata* in bloom. Four main shoots from the crown producing flowers at successive stages.
Hand pollination

These instructions are written for a novice worker in the lab to begin conducting hand pollinations with *A. fimbriata*. A basic knowledge of floral organs and plant anatomy is required: the student may wish to refer to a plant biology text in the lab if terms are unfamiliar (Raven et al., 1999; Simpson, 2006; Bliss et al., in preparation, 2008).

Select a trellised plant producing flowers at the shoot terminals (Figure B-12), and tie a pollination bag around the unopened flowers. The entire end of the shoot may be enclosed in the bag, selecting three or more flowers for subsequent hand pollination (Figure B-13). Loosely wrap the excess material of the pollination bag to cushion the stem and tie it below the petiole of the lowest flower, preventing pollinator access without crushing the stem (Figure B-14 A). Attach tags just below the pollination bag (Figure B-14 B). Check pollination bags daily for open flowers (Figure B-15). Note on the tag the date the flower first appears open (day one; first day of anthesis).
Figure B-13: Pollination bag. Four unopened flowers are enclosed at the end of one shoot.
Figure B-14: Pollination bag attachment detail. A. Tie pollination bag around stem (a) below petiole of lowest flower enclosed in bag. B. Tie tag below string of pollination bag.

Hand pollinations should be done on day one or day two flowers, as those are the days in which the stigma is most receptive (Bliss et al., in preparation, 2008). Pollen can most easily be collected from day two and day three flowers, but we have routinely done self-pollination of day one flowers. On the first day of anthesis, the anthers have not yet dehisced, so it is necessary to scrape open the anther to collect pollen on the tip of the toothpick, or pry an anther locule (or portion of a locule) from the anther and deposit it on the receiving stigma. If pollen will be collected from a different flower, it also should have been bagged to prevent contamination with foreign pollen from visiting insects.
Once the flower has been removed from the pollination bag, take care to prevent pollinator access. Cut the utricle above the gynostemium (Figure B-16 A) and inspect the utricle, limb, and gynostemium. If an insect is present the pollination bag has failed, so abort the hand pollination attempt. On the first day of anthesis, the lobes of the gynostemium should curve inward slightly (Figure B-16 B).

Figure B-15: Day of anthesis. Flower enclosed in pollination bag. Note backward bent fimbriae.
Collect pollen on a toothpick from the anthers, which are on the outside of the gynostemium (Figure B-17 A). Transfer pollen on the toothpick to the stigma (Figure B-17 B). Note that the flower in Figure B-17 is a “day two” flower, so its pollen has
already dehisced. Self pollinations are optimally conducted with day two flowers (Bliss et al., in preparation, 2008).

Figure B-17: Self-pollination in a day two flower A. Collect pollen from the anthers with a toothpick. B. Transfer it to the stigmatic surfaces of the receiving flower. (This is self-pollination: Both A. and B. are the same flower)
Use lab tape to tape closed the remnant of the perianth, preventing further access to the stigmatic surfaces. Fold the edges of the perianth as needed. The perianth, gynoecium, and tape will abscise above the inferior ovary (Figure B-18 A). If successfully fertilized, the ovary will begin to grow in a week (Figure B-18 B) and mature in a month.

---

**Figure B-18**: Inferior ovary. **A.** Perianth, taped shut, abscisses where indicated by tip of arrow. **B.** Maturing ovary, after perianth with lab tape has abscissed.
Recording pedigree

A method for labelling the plants was devised to convey the maximum amount of information possible on the tag in the pot in the greenhouse. Pedigree information can be determined from the label in the pot. An alternate method of labelling the plants would be to assign a serial number to the plant and tag, and record in a notebook the pedigree of the plant tagged with the serial number. In that system, the serial number would serve to connect the plant to the notebook and the notebook would connect the serial number to the pedigree information. Based on the experience I had keeping lineage information for our tissue culture work, I elected to use labelling system in the greenhouse which would communicate the pedigree directly from the pot label. Capsules are stored in envelopes with the pedigree information written on the outside. Here is the labelling and pedigree recording system, which was used to label the first and second generation lines used for AAGP tissue collection.

The name is written from left to right, adding on the right end as the depth of the pedigree increases. The label is interpreted from right to left, as follows.

a. Genotype = LR or CO (equivalent to VL and NV)

b. Seed/plant ID = 01-99. At one time, there were about one hundred plants each of LR and CO, all from the same parent plant. Eventually, many were eliminated to reduce maintenance.

c. Type of event generating the capsule. The letters S and O are reserved:
   S=selfing, O=open air pollination. Crosses recorded in the experimenters lab notebook can use other letters.
d. Year in which this capsule was collected = e.g., 04, 05, etc.

e. Capsule collected for this type of event, this year, using A-Z, 1-9, beginning with A.

f. Numerical designation of plants produced from the designated fruit

Letters c-f iterate through generations of descent from the original parent plant.

Example 1:

LR08 – S04A10– S06D01

Table B-1: Pedigree record, example 1.

<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>(f)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LR 08</td>
<td>S 04</td>
<td>A 10</td>
<td>S 06</td>
<td>D 01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reading from the right-most field, the label printed above Table B-1 indicates the first (01) seed germinated from the fourth (D) capsule collected in 2006 from selfing attempts done on plant LR08-S04-A10, which was grown from the tenth seed germinating from the first (A) capsule (LR08-S04A) collected in 2004 from selfing attempts done on plant LR08.
Example 2:

LR08 – S03A01 – S04A01 – S04B03 – S05C

Table B-2: Pedigree record, example 2.

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>(f)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>(f)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR08</td>
<td>S03</td>
<td>A01</td>
<td>S04</td>
<td>A01</td>
<td>S04</td>
<td>B03</td>
<td>S05</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The label above Table B-2 designates the third (C) capsule collected in 2005 (05) from selfing (S) attempts done on plant LR08 – S03A01 – S04A01 – S04B03, which was, itself, the third (03) seed to germinate from the second (B) capsule collected in 2004 (capsule can be identified as capsule LR08 – S03A01 – S04A01 - S04B), from selfing attempts done on a plant (LR08 – S04A01 – S04A01) that was, itself, the first plant to successfully germinate from seeds taken from capsule (LR08 – S03A01 – S04A), which was the first collected in 2004 from selfing attempts on plant LR08 – S03A01, which was the first (01) seed to germinate from the first capsule (LR08 – S03A) collected in 2003 from selfing attempts done on LR08, also known as VL08.

In this manner, up to 26 (A-Z) capsules per year can be collected from selfing events, and open air pollinations and each other type of event that can be designated by the other 24 letters in the alphabet or numbers 1-9. Up to 99 seeds per capsule generating new plants can be tracked.
Possible mutants

During AAGP seedling tissue collection several abnormal seedlings were observed by Paula Ralph. Two types are illustrated here. One type, we referred to as “tricot” appears to be producing three cotyledons (Figure B-19 A, B). These seeds (Table B-3) had three cotyledons when the seedling was extracted from the softened seed coat. Another unusual seedling type seemed to produce “basal” leaves (Figure B-19 C-E), and was apparently unable to extend the internode normally until after the earliest true leaves emerged.

Table B-3: Pedigree and incidence of “tricots.”

<table>
<thead>
<tr>
<th>Expt. date</th>
<th>Parent</th>
<th>Frequency (n=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/6/06</td>
<td>LR08 S04 A05</td>
<td>2/200</td>
</tr>
<tr>
<td>10/27/06</td>
<td>LR05 S04 A01 S05 A01</td>
<td>1/?</td>
</tr>
<tr>
<td>12/15/06</td>
<td>LR08 S04 A01 S05 B02</td>
<td>2/50</td>
</tr>
<tr>
<td>12/15/06</td>
<td>LR12 S04 A01 S05 C03</td>
<td>1/150</td>
</tr>
</tbody>
</table>
Figure B-19: Abnormal seedlings. A, B. "Tricots" C, D, E with basal leaves.
Literature cited


Burnham CR (1959) teosinte branched. Maize Genet Coop Newsletter 33: 74


Glass flowers from the Ware Collection in the Botanical Museum of Harvard University (1940). Harcourt, Brace and Company, New York, 58 pp


Kimura M (1968) Evolutionary rate at the molecular level. Nat Aust 217: 624-626


Kosugi S, Ohashi Y (2002) DNA binding and dimerization specificity and potential targets for the TCP protein family. Plant J 30: 337-348


Lee HS, Chen ZJ (2001) Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. Proc Natl Acad Sci USA 98: 6753-6758


Mizukami Y, Ma H (1995) Separation of AG function in floral meristem determinacy from that in reproductive organ identity by expressing antisense AG RNA. Plant Mol Biol


Murre C, McCaw PS, Vaessin H, Caudy M, Jan LY, Jan YN, Cabrera CV, Buskin JN, Hauschka SD, Lassar AB, et al. (1989b) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 58: 537-544


matK, and phyA genes, with special reference to differentiation of chromosome numbers. Syst Bot 31: 481-492


Sugawara T (1987b) Taxonomic studies of *Asarum* sensu lato III. Comparative floral anatomy. Bot Mag Tokyo 100: 335-348


Wei FS, Wing RA (2008) A fruitful outcome to the papaya genome project. Genome Biol 9: 4


VITA

Barbara J. Bliss

403 Life Science Building
University Park, PA 16802
(814) 441-4453
bjb316@psu.edu

466 Orlando Ave
State College, PA 16803
(814) 861-3553
bjbliss@gmail.com

EDUCATION
Ph.D. Biology (Molecular Evolutionary Biology) 2008 Pennsylvania State University, University Park, PA
M.S. Environmental Science (Toxicology) 1998 Huxley College, Western Washington Univ, Bellingham, WA
B. S. Biology magna cum laude 1995 Western Washington University, Bellingham, WA

PUBLICATIONS

TEACHING
Pennsylvania State University 6/2003-12/2008 Plant Taxonomy (co-instruct field course), Biological Concepts and Diversity (lab instruction)

AWARDS
University Graduate Fellowship, Pennsylvania State University 2002-2003
PNWSETAC annual student presentation award 1998
American Association of University Women Scholarship 1996-1997
Hughes Summer Research Intern Scholarship 1994
Western Foundation Scholarships 1993-1995
Academic Merit Scholarship 1993-1994