Title

Petaloidy and the Plant Bauplan: Using floral development in the Zingiberales (Angiospermae: Monocotyledoneae) as a test case to understand the evolution of plant form and function

Author

Almeida, Ana Maria Rocha de

Publication Date

2013

Peer reviewed|Thesis/dissertation

Petaloidy and the Plant Bauplan: Using floral development in the Zingiberales (Angiospermae: Monocotyledoneae) as a test case to understand the evolution of plant form and function

by

Ana Maria Rocha de Almeida

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Plant Biology
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Chelsea D. Specht, co-Chair
Professor Michael Freeling, co-Chair
Professor Sarah Hake
Professor Nicole King

Fall 2013

Petaloidy and the Plant Bauplan: Using floral development in the Zingiberales (Angiospermae: Monocotyledoneae) as a test case to understand the evolution of plant form and function

By

Ana Maria Rocha de Almeida

Petaloidy and the Plant Bauplan: Using floral development in the Zingiberales (Angiospermae: Monocotyledoneae) as a test case to understand the evolution of plant form and function

By

Ana Maria Rocha de Almeida

Abstract

Petaloidy and the Plant Bauplan: Using floral development in the Zingiberales (Angiospermae: Monocotyledoneae) as a test case to understand the evolution of plant form and function

By

Ana Maria Rocha de Almeida

Doctor of Philosophy in Plant Biology
University of California, Berkeley
Professor Chelsea D. Specht, co-Chair
Professor Michael Freeling, co-Chair

With more than 260,000 species, the angiosperms are the most diverse group of land plants on earth today. Many would argue that their striking diversity stems from the acquisition of the *flower* along this evolutionary lineage. The argument goes that by enclosing the plant's sex organs, especially the ovule, the flower provided angiosperms with special means to withstand a wide range of environmental conditions, while facilitating pollination or pollinator attraction and seed protection and dispersal. Regardless, the diversity of shapes, colors, and sizes of flowers across the angiosperms is irrefutable and fascinating. Understanding the mechanisms that underlie flower diversity leads us to the understanding, at least in part, of how evolutionary processes have enabled the origin of different forms in nature.

Although the Modern Synthesis has provided a solid framework for understanding how genes evolve in populations, it lacks a theory to satisfactorily explain the evolution of morphological diversity, as it largely marginalized the role of development in the evolution of biological form. Recently, however, an increasing attempt to understand the interrelationships between evolution and development has emerged as a new research field known as evolutionary developmental biology, or, for short, evo-devo. The study of genes involved in different developmental processes, and how changes in these genes or on their regulation can lead to changes in organismal form has become an insightful field. This dissertation focuses on the evolution and diversification of floral morphology in the Zingiberales and their implications for our understanding of the evolution of plant bauplan. The tropical monocot order Zingiberales provides an excellent framework for evolutionary developmental studies, as changes in floral form throughout the evolution of this group are mainly due to changes of form and function in the petal and stamen whorls, where stamens become infertile and petaloid.

The first part of this dissertation describes how changes in classical floral organ identity genes result in changes in floral organogenesis throughout the evolution of the

Zingiberales. First, through a combination of careful morphological studies and genetic approaches, I establish the homology of floral organs, particularly the nature of the so-called 'petaloid appendages' on fertile stamens of the ginger group. Second, I show that positive selection is acting upon the *AGAMOUS* (*AG*) lineage, and changes in the *AG* protein suggest a mechanism capable of explaining the morphological changes observed in the Zingiberales flowers.

The latter part of this dissertation goes beyond organ identity genes to investigate the development of organ morphology. In this section, I demonstrate the involvement of the abaxial-adaxial (ab-ad) polarity gene network on the evolution of filament morphology, not only within the Zingiberales but also across all angiosperms, and provide evidence that morphogenetic processes, not just organ identity *per se*, are driving the evolution of floral form across the order. By studying ab-ad polarity genes, well-known for the establishment of abaxial and adaxial surfaces of leaves, sepals, and petals, I show how the same gene regulatory network has been co-opted during the evolution of angiosperms to shape filament morphology in flowering plants.

I conclude this dissertation by discussing the implications of these findings to our understanding of the mechanisms of plant bauplan evolution. Lastly, I analyze the floral evo-devo research program through a historical and philosophical perspective, hoping to shed light on future directions of research in the field of plant evo-devo, as a consequence of important conceptual changes that this field has undergone in the past two decades.

Dedication & Acknowledgements:

I'd like to start by acknowledging Chelsea D. Specht and Mike Freeling, whose support and encouragement have been fundamental. During these five years in Berkeley, I learned a lot from you, not only about biology but also about people, and about life. I will always be thankful for the trust both of you put in me. I hope I have succeeded in showing that your trust was well placed.

I am also grateful for my Dissertation Committee, especially Sarah Hake and Nicole King who watched over me throughout these years, and whose guidance helped me to continually improve.

During my time in Berkeley, I met innumerable amazing and supportive people. The Specht Lab current and former members, especially Roxana, Carolina, Tanya, Shayla, Heather, and Eduardo made my long days in the lab much more fun, pleasant, and fruitful. I thank each of you for your support and friendship. You rock! I also had the pleasure to work with very talented undergrads – Andrew, Chris, Ben, Miranda, Kelsie, Duda – whose dedication and hard work made me always feel it was worth it. I would also like to thank the incredible members of the Freeling Lab. Thank you for all the computational expertise, and the always invigorating lab meetings. Within the Freeling Lab members, I certainly owe a special thank to James Schnable, whose knowledge and friendship were always there when I cried for help. J., as I always say, you are the best!

In Berkeley, I had the pleasure to work with countless gifted and supportive colleagues. In particular, Holly Forbes and Barbara Keller from the UC Botanical Garden, and Steve Ruzin and Denise Schichnes form the CNR Imaging Facility. I feel very lucky to have spent these years in Berkeley and its amazing scientific atmosphere. A special thanks goes to all MVZ members, and to the Integrative Biology Department. Even though I did not work with them directly, they were always a great source of inspiration, and I certainly learned a lot from them.

I would also like to thank the 2008 PMB graduate student cohort: your friendship and support, especially during my painful first year was instrumental. In particular, I'd like to thank Patrick Shih whose calls for coffee seemed to be always at the most needed moments. Thanks, Pat. I'll certainly miss our long talks.

I certainly owe a lot to my previous mentors and the various amazing scientists who I had the pleasure to meet. My special thanks goes to Charbel El-Hani, Peu Rocha, Stuart Newman, and Elena Alvarez-Buylla, among many others, who inspired me throughout these many years.

To my friends from here and from there, I extend a heartfelt thanks. Among many others, I'd like to extend special thanks to the ones that were here with me during these five years, Betita, Dudu, Angelica, Cacá, Gegê, Brodinho & cia, Gabriel & Clarissa, Dudete, Felipe, Carrrrlos Eduarrrrdo Barrrbosa, Carolina Sardinha, Lena & Dudu, Tati & cia,

Carla & John, Ale & Veinho and kids, Wagner & Lucia, The Hodoglugil, Marry Ellen, Mario and Mia, Jessy and Max, Tammy and Harold, Cleide and Kian, and Ana Brito and Brunão. I would also like to thank my old friends, whose distance prevented them from being physically with me. I know our minds are always connected, though. A special thank goes to Wly, my eternal and loyal friend, without whom I would never have dreamed I could attend UC Berkeley. You started it all, and I can never thank you enough.

I would like to thank very much the funding agencies that supported me and my research throughout these five years: CAPES/Fulbright, National Science Foundation (Dissertation Improvement Grant), The American Society of Plant Taxonomists, and The Heliconia Society.

Finally, I'd like to dedicate this thesis to the Almeida-Blair Clan (it's hard core – Luana and Peter; and it's soft core – The Almeida and Blair families), for making life such an enjoyable and loving journey. Without your love and support, none of this would be possible.

- Ana Maria Almeida, August 2013

Table of Contents

CHAPTER 1

Introduction to the study of floral morphological evolution in the Zingiberales: an evolutionary developmental framework

Molecular mechanisms underlying flower development

Angiosperms comprise approximately 250,000 species and are considered a textbook example of a rich taxonomic group that emerged over a relatively short period of time (Magallon and Sanderson 2005; Magallon and Sanderson 2006; Magallon and Castillo 2009). This extensive species radiation has been theoretically linked to a new mode of reproduction that resulted mostly from the acquisition of a novel structure: the flower (Hernandez-Hernandez, Martinez-Castilla et al. 2007). Although the Modern Synthesis has provided a solid framework for understanding how genes evolve in populations, it lacks a theory of form to satisfactorily explain the evolution of morphological diversity (Pigliucci 2007). Also, the Synthesis largely marginalized the role of development in the evolution of biological form (Gilbert, Opitz et al. 1996). During the last twenty-five years, however, an increasing attempt to understand the roles of genes involved in development and their implication on the evolution of form emerged as a new research field known as evolutionary developmental biology, or, for short, evo-devo (Jaramillo and Kramer 2007; Carroll 2008; Ioannidis 2008).

The genetic basis of flower development was established at the beginning of the 1990's based on studies in *Arabidopsis thaliana* and *Antirrhinum majus* (Bowman *et al*. 1991; Coen and Meyerowitz 1991; Jack *et al*. 1992; Weigel and Meyerowitz 1994). In the classical ABC model of floral development, differential gene expression results in the specification of the various floral organs. A-class genes (*APETALA2* (*AP2*), and *APETALA1* (*AP1*)) are involved in the specification of sepals (1st whorl organ), and together with B class genes (*PISTALLATA* (*PI*), and *APETALA3* (*AP3*)) they specify petal identity (2nd whorl). B class genes are also involved in the specification of stamen identity, together with the C class gene (*AGAMOUS* (*AG*)). Furthermore, *AG* is also responsible for the specification of carpel identity (Coen and Myerowitz 1991). Recently, the importance of other classes of genes (such as E class genes, *SEPALLATA1, 2*, and *3*; and a D class gene *SEEDSTICK* (*STK*) that specifies ovule identity), expanded the classical model to the current ABCDE model in eudicots (Pelaz *et al*. 2000; Colombo *et al*. 1995) (Figure 1A). With the exception of *AP2*, ABCDE class genes encode transcription factors that share a common DNA-binding domain (MADS-box domain) of approximately 180bp. Apart from the MADS-box domain, these proteins also embed an Intervening (I-box) domain (~90bp), a Keratin (K-box domain) of ~210 bp, and a variable C-terminus region, therefore being known as the MIKC-type MADS-box genes (Kaufmann *et al*. 2005).

It has been proposed that those proteins interact as dimers and subsequently as tetramers to bind DNA and regulate the expression of downstream genes during flower

development. The so called 'floral quartet model' (Honma and Goto 2001; Theissen and Saedler 2001) has been well studied *in vitro,* where heterodimerization is necessary for proper gene regulation by MADS-box genes in eudicots (Yang *et al*. 2003; Piwarzyk *et al*. 2007; Su *et al*. 2008; Melzer and Theissen 2009). However, it has been proposed that heterodimerization of MADS-box proteins evolved from a homodimerization state in gymnosperms, via facultative heterodimerization in basal angiosperms, where the two monomers of the same protein would interact to form dimers and, subsequently, tetramers (Winter *et al*. 2002) (Figure 1B).

A vast amount of knowledge has been generated on the genetic basis of floral development in eudicots and grass monocot species, such as *Zea mays* (Whipple *et al*. 2004; Whipple *et al*. 2007), but very little is known regarding the evolution of floral development in non-grass monocots. A few examples, such as those studies in orchids, are Ochiai *et al*. (2004), Nakamura *et al*. (2005), Tsaftaris *et al*. (2006), Mondragon-Palomino and Theissen (2008), and Tsai *et al*. (2008). In many monocot families, there is no obvious differentiation between organs of the 1st (sepal) and 2nd (petal) whorls (e.g. tulips). In those cases, a petaloid organ called the tepal comprises the first two whorls. In order to explain this flower morphology, a modified ABC model was proposed in which B class genes expand their expression into the 1st whorl (Kanno *et al*. 2003; Kanno *et al*. 2007). However, monocots exhibit a large spectrum of floral morphologies not all of which can be explained by this modified model. According to Litt and Kramer (2010) monocot morphology is extremely diverse and their flowers exhibit a number of complex forms of petaloid organs and inflorescence structures that have yet to be studied. In many cases, homologs of the ABC organ identity genes have been implicated in the development of such features and, again, gene duplications have clearly played a role in diversification of gene function.

Recently, however, a new interpretation of the genetic basis of flower development emerged. Most importantly, the A-, B-, and C-class genes were shown to be necessary, although not sufficient for appropriate floral organ specification. The classical ABC descriptive framework was revisited (Mendoza & Alvarez-Buylla 1998; Mendoza et al. 1999), and an elegant logical model was proposed as a more mechanistic re-interpretation of the ABC model. By mapping the landscape of the known gene interactions during floral development, Mendoza and coworkers (1999) were able to recover the stable states (i.e., attractors) that correspond to the gene expression patterns correlated to floral organs, in concordance with what has been described by the ABC model. In this model, several genes are involved in complex causal patterns that, during floral development, result in the expression of A-, B-, and C-class genes (Alvarez-Buylla et al. 2010a).

More interestingly, this gene regulatory network approach provides an interesting way for one to understand floral development. Interacting genes comprise modules. Each module, represented by several interacting genes, describes one aspect of floral morphogenesis. For instance, it has already been proposed (Alvarez-Buylla et al. 2010b) that common molecular modules act during early flower development. Three main modules have already been identified: the Primordia Positioning, the Primordia Polarity, and the Primordia Identity and Growth modules (Alvarez-Buylla et al. 2010b) (Figure 2). This

modules' approach provide a mechanistic understanding of floral development on which testable hypothesis can be generated.

Evolution of floral morphology in the Zingiberales

The Zingiberales are a group of herbaceous tropical monocots comprising eight families and approximately 2,500 species. They diverged from their sister order Commelinales (Bremer *et al*. 2009) approximately 80MY ago. Molecular studies have established the major phylogenetic relationships within the Zingiberales (Kress 1990; Kress *et al*. 2001) while careful morphological studies have characterized floral development in most families (e.g., Barker and Steward 1962; Kirchoff and Kunze 1995; Box and Rudall 2006; Kirchoff *et al*. 2009). Most major changes in floral morphology within the order are ascribed to the petal and stamen whorls. Three of these changes are well characterized: the acquisition of a complete differentiation between the first (sepals) and second (petals) whorls (that is, acquisition of a *biseriate perianth*) after the divergence of the banana lineage (Musaceae); the impressive reduction in the number of fertile stamens in the ginger group (from 5-6 to 1 in Costaceae and Zingiberaceae or ½ fertile stamen in Cannaceae and Marantaceae); and the acquisition of a novel floral structure – the labellum – as a result of the fusion of infertile stamens in Zingiberaceae and Costaceae (Kirchoff *et al*. 2009). This dissertation focuses on changes of the stamen whorl during the evolution of the Zingiberales order. These changes are depicted on Figure 3.

If one takes into account the gene regulatory network approach described above, one can map the evolution of floral morphology in the Zingiberales in terms of changes in the underlying molecular modules, as depicted by Figure 3. It is interesting to notice that at the base of the ginger clade, the correlation between petaloidy and reduction in the number of fertile stamens suggests a correlated change in two gene regulatory modules: the identity module and the polarity module. This point in the evolutionary history of the Zingiberales is, therefore, the object of study of this dissertation.

This dissertation aims at answering the following questions: Given the evolution of androecial morphology in the Zingiberales order, what is the involvement of floral organ identity genes on this observed evolutionary history? Also, is the polarity network implicated in the observed changes? While answering these questions, this dissertation tried to propose broader implications of these phenomena to the understanding of the evolution of land plant bauplans.

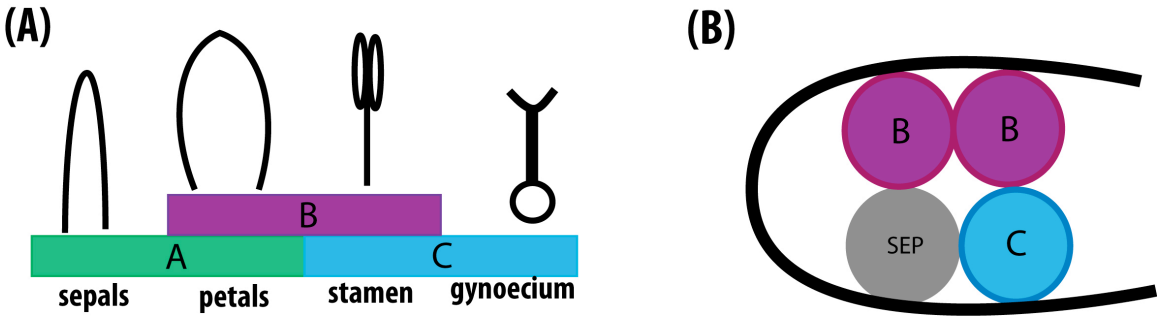

Figure 1. Schematic diagram of the (A) ABC, and (B) Quartet models. The ABC model of organ identity proposes that the domain of interaction of A-, B-, and C-class genes are correlated, in a combinatorial fashion, to the position of the different floral organ primordia during development. A-, B-, and C-class gene expression domains are depicted in green, pink and blue, respectively. According to the ABC model, sepals are specified whenever A-class genes are expressed alone. When a combination of A-, and B-class genes are expressed, petals are specified. Stamens develop as a result of a combinatorial expression of B-, and C-class genes, and gynoecium development is marker by the expression domain of C-class genes alone. The Quartet model proposes that the A-, B-, and C-class genes act via the formation of protein complexes. First, a dimmer is formed. Then, two dimmers are recruited to form a tetramer. According to the Quartet model, the regulation of transcription by the floral organ identity genes only occurs via the formation of the tetramer, as the active DNA binding unit. In (B) a hypothetical tetramer is depicted binding to a DNA loop (in black). This quartet might be former during stamen development, by a combination of B-class, and C-class gene products, as well as a SEPALLATA (SEP), frequently referred to as E-class genes. SEPALLATA genes are thought to be expressed in all floral organ primordia, during floral development.

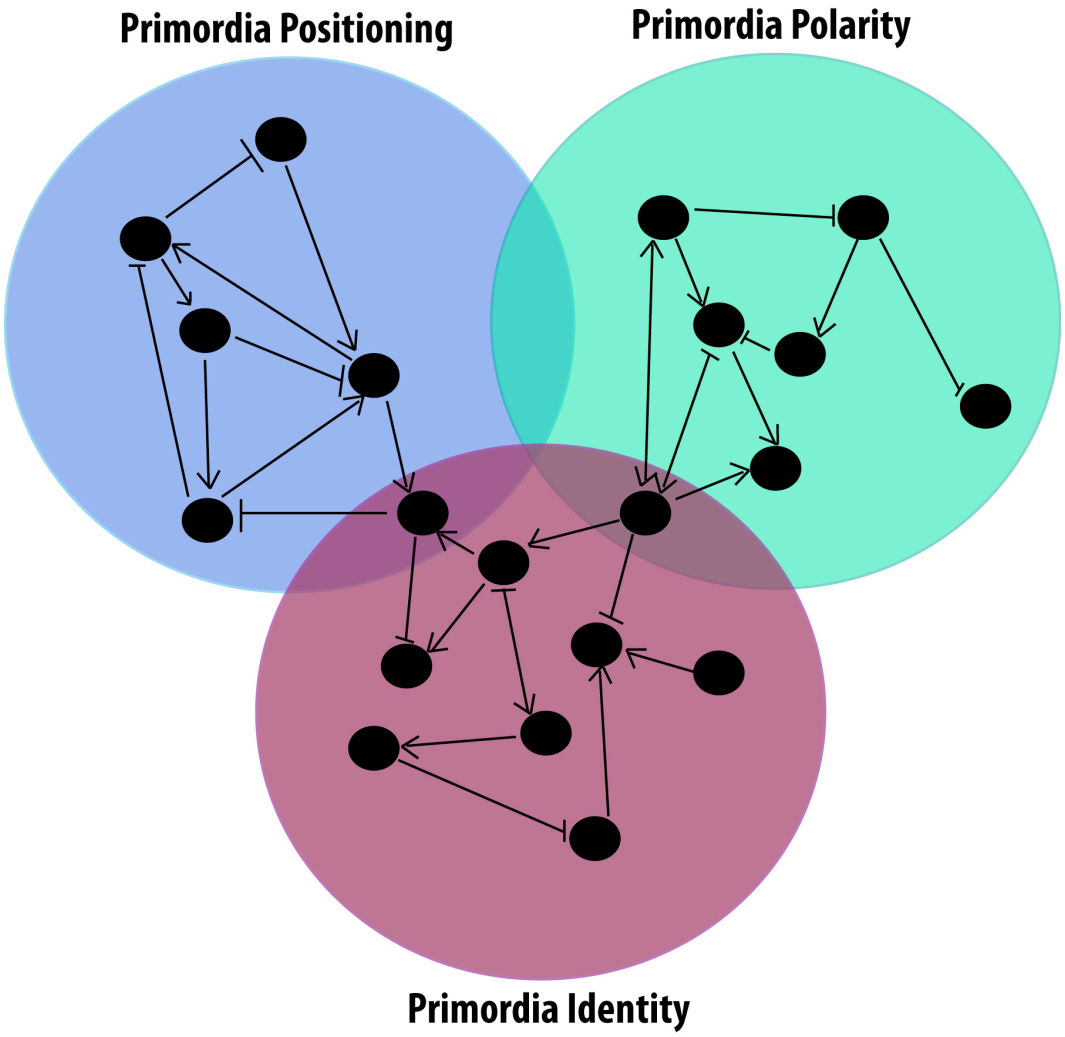

Figure 2. Hypothetical representation of the main Gene Regulatory Networks (GRN) operating during flower development. Genes are depicted as black dots, and their relationships as lines. → describes a positive interaction between two genes, while ⊥ describes a negative interaction in which one gene represses the expression of another gene. The Primordia Identity GRN is depicted in pink, and is responsible for specifying floral organ identity and growth. This GRN comprises not only the A-, B-, and C-class genes, but also other important genes, such as the ones involved in floral initiation and primordia size (e.g. WUS, CLV, LFY, etc). The Primordia Polarity GRN is depicted in green and is composed mostly of abaxial-adaxial polarity genes (e.g. KAN, YAB, PHV, REV, PHB, etc.). Primordia positioning is mainly a function of auxin-related genes (e.g., PIN, ARF, AXR1, etc.), and this GRN is depicted in blue. These GNR are described in detail elsewhere (Alvarez-Buylla 2010a), and at least the Primordia Identity module has been implemented computationally (Alvarez-Buylla 1998; 2010b), and was able to recover the gene expression patterns expected for the floral organs, and described by the ABC model of floral organ identity.

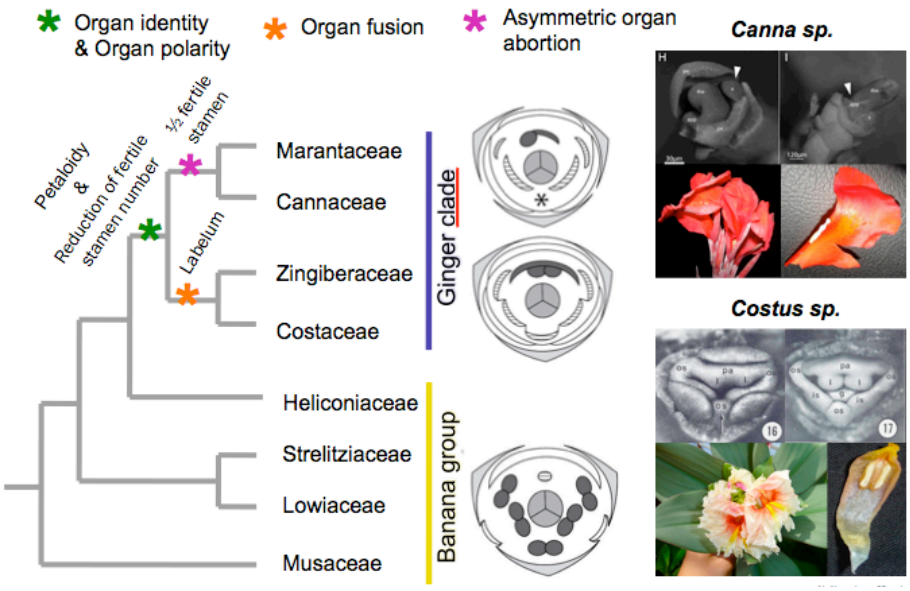

Figure 3. Evolution of floral morphology in the Zingiberales. The Zingiberales can be roughly divided into the paraphyletic banana lineages (Musaceae, Lowiaceae, Strelitziacea, and Heliconiaceae), and the ginger clade (Costaceae, Zingiberaceae, Cannaceae, and Marantaceae). The main morphological changes of the androecial whorl are depicted by * on the phylogenetic tree. After the divergence of the banana lineages, there is a striking reduction in the number of fertile stamens, from 5-6 fertile stamens in the banana group to 1 or ½ fertile stamen in the ginger clade. These infertile stamens (staminodes) also become petaloid leading to an interesting inverse correlation between pollen production and floral display (Specht et al. 2012). This change can be mapped to two distinct gene regulatory networks, the primordial identity and the primordial polarity network. Further changes in the androecial whorl occur within the ginger clade. In the families Costaceae and Zingiberaceae, different numbers of staminodes fuse, giving rise to the labellum, suggesting a shift in the organ boundary gene network. In the families Cannaceae and Marantaceae, the single fertile stamen is further reduced to a single theca, leading to an asymmetric organ abortion. It is also interesting to notice that within the ginger clade, the remaining fertile stamen bears what has been called a petaloid appendage (top and bottom right images). In the case of Costaceae (Costus scaber microscopy images modified from Kirchoff 1998), this petaloid appendage arises as an outgrowth of the filament. The development of Cannaceae petaloid appendage (top right corner) had not yet been studied previously. Drawings represent floral diagrams for Musaceae, Costaceae and Cannaceae. Floral diagrams are positioned in the same orientation for easier comparison across families. Fertile stamens are represented by dark grey circles; while * represents an aborted stamen. In Costaceae and Cannaceae, the petaloid appendage is also depicted in dark grey, attached to the single fertile stamen of these flowers. Cannaceae and Costaceae staminodes are depicted as stripped laminar structures in equivalent position to the fertile stamens in Musaceae. The light grey circle in the centre of each floral diagram represents the gynoecium, while the outer structures represent the sepals and petals of the perianth.

CHAPTER 2

Tracking the Development of the Petaloid Fertile Stamen in *Canna indica*: Insights into the origin of androecial petaloidy in the Zingiberales

The following chapter (excluding the preface) has been published as a peer reviewed article in the Annals of Botany PLANT:

Almeida, A.M.R.; Brown, A.; Specht, C.D. (2013) Tracking the development of the petaloid fertile stamen in *Canna indica*: insights into the origin of androecial petaloidy in the Zingiberales. AoB PLANTS 5, doi:10.1093/aobpla/plt009

PREFACE

I started my studies in Evolutionary Developmental Biology of flowers of the order Zingiberales by taking an approach that combines classical plant morphology technics with more recent advancements on the field of molecular development, in order to settle long lasting questions in floral morphology of this group of plants. One such question regards the development of the petaloid appendages in the remaining fertile stamen of the ginger families (Marantaceae, Cannaceae, Zingiberaceae, and Costaceae). Previous morphological analyses of species of Costaceae have suggested that the petaloid appendages of fertile stamens in this family develop in the position of the connective, as a laminarization of an otherwise radially symmetric filament. The question remained as to whether the petaloid appendages of the half fertile stamen of Cannaceae and Marantaceae also derived from the position of the connective or whether the petaloid appendage was a result of a homeotic conversion of the remaining infertile theca into a petal-like organ. Answering this question would result in a better understanding of the homology between the structures found in the fertile stamens of the ginger families, allowing us to pursue the molecular basis for the evolution of such traits. The paragraph that follows summarizes the research presented in this chapter.

Flowers of the order Zingiberales demonstrate a remarkable trend of reduction in the number of fertile stamens; from five or six fertile, filamentous stamens bearing two thecae each in Musaceae and Strelitziaceae to just a single petaloid stamen bearing a single theca in Cannaceae and Marantaceae. As one progresses from the ancestral to the more derived floral forms within the Zingiberales, the fertile stamens are replaced by petaloid staminodes in 4-5 of the six androecial members. In Cannaceae and Costaceae, all members of the androecial whorls exhibit petaloidy, including the fertile stamen. In Costaceae, for example, the single fertile stamen develops two thecae embedded on a broad petaloid appendage.

The most extreme case of petaloidy is exhibited in Cannaceae where the single fertile stamen is further reduced to a single theca with a prominent, expanded petaloid

appendage. Although Costaceae and Cannaceae are not sister lineages, they are closely related within the monophyletic ginger clade. However, whether petaloidy of the fertile stamen is a synapomorphy of the entire ginger clade (incl. Cannaceae, Costaceae, Zingiberaceae and Marantaceae), or the result of independent convergent evolution in Cannaceae, Costaceae, and some Zingiberaceae is unclear.

We combine a developmental series of the formation of the petaloid fertile stamen in *Canna indica* with data on the expression of B- and C-class floral organ identity genes to elucidate the organogenetic identity of the petaloid stamen and staminodes. Our data indicate that the single fertile theca in *C. indica* and its petaloid appendage are derived from ½ of the primordium of a single stamen, with no contribution from the remaining part of the stamen (i.e. the second theca primordium) which aborts early in development. The petaloid appendage expands later, and develops from the position of the filament/connective of the developing theca. Floral identity gene expression shows that petal identity genes (i.e., B-class genes) are expressed in all floral organs studied while C-class gene AG-1 is expressed in an increasing gradient from sepals to gynoecium, and AG-2 is expressed in all floral organs except for the petals.

The canonical model for molecular specification of floral organ identity is not sufficient to explain petaloidy in the androecial whorl in *Canna* sp. Further studies understanding the regulation of gene networks are required.

INTRODUCTION

The Zingiberales are a group of herbaceous tropical monocots comprising eight families and 2500 species. They diverged from their sister order Commelinales (Bremer et al. 2009) approximately 80 million years ago. In Zingiberales, the flowers are organized into five distinct whorls of three organs each: calyx (consisting of three sepals), corolla (consisting of three petals), two androecial. whorls for a total of six (three inner and three outer) stamens and the tripartite gynoecium (Kirchoff 1983).

The Zingiberales order has been traditionally divided into two groups based on overall floral morphology: the banana families, including families Musaceae, Lowiaceae, Strelitziaceae and Heliconiaceae, and the derived ginger families, a monophyletic lineage containing families Costaceae, Zingiberaceae, Marantaceae and Cannaceae (Fig. 1A). Most major evolutionary changes in floral morphology that define these two groups occur in the petal and stamen whorls. In particular, there is an impressive reduction in the number of fertile stamens across the order, from 5–6 fertile stamens in the banana families to a single fertile stamen in Costaceae and Zingiberaceae, and a half fertile stamen in Cannaceae and Marantaceae (Kirchoff et al. 2009). In the flowers of the ginger families, three to five infertile members of the androecial whorls develop as sterile petaloid structures (Kirchoff 1991).

In most Zingiberales flowers, the fertile stamens produce two mature pollen sacs or thecae. In the banana families, these fertile stamens have a narrow connective and thus are filamentous in form. Any petaloid members of the androecial whorls of the banana families are infertile, completely lacking thecae (Kirchoff et al. 2009). However, in the ginger clade a petaloid appendage can develop from the filament or connective of the fertile members of the androecial whorl (Fig. 1B) (Kirchoff 1991; Glinos and Cocucci 2011). This results in the potential for all members of the androecial whorls, whether fertile or sterile, to develop petaloidy. In Costus scaber, the anther consists of two locules, positioned adjacent to each other on the ventral surface of a petaloid structure in the inner androecial whorl (Kirchoff 1988).

Development of the petaloid component of the fertile stamen, which includes both filament and connective, is simultaneous with the development of the anther (Kirchoff 1988). The stamen primordium is divided into two parts—the ventral portion produces the anthers and the dorsal portion produces the petaloid filament and connective (Kirchoff 1988). Conversely, in the Zingiberaceae (sister to Costaceae; Fig. 1A), Leinfellner characterized the petaloid component of the fertile stamen as occurring late in development, thus classifying the petaloid portion as an accessory structure (Leinfellner 1956) and implying lack of homology between the petaloid structures in the fertile stamens of Costaceae and Zingiberaceae.

The concentric androecial whorls of *Canna indica* consist of 3–4 petaloid staminodes (sterile) and one-half of a single fertile petaloid stamen (Glinos and Cocucci 2011). The fertile stamen, labellum and inner staminode constitute the inner androecial whorl, while the outer androecial whorl is made up of the two (or sometimes one) remaining

staminodes (Eichler 1875; Rao and Donde 1955; Pai 1963; Kirchoff 1983). According to Kirchoff (1988, 1991), the fertile stamen is always found in the inner androecial whorl, which develops before the outer androecial whorl. However, the developmental origin of the petaloid appendage of the fertile stamen in Cannaceae remains unclear.

Our understanding of the molecular basis of floral development has increased greatly since the first descriptions of the genes responsible for specifying the identity of floral organs in Antirrhinum and Arabidopsis (Bowman et al. 1991; Jack et al. 1992). According to the canonical ABC model of floral development (Weigel and Meyerowitz 1994), differential gene expression results in the specification of the identity of the various floral organs. In Arabidopsis, A-class genes [APETALA2 (AP2) and APETALA1 (AP1)] are involved in the specification of sepals (first whorl organ), and together with B-class genes [GLOBOSA (GLO) or PISTALLATA (PI), and DEFICIENS (DEF) or APETALA3 (AP3)] they specify petal identity (second whorl). B-class genes are also involved in the specification of stamen identity when expressed together with the C-class gene [AGAMOUS (AG)]. Furthermore, AG alone is responsible for the specification of carpel identity (Coen and Meyerowitz 1991).

Although most components of the ABC model of floral development hold true for most model species studied so far, it is unclear to what extent this model can explain the morphological diversity and evolution of floral development across angiosperms. In the case of monocots, the most well-studied systems are among the grasses where the highly derived flower morphology of the Poaceae renders statements of homology a difficult task. In C. indica, it is unclear whether the petaloid appendage of the half fertile stamen is produced by the secondary expansion of residual meristematic tissue from the filament of a single fertile theca, or whether it is a result of a homeotic transformation of one of the thecae into a petaloid structure. Here, we use developmental studies to characterize the origin of the petaloid tissue in the Canna stamen and investigate whether the combinatorial expression of MADS-box genes can explain petaloidy in C. indica androecial whorls.

METHODS

Developmental series

Living material of Canna sp. was collected from the UC Berkeley Botanical Gardens, the Specht Lab diversity collection at the Oxford Tract Greenhouses, from residential neighbourhoods in the Berkeley hills (with consent from home-owners) and from the UC Berkeley Student Organic Garden (SOGA) (Table 1). In total, 30 inflorescences were collected from C. indica (18), Canna edulis (4), Canna tuerckheimii (4) and Canna sp. (4). Although several Canna species were observed in order to characterize any potential differences across Cannaceae, the developmental series portrayed and the molecular characterization focus specifically on the development of C. indica.

Inflorescences were dissected from living material, removing the outer bracts to expose most floral buds and floral organ primordia at the inflorescence apex. The apices were

vacuum infiltrated for 10–20 min in formalin– acetic acid–alcohol (FAA) (3.7 % formaldehyde), and stored in cold FAA for up to 2 weeks. Tissue fixation was carried out using a standard microwave procedure (Schichnes et al. 1999) as follows: three rounds of 15 min microwave cycles at 37 8C, followed by an ethanol. dehydration series (50, 70, 95 and 100 % ethanol) for 5 min at 67 8C for each ethanol concentration. Tissue was stained in 1 % w/v fast green FCF in 100 % ethanol for 2–3 days at 4 8C. Subsequently, tissue was destained with 100 % ethanol for 2–5 days at 4 8C, as necessary for final dissection, observation and photography (Sattler 1968).

Inflorescences were further dissected under an Olympus dissecting scope, and photographs were taken using a °—3.8 Ultrapak epi-illumination objective (Posluszny et al. 1980; Charlton et al. 1989) on a Leitz Orthoplan microscope equipped with a Nikon Digital Sight 5M digital camera, as described by Bartlett et al. (2008). NIS Elements software was used to process the images taken at different focal points (Bartlett et al. 2008) to expand the depth of focus.

Gene expression

Canna indica flowers were dissected from the same plants as used above. Fresh flowers were quickly dissected, separating sepals, petals, staminodes, petaloid part of the fertile stamen, anther of the fertile stamen and gynoecium into separate vials. RNA was extracted from each of the floral parts individually. RNA extraction was carried out fromfresh tissue with Plant RNA Reagent (Invitrogen), according to the manufacturer's guidelines. cDNA was synthesized after DNase treatment of each sample (Fermentas) using a BIO-RAD iScript Reverse Transcription Supermix kit with poliT primers.

Reverse transcription (RT) primers were designed for AGAMOUS-1 and AGAMOUS-2 AG-1 and AG-2), DEFICIENS (DEF), and GLOBOSA-1 and GLOBOSA-2 (GLO-1 and GLO-2). GLO sequences were downloaded from NCBI (GU594924.1 and GU594945.1) and used for RT primer design. DEF and AG genes were first amplified using degenerate primers. Polymerase chain reaction (PCR) products were cloned into Top10 cells and sequenced using an ABI Big Dye Terminator kit on a 3700 sequencer. Recovered sequences were used to develop copy-specific RT primers. Primer sequences are as follows: GLO1 Forward, CCC TTC CAC GTT ATC GAC GAT T; GLO2 Forward, CGT CCA CCT CGT TGT CTG AG; GLO Reverse, TTG TGC ATC TTC CAA ATC TCC; DEF Forward, CCT CCA CTG AAA CAA AGA AGA TT; DEF Reverse, CAG TTC ATG CAG CAA GTT CC; AG1 Forward, AGC CTA TGA ATT GTC GGT CTT G; AG1 Reverse, AGC TGA GAG ACT CAC CCA TCA; AG2 Forward, CGT ACG AAT TGT CCG TGC TT; AG2 Reverse, TCT GCT CTC GAG TTG CTT CA.

Reverse transcription-polymerase chain reactions were carried out using a Phire DNA Polymerase kit (Finnzymes) and the following: 2 mL of 5°— Phire buffer; 0.2 mL of 10 mM dNTPs; 0.5 mL of each primer; 0.1 ml of Phire polymerase; 1 mL of [1:10]cDNA; and ddH2O, for a total volume of 10 mL. Thermocycling. conditions followed the manufacturer's recommendations, and the following annealing temperatures for 30 cycles: GLO1, 668C; GLO2, 688C; DEF, 638C; AG1 and AG2, 70 8C. Reverse

transcription-polymerase chain reactions were visualized on 1 % agarose gels, and stained with GelRed™ (Phoenix Research Products) according to the manufacturer's protocol.

RESULTS

Canna indica fertile stamen development

Canna indica early floral development has been described previously (Kirchoff 1983). Here, we present only our new developmental data focused on the fertile stamen in order to understand the origin of the petaloid appendage. Therefore, early stages of floral development are only briefly discussed. The earliest discernible stage in *C. indica* floral development (Stage 1; Fig. 2A) is represented by the development of two meristematic bulges, previously described as the sepal primordia (Kirchoff 1983). As the floral bud continues to develop, the apex flattens out, forming a disc-shaped structure, the 'floral cup' (Stages 2, 3; Fig. 2B). The periphery of the floral cup continues to grow and differentiate, eventually becoming delineated into the distinct petal and stamen primordial (Stages 4, 5; Fig. 2C, D).

At about Stage 6 (Fig. 2E–G), the young fertile stamen protrudes out of the floral cup, distinguishing itself from the young petals. These observations are consistent with *Canna* floral development that has been well documented and described until Stage 6 (Rao and Donde 1955; Pai 1963; Kirchoff 1983). Stages 7 and 8 (Fig. 2H, I) depict the continued growth of the fertile stamen and the determination of organ identity. By Stage 7, the rapid development of the fertile stamen and its accompanying petaloid appendage becomes evident, and becomes a distinct feature in the floral bud (Fig. 2H). One theca continues to develop, while the other becomes comparatively reduced in size and discontinues growth or expansion (Fig. 2H).

The petaloid appendage is connected to the developing theca along the filament and apparently below where the connective would normally develop (Fig. 2I). Owing to the abortion of the second theca, no connective region is apparent. At Stage 8 (Fig. 2I), the final stage of this developmental series, the nearlymature fertile stamen is represented by a single developed theca that is connected to a rapidly expanding petaloid appendage emerging from the filament. A line of cleavage separates the aborted theca from the growing fertile theca with its petaloid appendage.

Gene expression during floral development

Reverse transcription-polymerase chain reaction for *C. indica* was used in order to assess the expression pattern of B- and C-class MADS-box genes in various floral organs (Fig. 3). Sepals (sep), petals (pet), staminodes (std) and gynoecium (gyn) were studied in their entirety. For a better account of gene expression patterns on *Canna* organs, the fertile stamen was divided into petaloid appendage (pap) and theca (the), which were studied independently. *Canna indica* has at least one copy of DEFICIENS (DEF), two copies of GLOBOSA (herein referred to as GLO-1 and GLO-2) and two copies of AGAMOUS

(AG-1 and AG-2) (A. M. R. Almeida et al., unpubl. data). B-class MADS-box genes (DEF, GLO-1 and GLO-2) are expressed in all the floral parts studied (Fig. 3).

It is interesting to note that expression of these genes is reduced in sepals, especially for DEF and GLO-1. B-class gene expression shows an expanded pattern when compared with the Arabidopsis ABC model, where expression of the B-class genes is restricted to petals and stamens. C-class MADS-box genes (AG-1 and AG-2) also show an expanded pattern of expression when compared with the expected expression pattern based on the canonical ABC model (Fig. 3): AG-1 seems to be expressed in a gradient, increasing from sepals (low) to gynoecium (high), while AG-2 is evenly expressed in all floral parts studied with the exception of the petals, where no expression was observed.

DISCUSSION

The initial stages of organogenesis in this developmental series confirm past studies on *Canna* floral development (Rao and Donde 1955; Pai 1963; Kirchoff 1983). Here we focus on the development of the fertile stamen with particular attention given to its petaloid appendage. Petaloidy is a striking trend in the evolution of Zingiberales floral morphology, especially in the ginger clade where the number of fertile stamens is drastically reduced and the remaining infertile androecial members are petaloid.

The extreme case is observed in Cannaceae flowers, in which all androecial elements are petaloid and the half fertile stamen has a marked petaloid appendage (Fig. 1B). In this case, only one theca is apparent at anthesis, and the question remains whether (i) the petaloid appendage of the fertile stamen develops from the filament and connective of the same primordium that gives rise to the single theca, or (ii) the appendage is the result of the growth and expansion of a separate theca primordium that undergoes homeotic transformation into a sterile, petaloid structure. In the first case, only half of the original stamen primordium would develop fully, forming both the anther and the petaloid appendage (see Fig. 4B, x). In the second case, the entire stamen primordium would grow and mature, with one-half giving rise to a petaloid structure and the other half forming an anther.

The morphological series presented here (Fig. 2G–I) provides evidence for the first hypothesis: that the petaloid appendage of the *Canna* fertile stamen develops from the same primordium that produces the theca, emerging from the position of the filament. This finding has implications for understanding fertile stamen development in other genera within the ginger lineage. For instance, because it appears that the entire structure (theca and petaloid appendage) is produced from a single half (stamen) primordium, other fertile stamen configurations, such as those observed in Costaceae and Zingiberaceae, could very probably result from concerted laminar development of the filament and connective associated with both fertile thecae. In order to investigate the molecular mechanisms associated with androecial petaloidy in *C. indica*, the expressions of class B and class C MADS-box genes were analyzed in various floral organs. We did not investigate A-class gene expression, as the role of the A function genes outside of Arabidopsis is unclear; alternatively, B and C function has been shown to predict the

stamen and petal development model for several groups of monocots (Kim et al. 2006; Tang et al. 2007).

The canonical expression pattern for B- and C-class MADS-box genes (Fig. 4A) does not appear to hold for *Canna*. We expected to find B-class genes in the petal and stamen whorls, and C-class genes in stamen and gynoecium whorls, with perhaps some changes in expression defining the differences between petaloid vs. fertile stamens within the androecial whorls. Instead, B-class (GLO and DEF) genes have expression domains that are expanded in both directions to include the first whorl and the gynoecium. C-class (AG) genes also show a broad expression pattern and are found in petals (AG-1) and sepals (AG-2) as well as the androecial and gynoecium whorls (Fig. 4B). There was no differentiation between fertile or sterile elements within the androecial whorl, nor was there a combination that seemed to define petaloidy regardless of its whorl of origin.

As petaloidy in *C. indica* is, however, not restricted to the corolla (petal) and androecial (stamen) whorls, the extension of B-class gene expression into the gynoecium might explain the laminar morphology of the carpels in *Canna* (Glinos and Cocucci 2011). Most of the *Canna* flower shows simultaneous B- and C-class MADS-box gene expression, which in the classical ABC model would result in the specification of stamen identity. Clearly, this combination is not functioning as stamen identity in the *Canna* flower, with its single half fertile stamen. This expression pattern implies that *Canna* petaloidy, whether in the petals, stamens or even the carpels, is probably not a simple result of re-deployment of the classical petal specification mechanisms (A- and B-class MADS-box gene expression), and potentially involves an as yet uncharacterized molecular basis.

Considering the origins of stamens from a petal-like organ (Goethe 1790), it is possible that the filamentous stamens that are ancestral to Zingiberales and that characterize Musa flowers are the result of a restriction of laminar growth associated with the development of fertility. The lack of pollen sac production in the majority of androecial members of the ginger families might cause a de-repression of laminar growth, resulting in the production of petaloid organs in the androecial whorls. When petaloidy is found in organs that do contain fertile thecae, it is unclear as to the mechanisms that enable laminar growth in the presence of pollen sac production. Current studies are focusing on the role of polarity genes that establish the abaxial/adaxial boundary and regulate laminar vs. radial morphology of lateral organs.

CONCLUSIONS

It is possible to conclude, based on the data from this study, that the developmental mechanisms resulting in petaloid floral organs are different even in closely related taxa such as Cannaceae and Costaceae. It appears that the development of a petaloid appendage on the filament of a single theca in *C. indica* might be the result of ectopic development resulting in the appearance of a half fertile, petaloid stamen. In contrast, in Costaceae the petaloid stamen might be the result of laminar growth of the filament and connective, returning to an ancestral leaf-like laminar development as seen in the petaloid

stamens of early diverging angiosperms (e.g. Nymphaea). Investigations on candidate gene expression during development of the stamens in Costaceae and Cannaceae will be necessary to determine if the genetic mechanisms underlying the development of the petaloid stamens are indeed different in these two families, indicating that homoplasy can be at work even in closely related species.

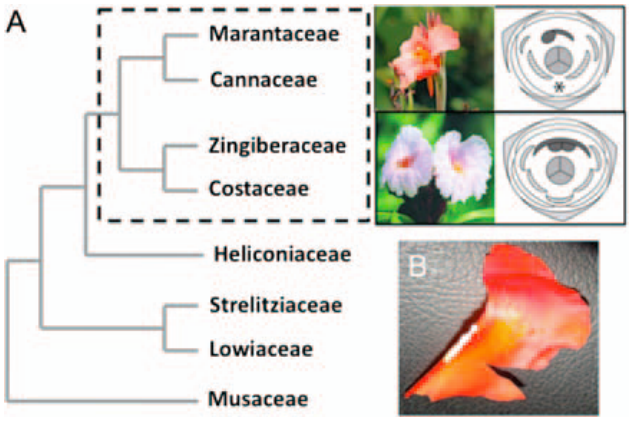

Figure 1. Phylogenetic context for studying comparative organogenesis in Zingiberales. (A) Zingiberales phylogeny according to molecular and morphological characteristics (Kress 1990; Kress et al. 2001). The dashed square highlights the ginger clade, comprising a monophyletic group of four families (Costaceae, Zingiberaceae, Cannaceae and Marantaceae). Photographs: *C. indica* (top); *Costus spicatus* (bottom). On the right, floral diagrams representative of flowers of the Cannaceae (top) and Costaceae (bottom) families. Light grey, sepals; white, petals; hashed, petaloid staminodes; dark grey, fertile stamen; *, aborted stamen; centre grey, gynoecium. (B) *Canna indica* half fertile stamen with petaloid appendage.

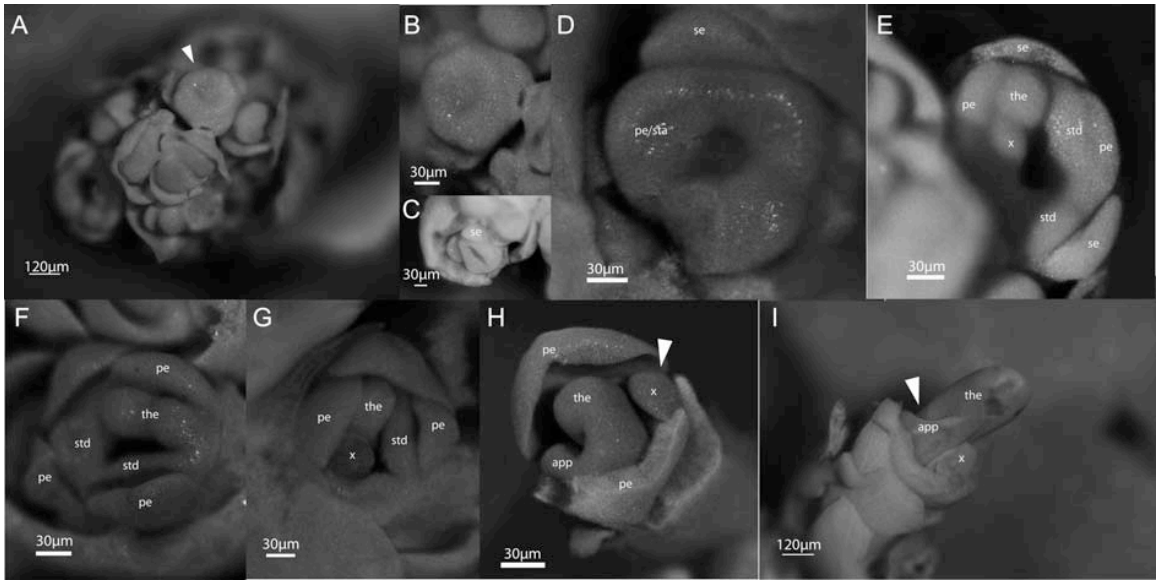

Figure 2. *Canna indica* floral development series. (A) Floral initiation showing the protrusion of the sepal primordial. The arrowhead points to a floral primordium amplified in (B); (B, C) development of the 'floral cup'; (D) sepal primordia already separated from the remaining floral primordium, and evident petal primordia; (E) early stages of fertile stamen development, with two theca primordia; (F, G) fertile stamen development; (H) later stages of fertile stamen development. A single theca has developed with its petaloid appendage, while the other theca arrests development (arrowhead); (I) an almost mature stamen with its petaloid appendage (arrowhead); and the aborted theca to its right. se, sepal; pe, petal; pe/sta, petal/stamen common primordium; std, staminode; the, fertile theca; app, fertile stamen appendage; x, aborted theca primordium.

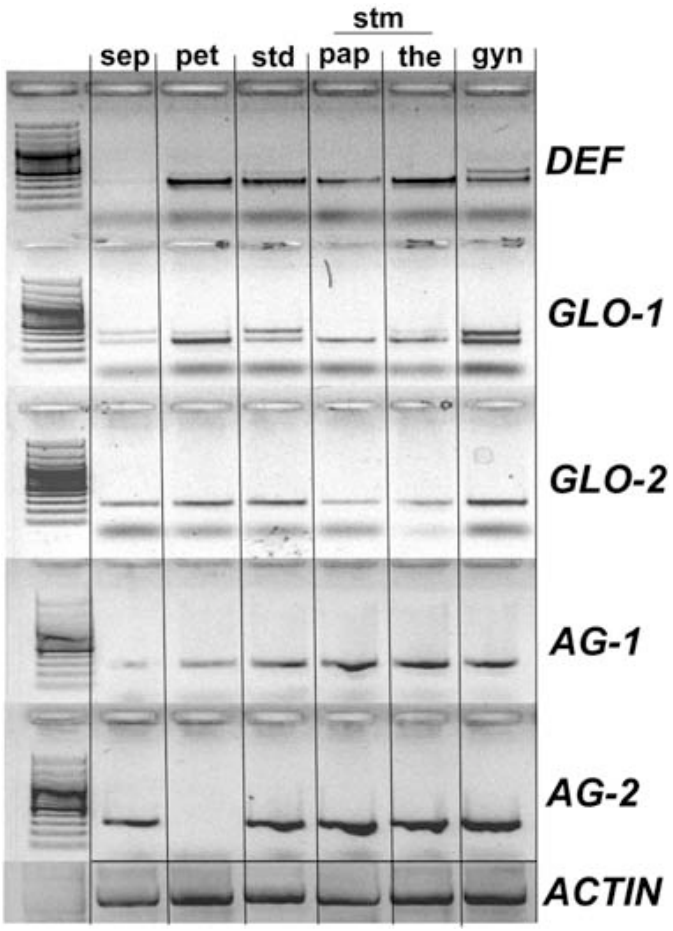

Figure 3. Expressions of B- and C-class MADS-box genes in the floral organs of *C. indica* as detected by RT-PCR. Each *C. indica* floral organ was dissected and RNA was extracted independently. The fertile stamen was divided into petaloid appendage and theca. sep, sepal; pet, petal; std, staminode; stm, stamen; pap, petaloid appendage of stamen; the, theca; gyn, gynoecium. Actin was used as an endogenous control for the cDNA synthesis. B-class genes: DEF, DEFICIENS; GLO-1, GLOBOSA-1; GLO-2, GLOBOSA-2. C-class genes: AG-1, AGAMOUS-1; AG-2, AGAMOUS-2.

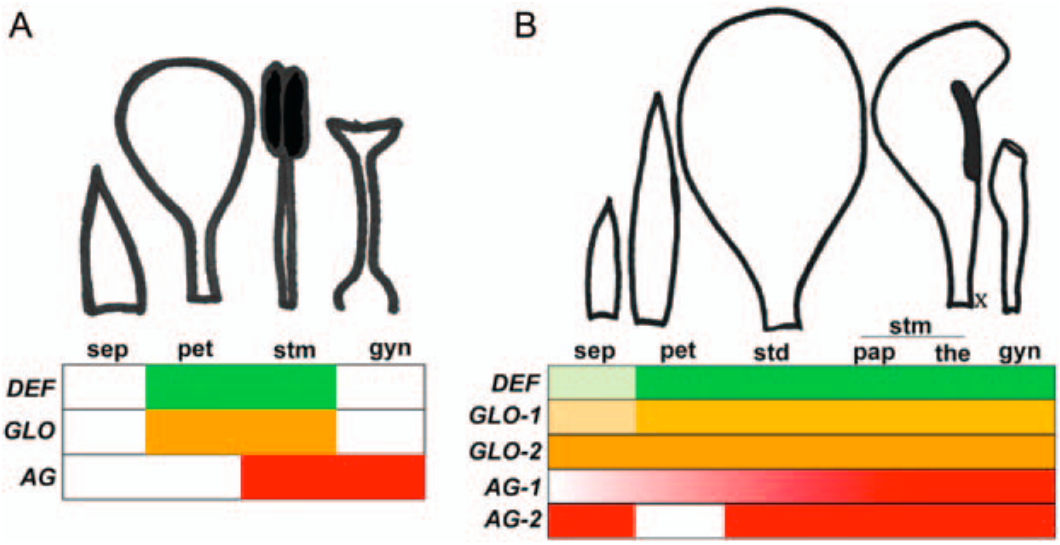

Figure 4. Summary results for gene expression and the corresponding floral organ morphology in *Arabidopsis* and *Canna*. (A) Classical ABC model of floral development based on *Arabidosis thaliana*. Only B-class (DEFICIENS and GLOBOSA) and C-class (AGAMOUS) MADS-box genes are depicted, as the role of A-class MADS-box genes in floral development in monocots awaits further investigation. In the classical ABC model, petal identity is a result of A- and B-class MADS-box gene expression, while stamen identity results from concomitant expression of B- and C-class MADS-box genes. (B) *Canna indica* B- and C-class MADS-box gene expression pattern. *Canna indica* contains two GLOBOSA genes (GLO-1 and GLO-2) and two AGAMOUS genes (AG-1 and AG-2). B- and C-class MADS-box genes are expressed in most of the floral parts studied here, and when compared with the classical ABC model, show an expansion in their expression domains. x, position of the aborted theca primordium relative to the half fertile stamen.

Table 1. Accession of *Canna sp*. used in morphological and molecular studies of floral developmental evolution.

Accession no.	Voucher location	Species	Fig. 2
AB006	SOGA	*Canna indica* L.	2H
AB009	UC Botanical Gardens	*Canna edulis* Ker Gawl.	2C
AB017	UC Botanical Gardens	*Canna indica* L.	2G
AB020	SOGA	*Canna indica* L.	2A, B, D–F, I

CHAPTER 3

Positive selection on the K domain of the *AGAMOUS* protein in the Zingiberales suggests a mechanism for the evolution of androecial morphology in the order

PREFACE

The second chapter of this dissertation established the homology between the petaloid appendages of the fertile stamen across the ginger families, while pointing out limitations of the current molecular model of organ identity to explain the identity floral organs in *Canna indica*. In this chapter, we explore the evolution of one of the floral organ identity genes, the AGAMOUS gene (AG), in order to shed light on the potential molecular mechanisms shaping the evolution of floral morphology in the Zingiberales. In particular, we focus on the extensive androecial petaloidy that evolves at the base of the ginger clade as a derived characteristic of these lineages. The increased petaloidy of androecial members is inversely correlated to the number of fertile androecial members (fertile stamens) in the ginger families, reaching an extreme of one half of a fertile stamen in Marantaceae and Cannaceae.

In turn, AG has been implicated in the development of sex organ, as well as in meristem determinancy, in model plants such as *Arabidopsis thaliana*. In this plant, intermediate levels of AG lead to a variety of floral phenotypes, including the development of petaloid stamens, while severe AG mutants leads to a complete conversion of stamens and carpels into petals and sepals, respectively. These Arabidopsis mutants suggest AG as a good candidate gene potentially involved in androecial petaloidy in the Zingiberales.

Please, note that this chapter has not yet been published elsewhere. Below is a summary of the motivations, methods, and results described in this chapter.

The classical ABC model of floral development describes, at least in model species such as *Arabidopsis thaliana*, the molecular basis for organ identity in the angiosperm flower. Expression of C-class genes is tightly linked to stamen (together with B-class genes) and gynoecium organ identity. In *Arabidopsis*, the C-class is represented by a single copy of the gene *AGAMOUS* (*AG*). Missexpression of AG in *Arabidopsis* results in the formation of petaloid stamens.

The Zingiberales is an order of tropical monocots in which the evolution of floral morphology is characterized by a marked increase in petaloidy in the stamen whorl such that androecial petaloidy can represent the bulk of the floral display. In the Zingiberales, petaloidy is associated with increased number of infertile stamens (staminodes) that develop as laminar structures. The eight families of Zingiberales are roughly divided in two groups: the paraphyletic “banana families” and the derived monophyletic “ginger families”. Petaloidy is a derived characteristic of the ginger families, and seems to have

arisen at the base of the ginger clade. We hypothesize that gene duplication followed by gene divergence of *AGAMOUS* in the ginger clade explains the evolution of petaloidy in the androecial whorl.

Our results present an intricate story in which duplication of the *AGAMOUS* lineage has lead to the retention of at least two diverged Zingiberales-specific copies, *ZinAG-1* and *ZinAG-2*. While *ZinAG-2* forms a clade of highly similar sequences, *ZinAG-1* sequences across Zingiberales are highly diverged from one another. Positive selection is observed in the K and I domains of Zingiberales *AGAMOUS*. In particular, positive selection in the third alpha-helix of the K domain suggests a mechanism by which *AGAMOUS* gene divergence may explain observed morphological changes in Zingiberales flowers.

INTRODUCTION

The genetic control of flower morphogenesis has long been studied in *Arabidopsis thaliana* and *Antirrhinum majus* (Coen & Meyerowitz 1991). Classically, floral organ specification has been described by combinatorial patterns of gene expression in what has been well known as the ABC model of floral organ identity. The domains of expression of A-, B-, and C-class genes correlate with the appropriate position of the developing sepals (A-class genes only), petals (a combination of A- and B-class genes), stamens (a combination of B- and C-class genes), and gynoecium (C-class genes only). In *Arabidopsis thaliana*, there are two A-class genes (*APETALA-1* (*AP1*), and *APETALA-2* (*AP2*)), two B-class genes (*APETALA-3* (*AP3*), *PISTALLATA* (*PI*)), and one C-class gene (*AGAMOUS* (*AG*)). With the exception of *AP2*, all other genetic components of the ABC model are Type-II MIKC^c MADS-box genes as determined by their arrangement of protein domains. Their proper function as transcriptional regulators is dependent on the protein-protein interactions that occur between the A, B, and C-class genes, as well as with the *SEPALLATA* genes (Pelaz et al. 2000), to form protein dimers and functional tetramers. The protein-level explanation for A-, B-, and C-class function is known as the Quartet Model, and asserts that only in tetramers are A-, B-, and C-class proteins capable of regulating downstream genes (for review, Melzer & Theissen 2009).

Although this classical approach has proven fruitful in describing organ identity and floral organ development, it lacks a mechanism to describe how such robust gene expression patterns are established during development and may correlate with evolutionary changes in organ identity. In order to address this mechanism, the A-, B-, and C-class genes have been integrated into an elegant complex-system approach, capable of explaining the robustness of the ABC gene expression patterns during flower development (Mendoza & Alvarez-Buylla 1998; Mendoza et al. 1999). By mapping the landscape of known gene interactions during floral development in *Arabidopsis thaliana*, Mendoza and coworkers (1999) were able to recover the stable states (i.e., attractors) that correspond to the gene expression patterns correlated to floral organ identity, as described by the ABC model. In doing so, the authors provided a set of necessary and sufficient genes and their interactions capable of providing a dynamical and mechanistic explanation for the establishment of the ABC gene expression patterns (Alvarez-Buylla et al. 2010a, Alvarez-Buylla et al. 2010b). Within the mapped gene interactions that constitute the organ identity gene regulatory network (Figure 1), the *AGAMOUS* gene is one of the most highly interconnected genes, thus alterations in this node are likely to constitute important changes in the stable states of the system, leading to morphological variation and providing a potential nexus for evolutionary changes. It is also interesting to note that stamen and petal attractors differ exclusively by opposite states of *AP1* and *AG* expression; *AP1* is expressed in petals but not in stamens, while *AG* is expressed in stamens but not in petals (Barrio et al. 2010).

The *AGAMOUS* gene was first isolated from *Arabidopsis thaliana* over two decades ago (Yanofsky et al. 1990), when fully penetrant mutations were shown to cause abnormalities in the development of the floral reproductive organs. *AGAMOUS* has since been implicated in proper development of reproductive organ identity across flowering

plants, and is thought to play an additional role in ovule development and meristem determinancy in some lineages (Yamaguchi et al. 2006; Dreni et al. 2011). *AG* itself is part of a large subfamily of transcription factors, with four copies in both *Arabidopsis* and in *Oryza* (Kramer et al. 2004; Yamaguchi et al. 2006; Dreni et al 2011) indicating that, throughout the evolution of flowering plants, the *AG* subfamily has undergone several duplication events followed by gene retention and subfunctionalization.

Phylogenetic analyses of the AG subfamily have demonstrated that a duplication event early during angiosperm evolution resulted in the origin of two major lineages: the AG and the AGL11 lineages (Kramer et al. 2004; Zahn et al. 2006). In the *Arabidopsis* ABC model, the AG and AGL11 lineages correspond functionally to class C and class D homeotic genes, respectively, in which C-class homeotic genes are involved in stamen and ovule identity and D-class genes are more specifically involved in ovule and fruit development. The AG lineage itself has undergone subsequent gene duplications, and parsing of function of duplicated copies is thought to have occurred independently in several angiosperm lineages. In *Antirrhinum*, the *AG* lineage genes *PLENA (PLE)* and *FARINELLI (FAR)* contribute unequally to specify male and female reproductive organs (Davies et al 1999; Airoldi et al 2010), while in petunia *FPB6* and *PMADS3* act redundantly as C-function genes (Heijmans et al. 2012). In *Zea mays,* however, the *AGAMOUS* paralogs *ZAG1* and *ZMM2* appear to be expressed in spatially distinct domains of the developing flower, and might have subfunctionalized into carpel- and stamen-specific paralogs (Mena *et al*. 1996), while in rice, *AGAMOUS* paralogs *OSMADS3* and *OSMADS58* are essential for reproductive organ identity and together with AG lineage *OSMADS13* are important for floral meristem determinancy (Yamaguchi et al. 2006; Dreni et al. 2011).

In addition to its role in reproductive organ identity, differential expression of *AGAMOUS* in *Arabidopsis thaliana* has also been shown to be involved in the development of petaloidy in the androecial whorl. For instance, it has been reported that the ag-11 mutant allele, a single point mutation in the regulatory region of *Arabidopsis AG*, results in the transformation of stamens into petaloid organs (Hong et al. 2003). Also, down-regulation of *AGAMOUS* by anti-sense RNA in *Arabidopsis thaliana* leads to a variety of floral morphologies, including petaloid stamens (Mizukami & Ma 1995).

Although the evolution of androecial petaloidy in the angiosperms is yet to be studied, petal-like stamens are present in a variety of basal angiosperm lineages. In *Amborella trichopoda*, for example, stamen filament is expanded into a petal-like structure. In basal angiosperms several lineages display a gradual transition between petal and stamen, with multiple levels of androecial petaloidy. In basal angiosperms, a gradient of *AG* and B-class gene expression has been implicated in the gradual morphological transition between petals and stamens (Soltis et al. 2007). In order to further explore the role of the *AGAMOUS* gene in the evolution of androecial petaloidy, we focus our study on the Zingiberales, a group of monocots that exhibits extensive androecial petaloidy. Zingiberales are an order of tropical monocots comprising approximately 2,500 species. The order is divided into eight families, phylogenetically organized into the paraphyletic banana families and the derived ginger family clade (Kress et al. 2001). In the

Zingiberales, androecial petaloidy is an important component of floral morphological diversity. Across the order, the derived lineages have a marked reduction in the number of fertile stamens and the infertile androecial members develop as petaloid structures. These petaloid staminodes usually constitute the bulk of floral display, as the petals of the same flower are frequently inconspicuous (Figure 2).

Given the involvement of the *AGAMOUS* gene lineage in reproductive organ development and its interconnectivity within the floral organ gene regulatory network (FOS-GRN, Figure 1), we hypothesize that gene duplication followed by gene divergence of the *AGAMOUS* lineage in the ginger clade is correlated to the evolution of petaloidy in the androecial whorl. In order to test our hypothesis, *AG* lineage genes were amplified from across the Zingiberales, and their expression was assessed during flower development in representative species. Tests of selection were carried out to investigate the role of selection on gene evolution and function. Our results suggest that positive selection has played a role in the evolution of *AG* across the Zingiberales order, particularly in protein divergence within the K domain. These protein modifications suggest a mechanism by which androecial petaloidy may have evolved in the Zingiberales, supporting the hypothesis that modifications in *AG* expression and function are correlated with androecial petaloidy.

METHODS

Plant material, RNA extraction and cDNA synthesis.

Twenty species from seven of the eight Zingiberales families were sampled in order to represent the diversity of floral form observed in the order (Table 1). Fresh flowers were collected and stored in "RNA-later" (recipe available upon request) for up to two weeks prior to RNA extraction. Total RNA was extracted from floral material using Plant RNA Extraction Reagent (Invitrogen, Carlsbad, CA, USA), according to the provided protocol. RNA was stored at -80°C until further use. Prior to cDNA synthesis, cDNA was treated for DNA contamination (DNAse treatment Fermentas®). cDNA synthesis was performed using iScript select cDNA synthesis Kit (Bio-Rad, Hercules, CA, USA) and poliT primers. Amplification of the ß-actin gene as a positive control for cDNA synthesis was performed using PCR primers (F: 5'GGA CGA ACA ACT GGT ATC GTG CTG3' and R: GAT GGA TCC TCC AAT CCA GAC ACT GTA3') (Bartlett and Specht 2010). Reactions without reverse transcriptase (no-RT) were used to control for DNA contamination.

Amplification of AGAMOUS genes in the Zingiberales

A multiple sequence alignment (MSA) for the *AGAMOUS* (*AG*) gene lineage was generated from sequences downloaded from NCBI (Table 1). The MSA was used to design general primers for amplification of Zingiberales AG genes. Multiple primer combinations, with different degrees of degeneracy, were used in order to improve chances of assessing all copies of the AG gene lineage within the Zingiberales. Primer

sequences follows. Forward primers: 5' ACI AAY MGI CAR GTI ACI TTY TG 3'; 5' ATG GSI MGI GGI AAR ATI SAR AT 3'; 5' CAR GTK ACC TTC TGC AAG 3'; 5' ATC CCA TGG AGC ATA AAG CA 3'; 5' GRG GRA AGA TCG AGA TCA AG 3'. Reverse primers: 5' ACC CTA TCA GTC TCG GCG ATC TTG TTC C 3'; 5' TCA TCG TTC AAC CAA AGT GG 3'; 5' TTG MAK RAA GTT CCY TGA RTM RT 3'.

PCR reactions were carried out using Phire DNA Polymerase kit (Finnzymes) and: 2µl of 5XPhire buffer; 0.2µl 10mM dNTPs; 0.5µl of each primer; 0.1µl Phire Polymerase; 1µl [1:10]cDNA; and ddH_2O, for a total volume of 10µl. Thermocycling conditions followed manufactor's recommendations. PCR products were visualized on a 1% agarose gel stained with GelRed™ (Phoenix Research Products) according to manufacture's protocol. PCR products were cloned into Top10 cells and sequenced using ABI Big Bye Terminator kit on a 3700 sequencer. At least sixteen clones were sequenced for each of the species sampled. Over forty clones were sequenced for *Costus spicatus*, in order to insure deep sampling of gene copy number in this species.

Gene family evolution and selection tests

Zingiberales sequences were aligned to outgroup sequences downloaded from NCBI (Table 1). A multiple sequence alignment was generated using MacClade 4.06 OS X, and used for downstream phylogenetic analysis. MrModeltest (Nylander 2004) was used to access the best-fit model. The best-fit model was implemented in MrBayes and PHYML in order to assess gene tree topology. Variations on the best-fit model were also tested, and the gene tree topology compared under likelihood and Bayesian frameworks. Posterior probabilities and bootstrap support (from 1,000 replicates) were calculated for MrBayes and ML analysis respectively, and are used as branch support in the gene tree. Likelihood tests for constrained topologies were performed using PAUP* (Swofford 2002).

Selection tests were performed in order to assess signals of selection across the AGAMOUS subfamily. Branch selection was assessed using PAML (Yang 2007) models M2 and M0, while site selection was evaluated using the FEL package in HYPHY (Pond et al. 2005) under a stringent cut-off of 0.1.

Gene expression of *AGAMOUS* and *APETALA1* in Zingiberales developing flowers

Expression of *AGAMOUS* and *APETALA1* was assessed via organ specific transcriptomes. Illumina libraries were prepared using the TruSeq RNA sample prep kit v2. Two libraries each were prepared from *Musa basjoo* stamen filament, theca, and free petal RNA, and *Costus spicatus* stamen filament, theca, petal, and labellum (fused petaloid staminodes). Libraries were multiplexed using barcoding set A. Samples were run on a HiSeq2000 at IIGB HT Sequencing Facility at the University of California, Riverside. Raw reads were trimmed to remove adapters and regions of poor quality with *cutadapt*. *Costus spicatus* sequences were assembled into a reference transcriptome using Trinity. GSNAP was used to align *Musa basjoo* trimmed reads to annotated CDS from the published *Musa acuminata* genome, while *Costus spicatus* trimmed reads were

aligned to the *Costus spicatus* reference transcriptome. Expression of *AG* and *AP1* was estimated using eXpress in units of FPKM (frequency per kilobase of exon per million aligned reads). Replicates were independently processed, and gene expression was compared between libraries for consistency. *ACTIN1* expression was used to normalize targeted gene expression across *Costus spicatus* libraries.

RESULTS

AGAMOUS sequences were obtained for all families within the order, with the exception of Lowiaceae (*Orchidantha*). A multiple sequence alignment (MSA) of 558bp was generated and encompasses all protein domains, with the exception of the first nine codons of the MADS domain and the end of the C-terminal domain, for which alignment to outgroup sequences became increasingly challenging. The final MSA comprises 37 ingroup and 13 outgroup sequences. This MSA was used as the input to MrModelTest, and determined the best-fit model as the GTR+I+G model. The best-fit model, as well as variations of the model (GTR, GTR+I, and K8), was implemented in both MrBayes and PHYML. Tree topology across methods and models was consistent (Figure 3).

All *AGAMOUS* sequences from Zingiberales form a monophyletic group with high support (posterior probability of 1 and 99% bootstrap support from 1,000 replicates). According to the species distribution on the gene tree there are, at least, two copies of the *AGAMOUS* gene in the Zingiberales, herein called *ZinAGcp1* and *ZinAGcp2* (Figure 3B). *ZinAGcp2* sequences form a monophyletic group, suggesting an orthologous relationship between copies found in the banana and ginger lineages, while *ZinAGcp1* shows a more complex evolutionary history. It is interesting to note that although ginger sequences form a paraphyletic lineage at the base of the *ZinAGcp2* clade, *ZinAGcp1* copies of *Costus spicatus* and *Canna sp*. (also ginger species) are found within the first branching lineage, together with banana group sequences (Figure 3).

Based on the phylogenetic analyses, several different evolutionary scenarios could account for the recovered topology for *ZinAGcp1* sequences (Figure 4A). Although they appear paraphyletic in the phylogenetic analysis, these lineages could have evolved from a single duplication event predating the divergence of the Zingiberales, followed by subsequent differential sequence divergence resulting from distinct evolutionary pressures (Figure 4A, Scenario 1). In this case, the phylogenetic analysis would recover an incongruence between the gene tree (placing the lineages as paraphyletic) and the organismal tree, probably due to distinct evolutionary forces acting in specific lineages. Alternatively (Figure 4A, Scenarios 2 and 3), a second duplication event may have occurred in the Zingiberaceae after its divergence from Costaceae, leading to a third lineage specific *AG* copy (cp3) in the Zingiberaceae. This copy would have been retained while the paralogous duplicate was then lost from the Zingiberaceae, yielding only two copies but with less clear orthology to the two copies found in the remaining Zingiberales lineages.

In order to test the likelihood of the different scenarios, we performed a K-H test using a constrained gene tree in which the two first branching lineages (all "*ZinAGcp1*")were forced to form a monophyletic group (Figure 4B), against the unconstrained gene tree (Figure 3). According to the K-H test, the constrained gene tree shows a significantly better likelihood score, suggesting that Scenario 1 (Figure 4A) is the most likely evolutionary history of *ZinAG* copies.

Following a gene duplication event, differential selection may occur on each of the duplicated genes, and that selection can also differ across the gene particularly in regions of coding and non-coding sequence leading to differential divergence. Sequence divergence post duplication may result in the evolution of a new gene function for one of the copies (neofunctionalization), or the partition of function between the two gene copies (subfunctionalization). In these cases, both gene copies will be retained by selection but may undergo significantly different sequence evolution as a result of neutral processes as well as selection. In addition, one of the copies may loose its ability to function, resulting in subsequent loss of the gene or inability to recover the sequence due to the accumulation of random mutations. Given the potential orthology between *ZinAGcp1* and *ZinAGcp2*, selection tests were carried out using the gene tree (Figure 3A), in order to search for differential post-duplication signatures of selection. Branch-selection tests (PAML) show significant positive selection (*w* (omega) = 1.2059; LRT = 1622.95153, p=0.000) at the base of the Zingiberales clade, suggesting that functional divergence between lineages happened early after the duplication event (Figure 5A).

Sites under selection were identified using the FEL package of HyPhy. As expected, most sites are under balancing selection, while three sites show signs of positive selection (Figure 5B). Codon position 75 in the I domain, and codon positions 124 and 142 in the K domain show signs of positive selection (Figure 5B). Comparing these sites between species of the banana group (e.g. *Musa acuminata*) and the ginger group (e.g. *Costus spicatus* and *Canna* sp.), it becomes clear that most of the changes, although fixed between the two groups, do not result in any changes to the chemical properties of the amino acids in these positions (Figure 6). The single exception is the change observed at codon position 142 of *ZinAGcp1*. In *Musa acuminata*, position 142 is occupied by amino acids with charged polar side chains, such as asparagine (N) and histidine (H), while in *Canna* sp. and *Costus spicatus* this position is occupied by tyrosine (Y), an amino acid with an uncharged side chain. Codon position 142 is part of the third alpha-helix of the K domain, also known as K3. The K domain of the MADS proteins are involved in the formation of protein complexes for DNA-binding. In particular, K1 and K2 helices are in involved in dimmer formation, while K3 is involved in the formation of tetramers (Yang & Jack 2004; Kaufmann et al. 2005; Immik et al. 2009; Gramzow & Theissen 2010).

Given the amino acid changes observed in *ZinAGcp1* between the banana group and the ginger group at the K3 position, one can hypothesize that amino acids with uncharged side chains are less likely to chemically react than amino acids with charged side chains. Thus, banana group and ginger group *ZinAGcp1* proteins have different abilities to form higher level complexes, while maintaining similar abilities to form protein dimers. This suggests an interesting mechanism in which *ZinAGcp1* from the ginger group could act as

a negative regulator of tetramer formation: While binding to AG interacting proteins to form dimers, this complex would be less likely involved in the formation of quartets, resulting in a post-transcriptional down-regulation of *AGAMOUS* downstream targets.

In order to test this hypothesis, we investigated the expression of AG genes in *Musa acuminata* and *Costus spicatus*. *AGAMOUS* is expressed in filaments and theca of *Musa*, and in very low levels in the free petal (Figure 7A). In *Costus spicatus, AG* expression is dominated by *ZinAGcp1*, and extremely low levels of *ZinAGcp2* are only observed in the petal (Figure 7B). If one assumes that *ZinAGcp1* in the ginger clade inhibits quartet formation, and thus its expression leads to the down-regulation of downstream targets in *Costus spicatus* (as suggested by the amino acid change; Figure 6), we expect that high levels of *Costus ZinAGcp1* would lead to a stronger suppression of downstream genes, and a more petal-like phenotype in the stamen whorl. The correlation between higher levels of *ZinAGcp1* in *Costus* labellum and the filament and a petaloid phenotype of these organs is consistent with the decreasing levels of *ZinAGcp1* expression towards the fertile theca.

In *Arabidopsis thaliana*, *AG* is known to repress *AP1* expression (Gustafson-Brown et al. 1994); while *AP1* is expressed in petal primordia (to the exclusion of *AG*), *AG* is expressed in developing stamens (to the exclusion of *AP1*). This mutual exclusion expression pattern can also be observed in the modeled Gene Regulatory Networks (Barrio et al. 2010), in which the main difference between stamens and petals as attractors relates to their differential expression of *AG* and *AP1*. To further support the hypothesis that *ZinAGcp1* has evolved to function as a dominant negative regulator of downstream gene expression specifically in the ginger families, we investigated *AP1* expression in comparison to *AG* expression (Figure 7). In *Musa acuminata*, *AP1* expression largely agrees with that anticipated based on a hypothesis of mutual exclusion: *AP1* is mostly expressed in petals where there is very low expression of *AG*, while in stamens *AP1* expression is almost abolished and *AG* is highly expressed (Figure 7A). However, *Costus spicatus* gene expression tells a different story. *AG* and *AP1* are simultaneously expressed in the androecial whorl (labellum, filament, and theca), suggesting that *ZinAGcp1* is not capable of suppressing *AP1* expression in these organs. This lack of *AP1* suppression leads to a mixed petal-stamen attractor, a feature that is likewise observed in transgenic *Arabidopsis* plants with reduced levels of *AG* expression (Mizukami & Ma 1995). In these plants, as well is in gingers, petaloidy is observed in the androecial whorl.

DISCUSSION

The *AGAMOUS* gene subfamily has been extensively implicated in the development of reproductive organs (caperls and stamens) and meristem determinancy in angiosperms. In both monocots and eudicots, the conservation of these functions by *AGAMOUS* lineage genes is remarkable considering multiple gene duplication and subfunctionalization events (Mena et al. 1996; Yamaguchi et al. 2006; Zahn et al. 2006), although *AGL11* lineage genes might act redundantly in some lineages (Heijmans et al 2012; Dreni et al. 2011). In the Zingiberales, at least one lineage-specific duplication event is observed

within the AG lineage. Sequence divergence between the two copies (*ZinAGcp1* and *ZinAGcp2*), as well as their expression patterns, suggests the involvement of Zingiberales *AG* genes in the evolution of reproductive organ development and the evolution of petaloidy in the order.

Early after the Zingiberales-specific duplication event, selection acted differently upon the two *AGAMOUS* lineage copies. Most of the observed selection is due to balancing selection, as expected for genes that have a high degree of interactions that must be maintained in order to function. However, three residues in the Zingiberales show signs of positive selection. In particular, residue 142 of the K3 domain of *ZinAGcp1* is under positive selection, and fixed differences in this residue among members of the ginger clade suggest that the chemical modification in this position are implicated in the morphological changes observed in the androecial whorl of the Zingiberales.

In *Antirrhinum*, a single amino acid change has been implicated in differences in the establishment of male and female identity between *AG* lineage genes *PLENA* (*PLE*) and *FARINELLI* (*FAR*). A single glutamine insert in the K3 domain of FAR leads to a limited protein-protein interaction between *AG* and *SEP* genes, underlying the functional differences observed between *FAR* and *PLE* in determining reproductive organ identity (Airoldi et al. 2010).

In transgenic *Arabidopsis* plants carrying *AG* anti-sense RNA a range of floral organ phenotypes is observed (Mizukami & Ma 1995). Flowers exhibiting intermediate levels of *AGAMOUS* espression show a large range of morphologies that includes the occurrence of petaloid stamens. Mutations in the regulatory site of *AG* in *Arabidopsis* can also lead to the development of petaloidy in the androecial whorl (citation?). Here, we show that androecial petaloidy in the Zingiberales is likewise associated with the *AG* lineage, and may result from a single amino acid change in the K domain of *ZinAGcp1*, after the divergence of the banana lineages and the ginger clade.

The role of subdomains of the K domain in MADS-box protein-protein interactions has been studied, especially in the formation of dimmers between B-class genes and *SEPALLATA* genes (Yang & Jack 2004). Also, the K domain has been implicated in the specification of heterodimerization of B-class proteins in *Arabidopsis* (Yang et al. 2003). Although the K3 domain has not been implicated in the dimerization of MADS-box proteins studied so far, this domain has played an important role in the formation of AP3/PI/SEP1 complexes (Yang et al. 2003), as well as to the promotion of DNA binding ability of SEP3 complexes (Melzer et al. 2009). It is also known that SEP3 proteins lacking the C-terminal end of the K3 domain are unable to form multimeric complexes with AG (Immik et al. 2009). This experimental evidence supportes the role of the K3 domain in the formation of multimeric complexes agrees with our hypothesis that a single amino acid alteration in this area of *AGAMOUS* might lead to a reduced binding/interacting affinity, leading to a post-transcriptional down regulation of *AG* function that is capable of explaining the morphological changes observed in the flowers of the Zingiberales order.

This view is supported by changes in the expression profile of *APETALA1* (*AP1*) across floral organs of the Zingiberales. In *Arabidopsis thaliana*, relatively high levels of *AP1* were detected in petaloid stamens and sepaloid carpels of flowers with reduced levels of *AG* due to anti-sense (RNAi) knockdown (Mizukami & Ma 1995). Accordingly, petaloid organs in the androecial whorl, as the ones observed in *Costus spicatus* are characterized by simultaneous expression of *AGAMOUS* and *AP1* indicating a lack of negative interaction between these two gene families. Morphologically, this expression profile corresponds to a 'hybrid' between stamen and petal attractors in *Arabidopsis thaliana* (Barrio et al. 2010). Although this 'hybrid' attractor has not been observed as a stable state of the *Arabidopsis thaliana* FOS-GRN, this observed change have been positively selected for in the Zingiberales. Perhaps, multiple duplication events within the FOS-GRN genes observed in the Zingiberales could lead to stable states that are not observed in *Arabidopsis* due to different lineage-specific duplication events and/or the loss of duplicated copies.

Based on these studies, positive selection acting upon *AGAMOUS* genes in the Zingiberales has resulted in a fixed change in the K3 domain responsible for the androecial petaloidy and infertility observed in the Zingiberales. If it is the case that the changes in the AGAMOUS genes are responsible for the observed changes in floral morphology across Zingiberales, a clear trade-off between production of fertile stamens and increased petaloidy has been fixed by positive selection in this order. Although androecial petaloidy is a remarkable feature of Zingiberales floral evolution, no changes have been observed in meristem determinancy. This might be explained, at least in part, by the potential functional redundancy between *AG* and *AGL11* lineage genes, as already reported for in rice and petunia (Dreni et al. 2011; Heijmans et al. 2012). Further studies of the AG subfamily genes in the Zingiberales will help understand the complete role of the AG subfamily in floral development and evolution across the Zingiberales.

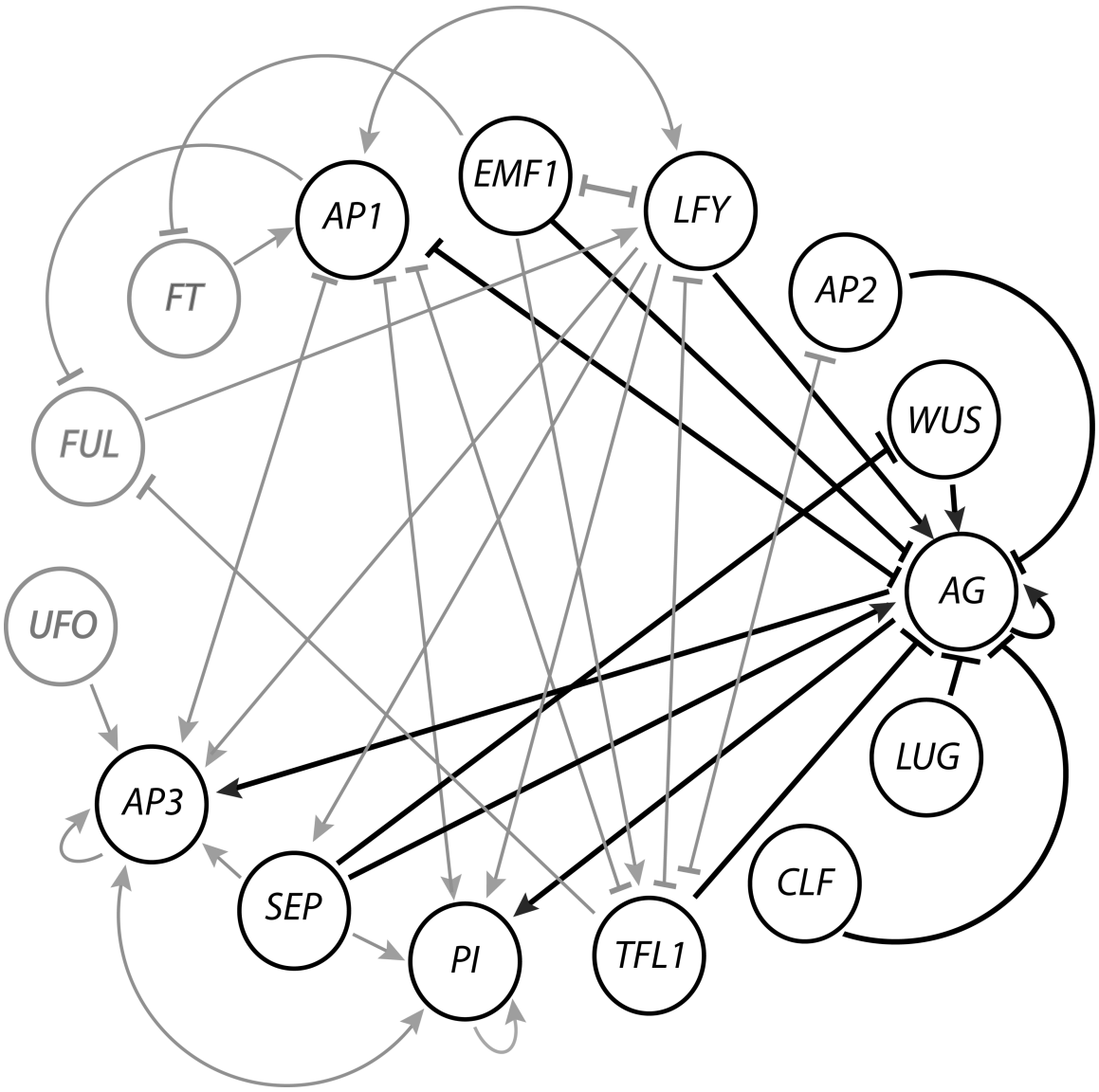

Figure 1. Floral Organ Specification Gene Regulatory Network (FOS-GRN), modified from Alvarez-Buylla et al. (2010a). The circles (nodes) represent genes or proteins that are experimentally known to participate in floral organ specification during flower morphogenesis in *Arabidopsis thaliana*. (→) describes positive interactions between nodes, while (—|) describes negative interactions between nodes. Fifteen genes are depicted: EMF1, LFY, AP2, WUS, AG, LUG, CLF, TFL1, PI, SEP, AP3, UFO, FUL, FT, AP1. Direct gene interactions between AG and other genes in the FOS-GRN are highlighted in black. Other interactions within the FOS-GRN not directly involving AG are depicted in grey, as well as genes that are not direct interactors of AG.

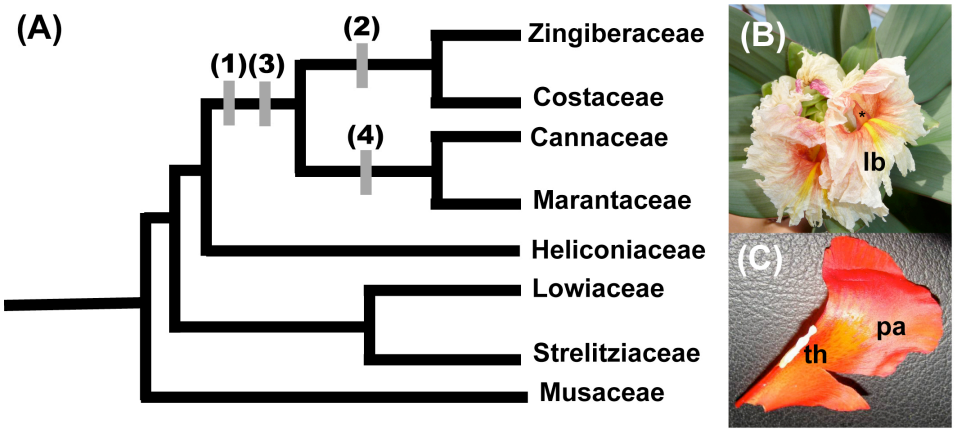

Figure 2. (A) Most important changes during morphological evolution of the androecial whorl are mapped onto the Zingiberales phylogeny (Kress et al. 2001). The eight Zingiberales families are divided into two groups: the first diverging banana lineages (Heliconiaceae, Strelitziaceae, Musaceae, and Lowiaceae), and the derived ginger clade (Zingiberaceae, Costaceae, Marantaceae, and Cannaceae). Main changes in androecial morphology are depicted with numbers. (1) Reduction in the number of fertile stamens, from 5-6 fertile stamens in the banana lineages, to 1 fertile stamen in Zingiberaceae and Costaceae or ½ of a fertile stamen in Marantaceae and Cannaceae; (2) Fusion of petaloid staminodes leading to the formation of the labellum. Five infertile stamens fuse in Costaceae, while variable number of staminodes form the labellum in the Zingiberaceae; (3) Laminar extension of the filament of the fertile stamen; (4) Abortion of ½ a theca of fertile stamen. (B) *Costus sp*. flowers. (lb) labellum; (*) abaxial side of laminar connective of fertile stamen. (C) *Canna indica* ½ fertile stamen. (th) single theca; (pa) petaloid appendage resulting from the laminar expansion of the filament (Almeida et al. 2013).

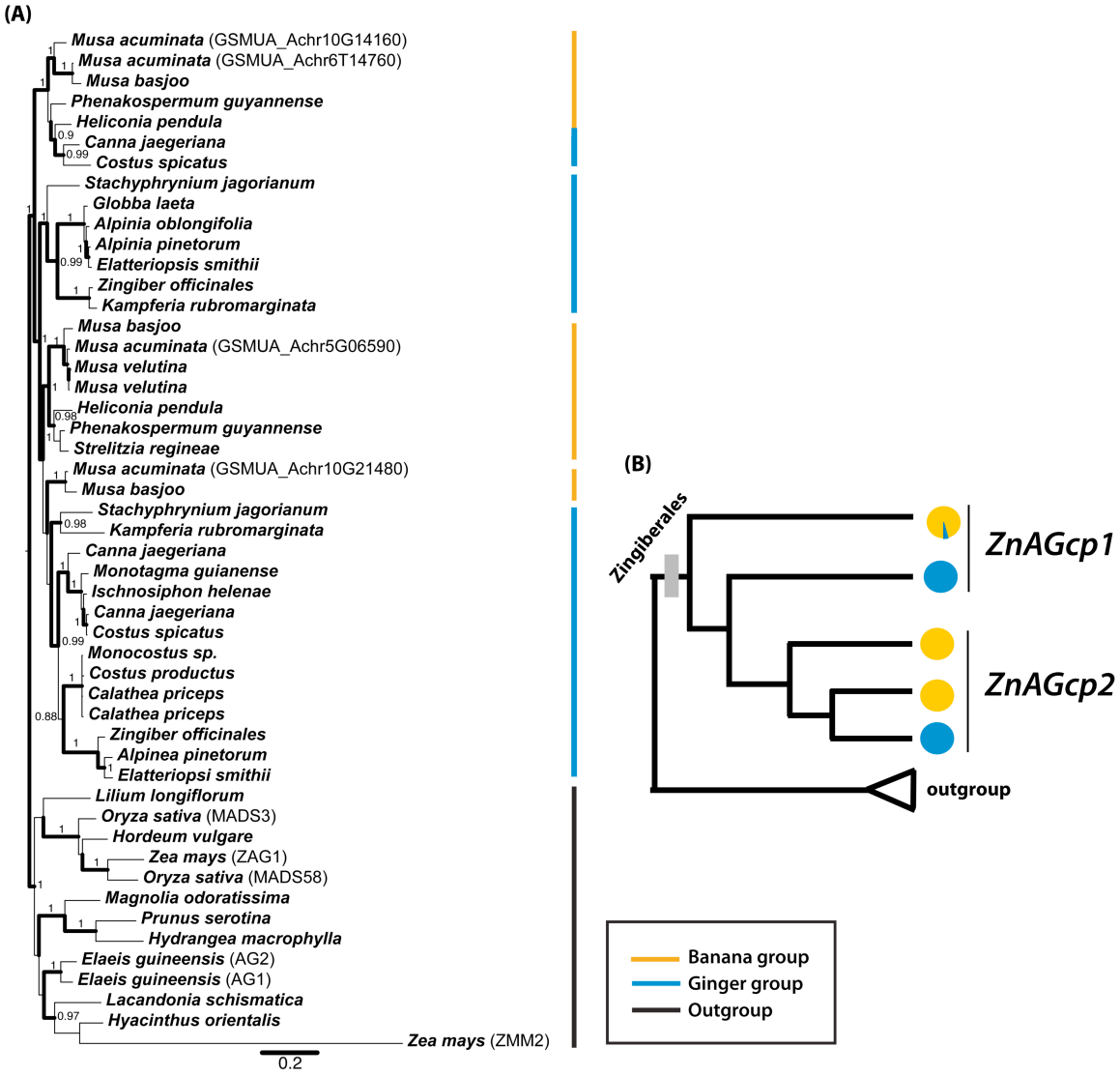

Figure 3. (A) Zingiberales *AGAMOUS* gene tree. Tree was generated using MrBayes and 3M generations, under the GTR+I+G model (best fit model according to MrModeltest). The general tree topology agrees with results generated by Maximum Likelihood (ML) analysis (PHYML), as well as under different models both in Mr. Bayes and in ML analysis. Partition of the data set according to codon position rendered an unresolved tree with poor likelihood (data not shown). Only posterior probabilities (PP) above 0.8 are presented. Branches that show >90% bootstrap support are depicted in bold. At least two copies of the gene *AGAMOUS* can be identified, according to the distribution of ginger and banana group species in the gene tree. (B) Schematic representation of the *AGAMOUS* gene tree. Colored circles represent clades, and each circle mirrors the position of the color-coded clades in the gene tree. This schematic tree depicts one nested clade (*ZinAGcp2*) comprised by ginger and banana group sequences, and two first branching paraphyletic lineages with banana and ginger group sequences (*ZinAGcp1*).

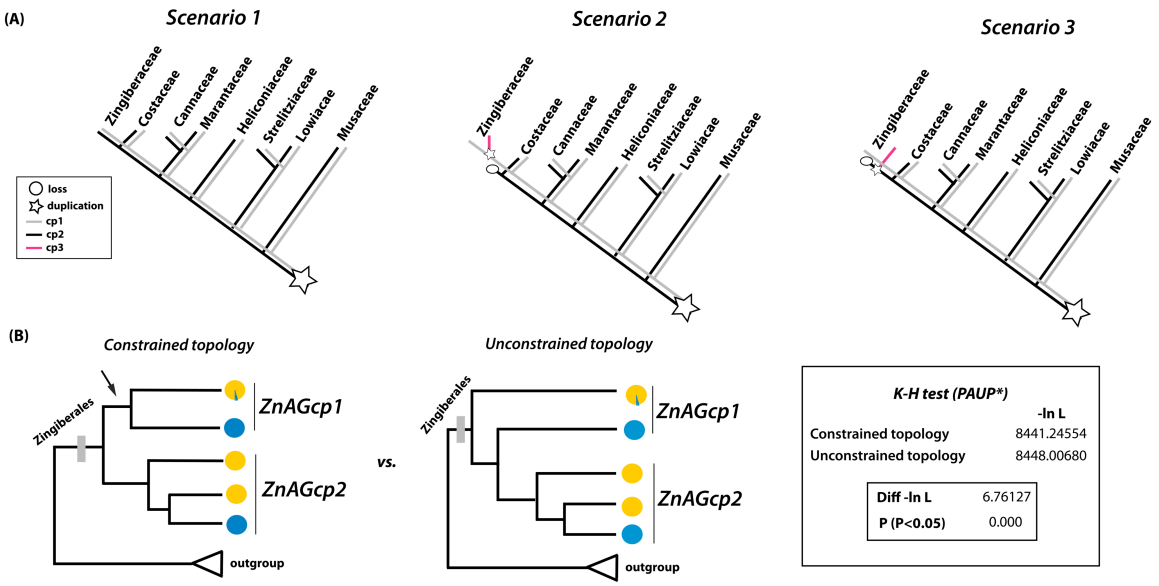

Figure 4. (A) Potential AGAMOUS gene-copy histories within the Zingiberales. Scienario1 assumes one single duplication event at the base of the Zingiberales order, leading to two distinct orthologous AGAMOUS lineages (cp1 and cp2). Scenarios 2 and 3 depict alternative histories of duplications and losses of the AG copies, particularly in the Zingiberaceae lineage. In both cases, orthologous relationships would be complicated by the existence of subsequent duplication events, unique to the Zingiberaceae lineage, leading to the evolution of yet another copy of AG, cp3. (B) Kishino-Hasewaga test (Kishino & Hasegawa (1989), K-H test) was performed using PAUP* on a constrained topology, where the two first paraphyletic lineages were forced to form a monophyletic clade. The likelihood score for the constrained topology was compared to the likelihood score of the unconstrained gene tree, as obtained on bayesian and maximum likelihood phylogenetic analysis (Figure 3). The constrained topology shows a better likelihood score than the one calculated for the gene tree topology presented here, supporting the idea that the first two paraphyletic lineages are actually derived from a single duplication event. This result supports the evolutionary history depicted by Scenario 1, in which *ZinAGcp1* and *ZinAGcp2* are orthologous lineages.

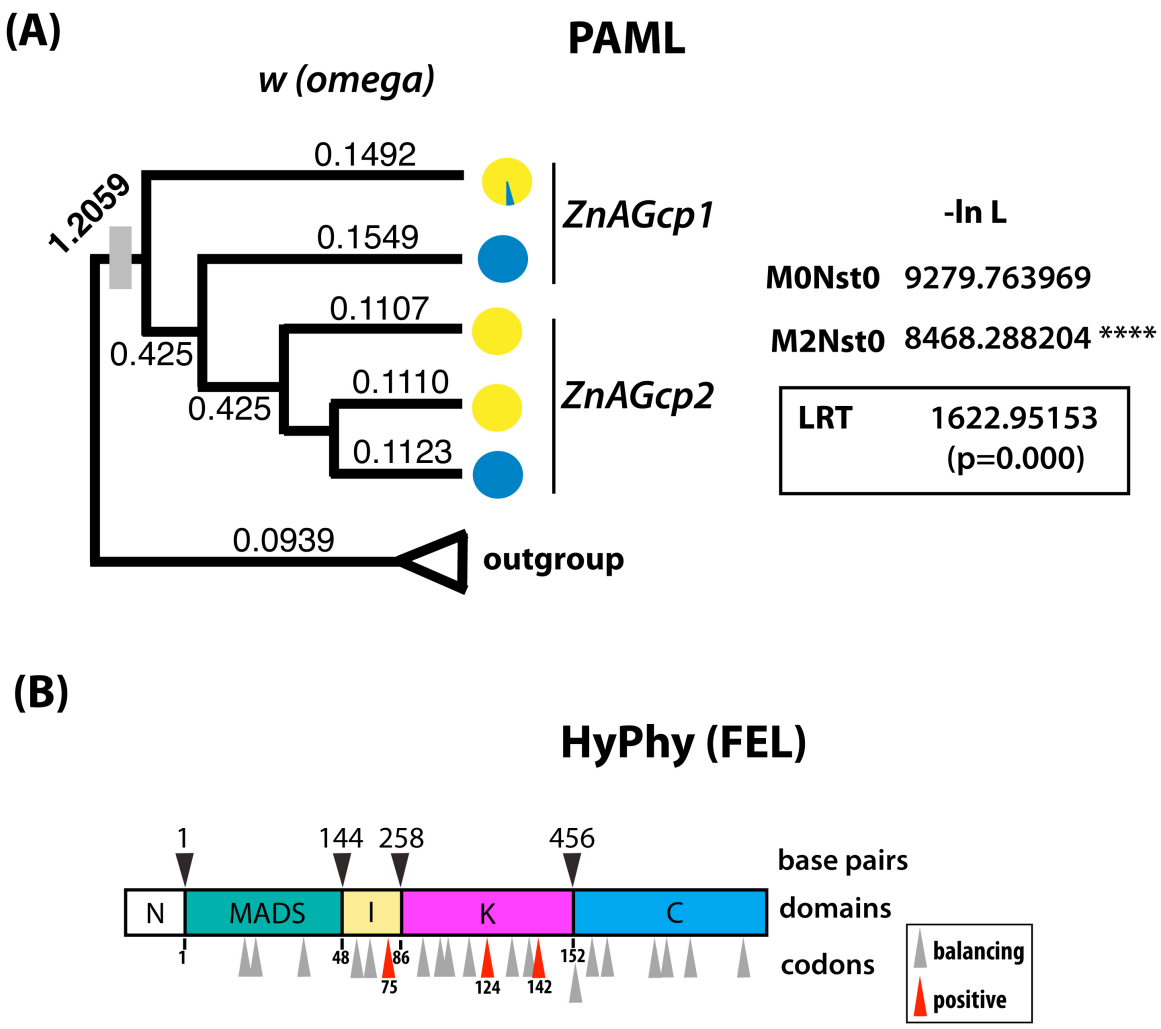

Figure 5. Selection tests. (A) PAML branch selection test. Omega (w) values are depicted for each branch in the gene tree. A likelihood ratio test (LTR) for branch models (M2Nst0 and M0Nst0) was performed. PAML detects a strong selection signal at the base of the Zingiberales sequences, but nowhere else in the gene tree. (B) HyPhy (package FEL) site-selection test. Balancing and positive selected residues are marked along the *AGAMOUS* protein domains. As expected FEL, detected various sites under balancing selection, while three sites were detected to be under positive selection, particularly in the I and K domains.

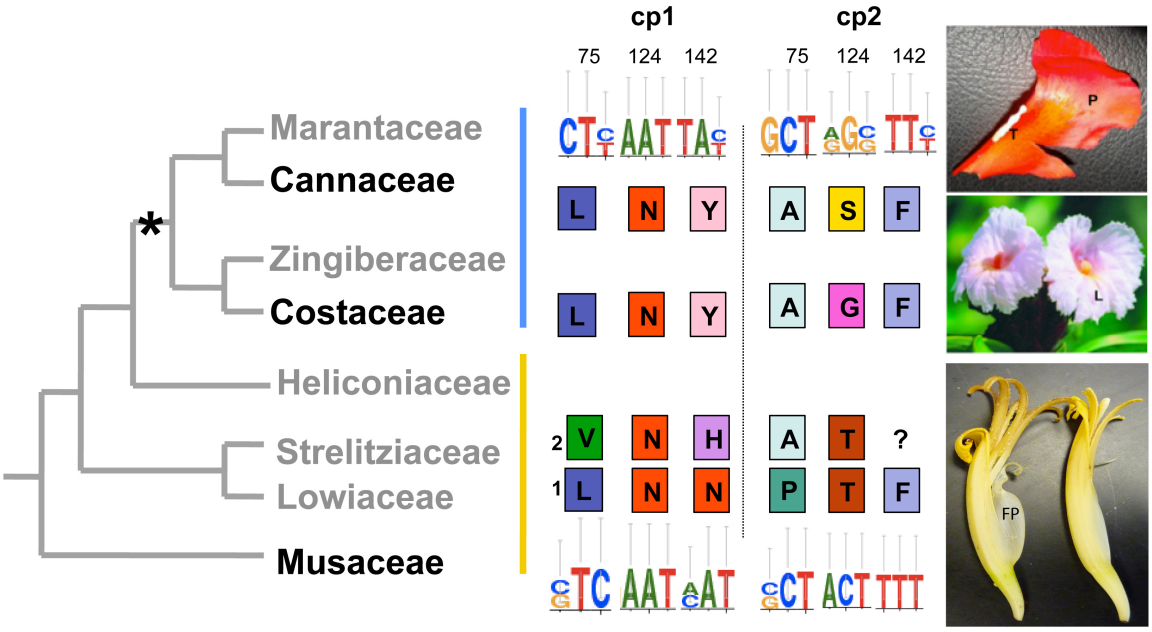

Figure 6. Aminoacid changes within positive selected sites in the two copies of the AGAMOUS gene across Zingiberales species. The (*) depicts the evolution of androecial petaloidy within the Zingiberales order. Note that it also corresponds to the base of the ginger clade (in blue). Marked in yellow are the paraphyletic lineages of the banana group. For aminoacid comparisons, *Musa acuminata* (Musaceae), *Costus spicatus* (Costaceae) and *Canna indica* (Cannaceae) AGAMOUS sequences were used. Logos for the specific codons of the banana group (bottom) and ginger clade (top) are shown. Single-letter aminoacid traditional names were used (colored boxes). On the far right, images of *Canna sp*. fertile stamen (top; T-theca; P-petaloid filament); *Costus sp*. labellum (L, middle image); and *Musa basjoo* flower (bottom; FP-free petal) are shown. Also, note that *Musa acuminata* has four AGAMOUS sequences due to a subsequent whole genome duplication event after the divergence of the Musaceae lineage (D'Hont et al. 2012).

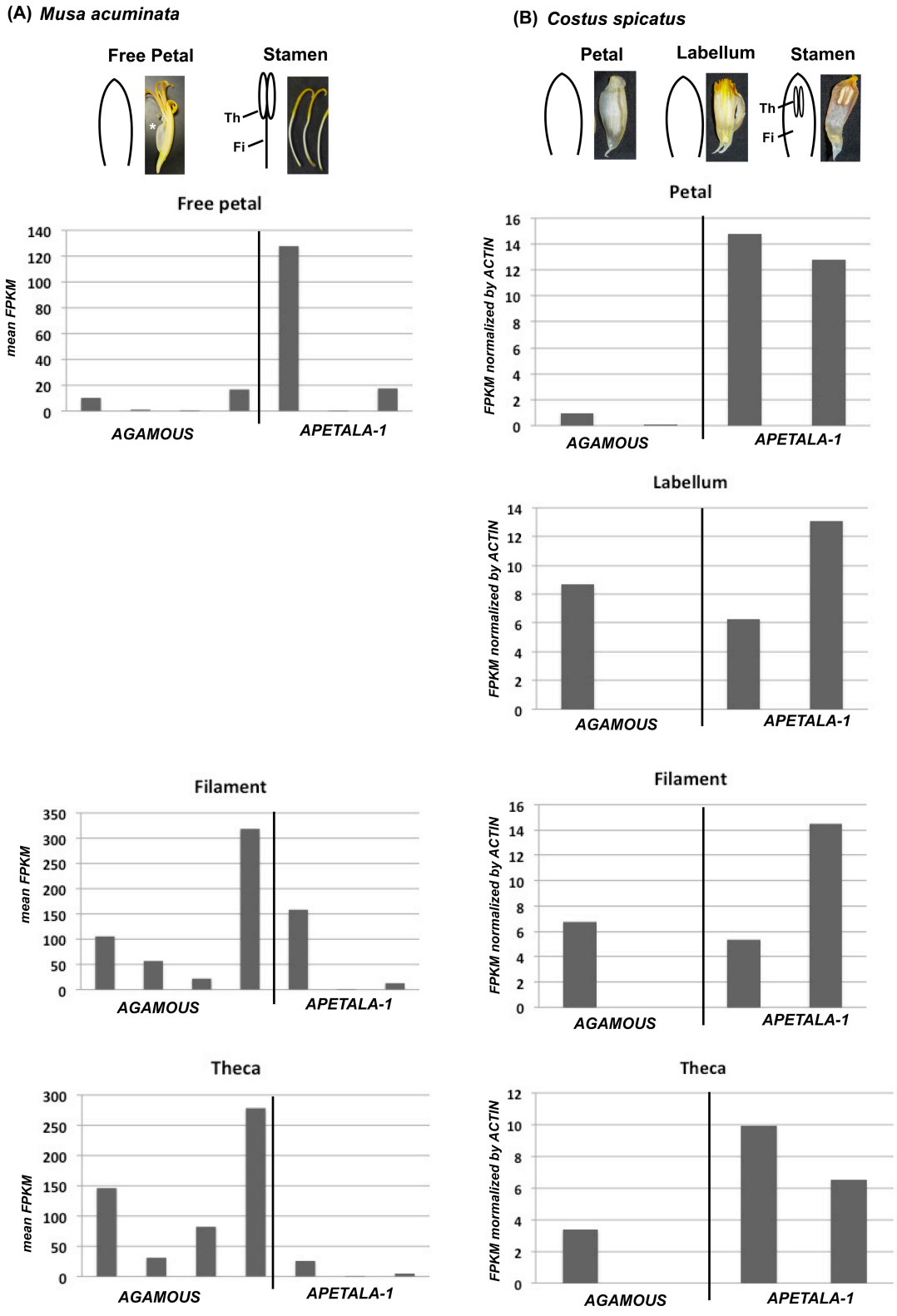

Figure 7. *Musa acuminata* and *Costus spicatus* gene expression based on transcriptomes of developing floral organs. (A) *Musa acuminata AGAMOUS* and *APETALA-1*

expression. *AGAMOUS* has four copies in *Musa*, due to an independent duplication event, are named distinguished by the letters 'a' and 'b'. Bars in each graph represent, from left to right, correspond to the expression of *ZinAGcp1a*, *ZinAGcp1b, ZinAGcp2a,* and *ZinAGcp2b* for AGAMOUS expression, and *AP1-1* and *AP1-2* for *APETALA-1*. Gene expression was estimated for *Musa acuminata* free petal (Fp), filament (Fi), and theca (Th). (B) *Costus spicatus* gene expression on petal (Pe), labellum (La), filament (Fi), and theca (Th). Bars in graph represent, from left to right, *ZinAGcp1*, *ZinAGcp2, AP1-1,* and *AP1-2*.

CHAPTER 4

Co-option of the polarity gene network shapes filament morphology in angiosperms

PREFACE

Despite the clear homeotic conversion of stamen primordia into petals-like organs (described in Chapter 1), as well as the potential involvement of the AGAMOUS gene in the evolution of androecial petaloidy in the Zingiberales (discussed in Chapter 3), it became increasingly clear that the floral organ identity genes, although important, are not the only ones involved in the evolution of stamen morphology in the order.

As described in Chapter 2, the laminar expansion of the filament, a derived characteristic of the ginger families (Cannaceae, Marantaceae, Costaceae and Zingiberaceae) potentially evolved from an almost radially symmetric filament, present in the banana lineages (Musaceae, Heliconiaceae, Strelitziacea, and Lowiaceae). In order to address the laminarization observed in the filament of the ginger families, I studies the abaxial-adaxial polarity gene regulatory network. In model plants like *Arabidopsis thaliana*, a balanced expression of abaxializing and adaxializing genes leads to the establishment of a middle zone where laminar expansion takes place. This phenomenon is observed in leaves, sepals and petals, as well as ovule integuments of Arabidopsis. However, whenever this abaxial-adaxial balance is broken, either via overexpression or underexpression of either abaxial or adaxial genes, the result is the radialization of an otherwise laminar structure.

Thus, I set to study the involvement of the polarity gene regulatory network in the evolution of filament morphology, not only in the Zingiberales, but also across angiosperms.

The chapter that follows has not been published elsewhere yet, and is summarized below.

The molecular mechanisms of abaxial-adaxial (ab-ad) polarity in plants have been well studied as a property of lateral and flattened organs, such as leaves. Results of studies focusing on the genetic mechanisms of leaf development suggest that laminar expansion takes place at the domain of interaction between abaxializing and adaxializing agents, in the presence of balanced ab-ad gene expression. Over- or under- expression of either abaxializing or adaxializing genes inhibits laminar growth, resulting in a mutant radialized phenotype. In the Zingiberales, an order of tropical monocots, the evolution of androecial morphology is characterized by a marked reduction in the number of fertile stamens and their replacement by infertile stamens (staminodes) that are flattened, laminar ("petaloid") structures.

We hypothesize that the evolution of petaloidy in these staminodes is tightly linked to the co-option of the leaf ab-ad gene network acting during development of the filament of the petaloid staminodes. Transcriptome data from *Costus spicatus* and *Musa acuminata* filaments show a 10-fold overexpression of *Arabidopsis thaliana* YABBY2/5 ortholog in the *Musa acuminata* radial filament when compared to the *Costus spicatus* laminar filament. To further test our hypothesis, we compared our results with gene expression data from the filament of *Brassica rapa* - an eudicot rosid from the order Brassicales. Ab-ad polarity gene expression in the radial *Brassica rapa* filament is consistent with that found in the radial filament of *Musa acuminata*, providing further evidence for the involvement of the ab-ad polarity network in filament morphology across angiosperms.

Disruption of a balanced ab-ad gene expression in the filament inhibits laminar growth, resulting in a radial structure. Here, we argue that the co-option of the leaf ab-ad polarity network is an important feature of filament morphology across angiosperms; the angiosperm filament is an abaxialized structure, equivalent to the radialized leaf of an ab-ad polarity mutant. Finally, when considered together with data from the literature, our results suggest that the polarity gene network is an ancestral regulatory module involved in shaping basic angiosperm form.

INTRODUCTION

The homology between floral organs and leaves has long been proposed as a fundamental principle of plant comparative morphology. Since 1790, Goethe envisioned what in more recent years has been described as 'serial homology' between vegetative and reproductive organs. For Goethe (1790), leaves and floral organs shared an underlying common theme that would manifest itself differently during specific phases of plant development, leading to a series of successive forms: leaves, bracts, sepals, petals, stamens (microsporophylls), and carpels (megasporophylls). More recently, the genetic mechanisms underlying plant organogenesis, as well as leaf and floral organ initiation, growth and development have been revealed (Goto et al. 2001, Pelaz et al. 2001, Litt & Kramer 2010). Indeed, many of the same mechanisms that operate in shaping leaves seem to play an important role in shaping floral organs, with perhaps the exception of the mega- and microsporangia (Lönnig 1994). For instance, polarity genes, such as those described in leaf abaxial-adaxial patterning, have also been implicated in establishing abaxial-adaxial polarity of various floral organs (e.g. polarity of sepals and petals, McConnell and Barton, 1998; Sawa et al., 1999; Pekker et al., 2005; polarity in stamens and ovules, Toriba et al. 2010; polarity during ovule development, Baker et al. 1997; Villanueva et al. 1999; Léon-Kloosterziel et al. 1994; Sieber 2004; McAbee et al. 2006).

Abaxial-adaxial (ab-ad) polarity in plants has been most widely studied as a property of flattened lateral organs (i.e. the leaf). The first ab-ad polarity genes were described in *Antirrhinum majus* and *Arabidopsis thaliana* mutants with cylindrical (= radial), or filamentous, leaves in place of the WT laminar structures (Talbert *et al*. 1995; Waites & Hudson 1995; McConnell & Barton 1998). Since then, much has been discovered about ab-ad polarity patterning and the underlying gene regulatory network, including the fundamental role of smallRNAs, such as tasiR-ARF and miR165/166 (see Kidner & Timmermans 2007, and review in: Alvarez-Buylla et al. 2010).

Most importantly, it has been shown that laminar development, of the type observed in leaves and most floral organs, is dependent on the proper specification of abaxial and adaxial identity. Disruption of either the abaxial or adaxial signaling pathway tends to produce a mutant phenotype of radialized leaves, with the epidermis displaying markers characteristic of the adaxial surface (in the case of abaxial signaling mutants) or the abaxial surface (in the case of adaxial signaling mutants).

The molecular processes implicated in laminar expansion in plants – the 'abaxial–adaxial juxtaposition hypothesis' (Waites & Hudson 1995) - is described as the contact zone of abaxializing (e.g., YABBY genes, KANADI genes) and adaxializing (e.g., REVOLUTA, PHABULOSA) gene products, where laminar outgrowth is initiated at this contact zone. It is interesting to note that down-regulation or up-regulation of either abaxial or adaxial genes leads to an impairment of laminar growth, suggesting that a proper balance between abaxializing and adaxializing gene products is essential for laminar expansion. Despite the relatively detailed knowledge of the gene regulatory networks involved in determining the polarity of leaves and floral organs in model systems, we still know relatively little about how such networks have changed though

evolutionary time or how they may be responsible for different leaf and floral morphologies observed during angiosperm evolution.

Here we present evidence for the involvement of polarity genes in stamen morphological diversity through the study of the evolution of androecial form across the Zingiberales. The evolution of floral morphology in this order is tightly correlated to increased petaloidy of the androecial whorl, which is, in turn, inversely correlated to the number of fertile stamens (Figure 1). The evolution of petaloid, laminar filaments and the development of both sterile and fertile petaloid stamens within the Zingiberales provide a unique opportunity for the study of the role of polarity genes in shaping filament morphology.

We hypothesize that laminar expansion in the filaments of most of the derived species in the order is a result of altered expression of polarity genes with compared to species with ancestral, radial filaments. According to this hypothesis, the gene expression pattern observed in petaloid stamens restores the balance between abaxial and adaxial gene products within the stamen primordium, leading to the elaboration of a margin that functions to signal laminar expansion in an otherwise radial filament. In order to test this hypothesis, we generated a set of organ-specific transcriptomes for floral organs of *Musa basjoo* and *Costus spicatus*, including radial (*Musa*) and petaloid (*Costus*) stamens. The results were compared to gene expression in the radial filament of *Brassica rapa* (Edger et al. unpublished).

Our results constitute a unique report on the role of polarity genes in stamen morphological evolution, and provide further support for the hypothesis that that the same underlying molecular mechanisms functioning in leaves have been co-opted during the evolution of the stamen whorl in Zingiberales, resulting in petaloid stamens and staminodes in derived members of the order. These polarity genes are likely to be responsible for shaping stamen morphology across angiosperms. Thus, our results suggest that the ab-ad polarity gene network is one of the fundamental modules involved in shaping plant morphology during angiosperm evolution, and particularly floral organ shape.

MATERIAL AND METHODS

RNA extraction and Illumina Library Preparation and Processing

Musa basjoo (UCBG) and *Costus spicatus* (Specht Lab Greenhouse, UCB) flowers were collected and floral organs were carefully dissected. Immediately following dissection, floral organs were flash frozen in liquid nitrogen and stored at -80°C. Organ-specific total RNA extraction was performed using INVITROGEN Plant RNA Reagent, according to the manufacture's guidelines. Illumina libraries were prepared using the TruSeq RNA sample prep kit v2. Two libraries each were prepared from *Musa basjoo* filament and free petal RNA, and *Costus spicatus* filament RNA and multiplexed using barcoding set A. Samples were run in a HiSeq2000 at IIGB HT Sequencing Facility at the University of California, Riverside.

Raw reads were trimmed to remove adapters and regions of poor quality with *cutadapt* (Martin 2011). *Costus spicatus* sequences were assembled into a reference transcriptome using Trinity (Grabherr et al. 2011).

Two independent replicates of RNA-seq data for filaments in *Brassica rapa* B3 were generated from greenhouse-grown plants, from which the central half of the filament was stored in RNAlater until RNA extraction was performed. Raw reads were downloaded from NIH, and FPKMs were generated using protocol detailed below for *Musa* short reads.

Polarity Gene Orthologs in *Arabidopsis thaliana*, *Musa acuminata*, and *Costus spicatus*

The polarity gene network for *Arabidopsis thaliana* is integrated elsewhere (Alvarez-Buylla et al. 2010). Sequences for each of the described *Arabidopsis thaliana* polarity genes were downloaded from NCBI: KANADI genes ATS, (NM_001125891.1) KAN2 (NM_102957.3), KAN3 (NM_117878.2), KAN4 (NM_121662.2); one AGONAUTE gene (AGO1, NM_001198240.1); YABBY genes INO (NM_102191.5), YAB5, (NM_179749.2) , YAB2 (NM_001084021.1), AFO (NM_130082.3), CRC (NM_105585.2), YAB3(NM_116235.2); two ASYMETRIC LEAVES, AS1(NM_129319.3), and AS2(NM_105235.4); NUB (NM_101210.2), JAG (NM_105519.3); PHB (NM_129025.3), REV (NM_125462.3), and PHV (NM_102785.4).

Coding sequences (predicted CDS) for *Musa acuminata* genes were downloaded from the CIRAD website (*Musa acuminata* genome CIRAD Website - http://banana-genome.cirad.fr/download.php; D'Hont et al. 2012) and reciprocal BLAST was used to search for *Musa acuminata* orthologs of the *Arabidopsis thaliana* polarity genes listed above. *Costus spicatus* orthologs were identified using reciprocal BLAST against the assembled transcriptome using the *Musa acuminata* gene list.

Polarity Gene Expression

GSNAP (Wu & Watanabe 2005; Wu & Nacu 2010) was used to align *Musa basjoo* trimmed reads to annotated CDS from the published *Musa acuminata* genome, and *Costus spicatus* trimmed reads to the *Costus spicatus* reference transcriptome. The expression of polarity genes was estimated using eXpress in units of FPKM (frequency per kilobase of exon per million aligned reads) (Robert et al. 2011). Replicates were independently processed, and gene expression was compared between libraries for consistency.

In order to confirm transcriptome gene expression of genes of interest, Real Time PCR (RT-PCR) was carried out. Total RNA was extracted from fresh tissues of labellum (fused petaloid staminodes) and the petaloid filament of *Costus spicatus*, as well as from the single free petal and the cylindrical filament of *Musa acuminata* (UC Davis) using

INVITROGEN Plant Reagent. Total RNA was treated with Ambion TURBO DNA according to the rigorous treatment for abundant DNA contamination, following the protocol provided by the manufacturer. cDNA was generated from 1μg of total RNA using SuperScript®, according to the provided protocol. Reactions without reverse transcriptase (no-RT) were performed as a control for genomic DNA contamination. Actin was used as a positive control for cDNA synthesis.

RT-PCR primers were designed using Primer3 (http://frodo.wi.mit.edu/) based on *Musa basjoo* and *Costus spicatus* transcripts. *ACTIN1* was used as an internal control for the RT-PCR reaction. Expression was confirmed for three genes for both *Musa acuminata* (*Ma*) and *Costus spicatus* (*Cs*): *YABBY2/5* orthologous genes; *AS1* orthologs, and a homolog of *KAN2/3*. *Musa acuminata* primers used were the following: *MaACTIN1* forward TCCATCATGAAGTGCGATGT, *MaACTIN1* reverse CTCTGCTTTTGCAATCCACA; *MaYABBY2/5* forward AGCATTGTAGCAGTGCGATG, *MaYABBY2/5* reverse GGACGCATAGGCAGCATAAT; *MaAS1* forward AGCGGTGGAGATCTGAAGAG, *MaAS1* reverse CGCGATCTTCTTCCACTTGT; *MaKAN2/3* forward CTTCCCAACACAGCCAGATT, *MaKAN2/3* reverse CGAAATTGGAGGTGGAAGAA. *Costus spicatus* primers, for orthologous genes, were the following: *CsACTIN1* forward GCATGAGCAAGGAGATCACA, *CsACTIN1* reverse CAAACATGACTTGGGTGTGC; *CsYABBY2/5* forward CGGTTAGTGTGCCAGGAAAT, *CsYABBY2/5* reverse CATGGGGAGCATCTGTTCTT; *CsAS1* forward GATCCGAAGAGGATGCGATA, *CsAS1* reverse ACTTGTTTCCGTGCTTTGCT; *CsKAN2/3* forward TTAGGAGGCCATGAGAGAGC, *CsKAN2/3* reverse GGATTAACCTCGGGCAAGTT. Primers were tested for single band amplification and primer dimer prior to the RT reaction.

RT-PCR was performed with iQ SYBR Green Supermix kit (BIO RAD) for a 20μl reaction (0,75μl of 10mM primers forward and reverse; 10μl 2X iQ™ SYBR® Green Supermix; 5μl of a 1:5 cDNA solution; and 3,5μl H_2O). The cycling protocol was as follows: 1X initial denaturation time at 95°C for 2 minutes, followed by denaturing and annealing-extension cycle 95°C for 10 seconds, and 55°C for 30 seconds, 44X; and a melting curve from 65°C-95°C in 0.5°C increments, for 10 seconds. A standard curve (1:1; 1:10; 1:100; 1:1,000; 1:10,000) was carried out for each primer pair.

RESULTS

Polarity Gene Orthologs in *Arabidopsis thaliana*, *Musa acuminata* and *Costus spicatus*

Based on the results of a reciprocal BLAST search, probable orthologs of *Arabidopsis thaliana* ab-ad polarity genes were identified from *Musa acuminata* and *Costus spicatus* (Table 1). In many cases, there was a one-to-many, or a many-to-one relationship between *Arabidopsis* genes and either *Musa* or *Costus* genes. Only genes that were present in *Costus spicatus* and *Musa basjoo* transcriptomes were used for downstream

expression analyses. With only few exceptions, BLAST results identified a one-to-one relationship between *Musa acuminata* and *Costus spicatus* genes, which facilitated downstream expression analyses.

Starting with twelve characterized *Arabidopsis thaliana* ab-ad polarity genes, twenty-one presumed orthologs were identified within the *Musa acuminata* genes, and fifteen orthologous genes were retrieved from the *Costus spicatus* transcriptome assembly (Table 1), in line with the lineage-specific whole genome duplication proposed for *Musaceae* by the *Musa acuminata* genome sequencing group (d'Hont et al. 2011).

Abadixal-adaxial Polarity Gene Expression

Abaxial-adaxial polarity genes' FPKMs were calculated using eXpress. Transcriptome replicates for all sampled floral organs resulted in similar FPKM counts for polarity genes in both *Costus spicatus* and *Musa basjoo* replicates (Supplemental material). A mean FPKM was thus calculated across both replicates (Table 01).

Overall, ab-ad polarity genes in the filament of *Musa basjoo* and *Costus spicatus* showed low expression values in the RNA-seq datasets (Figure 2a). The exception was *Musa GSMUA_Achr7G01330_001*, a *YABBY2/5* ortholog. In *Musa basjoo* (Figure 2b) filaments, this gene showed a 25-fold increase in FPKM when compared to *Costus spicatus* (Figure 2c) filament expression (from 15.7 in *Costus* to 395.8 in *Musa*). The orthologous gene from *Costus spicatus* transcriptome, *YABBY2/5* orthologue *comp35601_c0_seq1*, showed a low expression in the filament, equivalent to those found for other polarity genes in the filament. Interestingly, the expression of *Musa GSMUA_Achr7G01330_001* is 4-fold higher in the filaments of *Musa basjoo* than in the free petal – a laminar structure. When filament expression of other *YABBY2/5* orthologs was compared between *Musa basjoo* and *Costus spicatus*, no correlation was found between differential expression and differential morphology (Figure 2b). For all other *YABBY2/5* homologues, *Musa* and *Costus* filament showed very similar FPKM.

In *Brassica rapa,* a distantly related species with a cylindrical filament morphology, similar to that of banana (Figure 2d), the expression of *YABBY2/5* genes show a similar pattern. *Brassica rapa* has three *YABBY2* and one *YABBY5* homologues. In the *Brassica rapa* filament transcriptome, most genes, including *YABBY* genes, are expressed in very low levels in the filament. However, *YABBY2-2A* shows a 9-fold increase in FPKM when compared to other *YABBY2* or *YABBY5* gene expression in the same organ. (Edger et al., unpublished data). Also, YABBY2-2A has, at least, a two-fold increase in expression when compared to leaves or sepals. Although petals show high expression levels of YABBY2-2A, the levels of expression of this gene are still significantly lower than those from filament samples.

In order to confirm transcriptome results for genes of interest, RT-PCR was performed on *Musa acuminata* filament and free petal, and *Costus spicatus* labellum (fused petaloid staminodes) and petaloid filament. *YABBY2/5* ortholog (GSMUA_Achr7G01330_001, comp35601_c0_seq1); *KANADI2/3* ortholog (GSMUA_Achr10G06430_001,

comp30224_c0_seq1); and *AS1* (GSMUA_Achr9G03320_001, comp38033_c0_seq1) had their expression confirmed. Gene expression shows the same pattern: low expression overall for all genes in the filament when compared to YABBY2/5 expression levels in *Musa acuminata* (Figure 3). It is interesting to note that although transcriptome results pointed to a potential role of AS1 ortholog in the downregulation of YABBY2/5 in *Costus* filaments, quantitative PCR results suggests a potential role of AS1 in the labellum, but not in the filament. It might be the case that different regulatory mechanisms are operating in an organ-specific manner to down-regulate YABBY2/5 expression. A deeper understanding of YABBY2/5 downregulation in *Costus spicatus* awaits further studies.

DISCUSSION

The abaxial-adaxial polarity gene network has been independently implicated as a potential developmental mechanism for various morphological traits during land plant evolution, such as leaf lamina and ovule integument development (Bowman et al., 2002; Eshed et al., 2004; Kidner and Timmermans, 2007; Sarojam et al., 2010; Villanueva et al., 1999; McAbee et al., 2005). Most of the genes comprising this network are shared by at least all vascular plants (i.e. YABBY genes, Yamada et al. 2011; Nishiyama et al. 2003) or even by all land plants (Streptophytes) (i.e., HD-ZIPIII genes, Floyd et al. 2006), suggesting that this network, or at least some of its components are shared by all vascular plants.

In particular, the YABBY gene family has five distinct members in angiosperms: FILAMENTOUS FLOWER (FIL), YABBY2, CRABS CLAW (CRC), INNER NO OUTER (INO), and YABBY5 (Yamada et al. 2011). Functional studies in *Arabidopsis thaliana* have supported the idea that some of the members of the YABBY family (more specifically FIL, YABBY2 and YABBY5) play important roles in both abaxial cell fate, and abaxial/adaxial juxtaposition-mediated lamina expansion (Eshed et al. 2004) in angiosperm lateral organs, while others (particularly CRC and INO) are expressed specifically in the carpel, and are important in ovule integument development (Bowman & Smyth 1999; Baker et al. 1997).

Recent studies on YABBY diversification in angiosperms suggest that abaxial expression of YABBY5 and FIL in lateral organs is an ancestral pattern, as is carpel expression of INO and CRC. However, restriction of CRC expression to the carpel was acquired later after the divergence of the Nymphaeales (Yamada et al. 2011). Thus, throughout the history of the vascular plants the leaf ab-ad polarity network has been co-opted for the specification of floral organ polarity, such as the case for ovule integument polarity. If we define co-option as the use of pre-existing traits – including genes and/or entire gene regulatory networks or modules – in a new way, and presume that co-option can ultimately generate novelties (True & Carroll 2002), we can argue that the involvement of the leaf ab-ad polarity gene regulatory network underlying androecial petaloidy in the Zingiberales is a clear case of co-option. Moreover, extrapolating from the data obtained from *Brassica rapa* filaments, ab-ad polarity genes seem to play a role in stamen morphology throughout the angiosperms and appears thus to be a fundamental network

that regulates laminar growth across the entire plant body plan.

It is striking to notice that flattened petaloid filaments and radial filaments have distinct macro and microscopical morphology. While radial stamens, like those observed in *Musa* and *Brassica*, have a single vascular trace, petaloid stamens exhibit complex vasculature. These morphological differences parallel those described for the controversial distinction between megaphyllous and microphyllous leaves (Tomescu 2008). It has been shown that at least some of the molecular mechanisms specifying the development of microphylls and megaphylls are shared between microphyll-bearing lycopodes and megaphyll-bearing euphyllophytes (ferns, gymnosperms and angiosperms), demonstrating the importance of consideration of physical constraints when comparing structures of differing size but similar position and/or function (Harrison et al. 2005).

Studies on the molecular mechanisms of cladode development also demonstrate the role of YABBY genes and the ab-ad polarity network in the alteration of radial stem morphology to form of a laminar, leaf-like organ in Asparagales (Nakayama et al. 2010; 2012). Whether of determinate (e.g. leaf and flower) or indeterminate (e.g. stem) growth, lateral organ morphology in vascular plants seems to be intrinsically correlated with position relative to the shoot apex and the proper function of the ab-ad polarity network promoting (when expressed correctly) laminar expansion. In this broader sense, whether a lateral structure will develop into a radial or flattened organ is dependent on balanced expression of abaxializing and adaxializing gene products in the lateral organ. Unbalanced gene expression will result in the default state of radial growth.

Thus, a single molecular mechanism has been co-opted throughout the evolution of the angisoperms to shape lateral organ morphology regardless of the identity of the organ itself. Here, in line with previous studies of lateral organ development in angiosperms, we argue that the ab-ad polarity gene regulatory network is an ancestral developmental regulatory module guiding the development of flattened versus radial structures in angiosperm lateral organs, including floral organs.

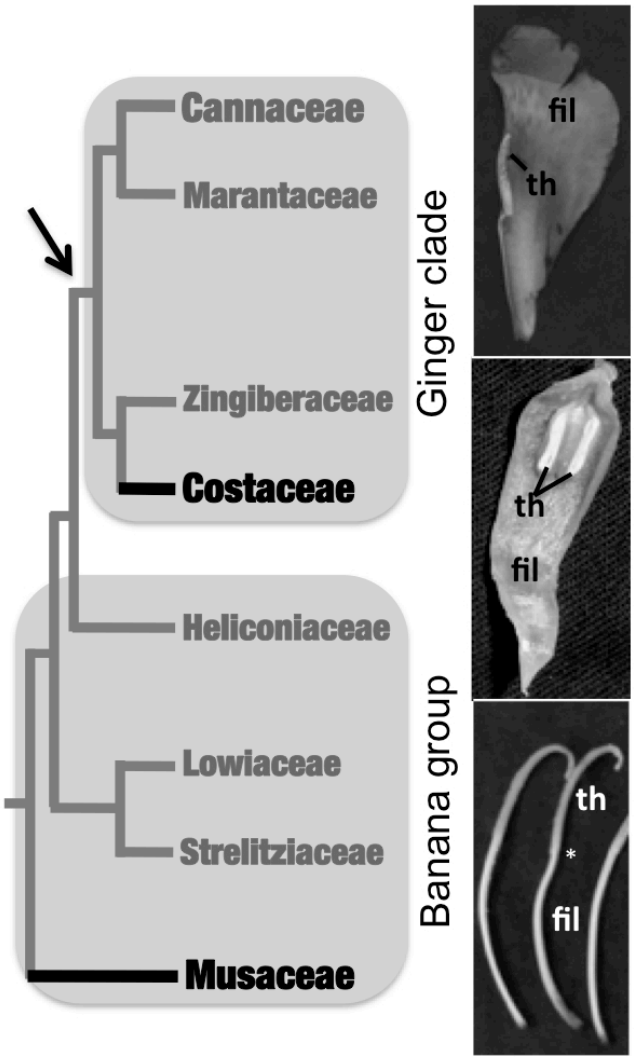

Figure 1 – Phylogenetic relationships of, and androecial morphological evolution within the Zingiberales order. This order of herbaceous tropical monocots of approximately 2,500 species comprises eight families which are generally divided into two groups: the basal paraphyletic banana group - including Musaceae, Strelitziaceae, Lowiaceae, and Heliconiaceae -, and the ginger clade comprising Zingiberaceae, Costaceae, Marantaceae, and Cannaceae. The arrow points to an important transition in the evolution of androecial morphology in the Zingiberales. At the base of the ginger clade, there is an impressive reduction in the number of fertile stamens, from 5 or 6 in the banana group to 1 or ½ in the ginger clade. This reduction in the number of fertile stamens is also associated with increased petaloidy of the infertile androecial members (i.e., staminodes). Furthermore, fertile stamen members in the ginger clade families, such as Costaceae and Cannaceae also show laminar expansion of the filament (Almeida et al 2013), leading to a petaloid fertile stamen. Photos depict the fertile filament of *Canna indica* (Cannaceae) at the top; *Costus spicatus* (Costaceae) in the middle; and *Musa basjoo* (Musaceae) at the bottom. th – thecae; fil – filament. The * marks the transition between the theca and the filament within the banana stamen.

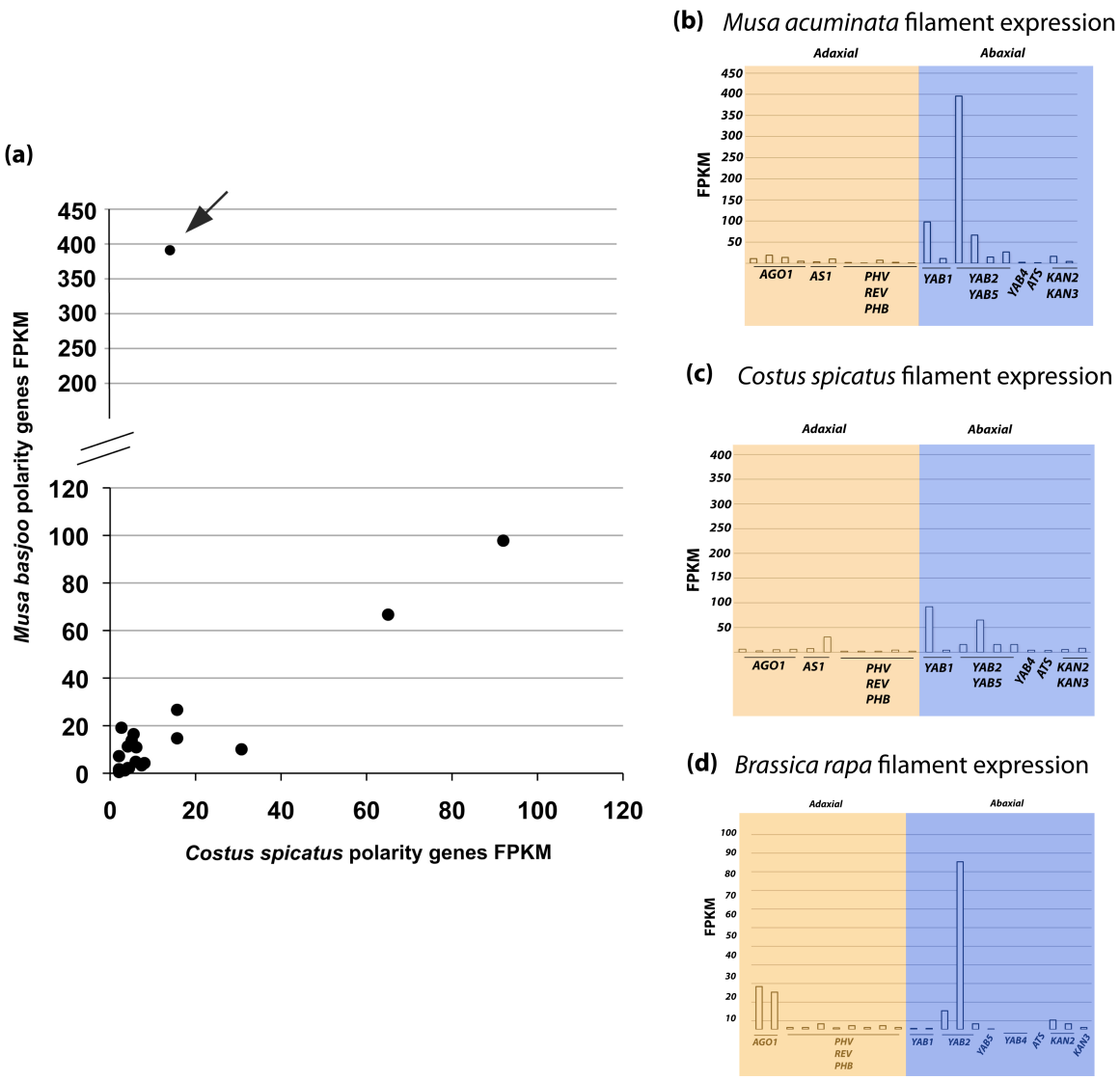

Figure 2 – (a) Fragments Per Kilobase of exon per Million fragments mapped (FPKM) for *Musa basjoo* and *Costus spicatus* ab-ad polarity genes. *Musa acuminata* genome was used as reference nomenclature for *Musa basjoo* gene expression. FPKM of all ab-ad polarity genes analyzed for (b) *Musa basjoo*, (c) *Costus spicatus*, and (d) *Brassica rapa* filaments.

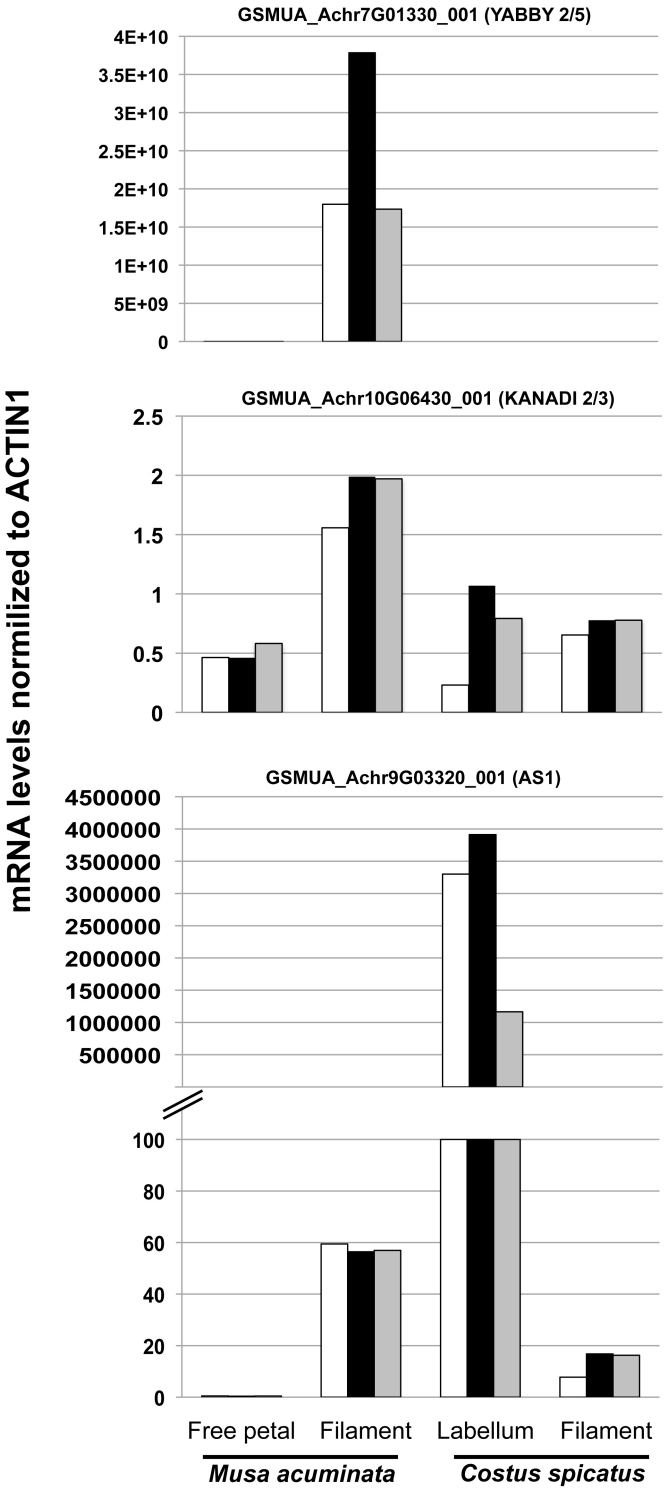

Figure 3 – Quantitative PCR for *Musa acuminata* and *Costus spicatus* YABBY2/5, KANADI2/3, and AS1 homologues. mRNA levels are normalized to ACTIN1 (see Supplemental material). Results are based on three replicates per sample (depicted by different bar colors).

Table 1 – Orthologous polarity genes in *Arabidopsis thaliana*, *Musa acuminata*, and *Costus spicatus*. Mean FPKM values for each gene, based on two filament transcriptome technical replicates, for *Costus spicatus* and *Musa basjoo* are also presented.

Polarity genes			**Mean FPKM**	
Arabidopsis	***Musa***	***Costus***	***Musa***	***Costus***
AGO1 (NM_1198240.1)	GSMUA_Achr3G27070_001	comp25496_c0_seq1	10.9	6.125
	GSMUA_Achr3G13970_001	comp25306_c0_seq1	19.13	2.645
	GSMUA_Achr1G17950_001	comp33286_c0_seq1	13.82	5.045
	GSMUA_Achr1G07120_001	comp33722_c0_seq1	4.81	5.98
AS1 (NM_129319.3)	GSMUA_Achr5G29070_001	comp7486_c0_seq1	3.32	7.365
	GSMUA_Achr9G03320_001	comp38033_c0_seq1	10.06	30.745
PHB (NM_129025.3)	GSMUA_Achr4G28360_001	comp13201_c0_seq1	1.675	2.075
	GSMUA_Achr7G01060_001	comp13201_c0_seq1	0.51	2.075
REV (NM_125462.3)	GSMUA_Achr4G31660_001	comp13201_c0_seq1	7.195	2.075
PHV (NM_102785.4)	GSMUA_Achr5G02420_001	comp28582_c0_seq1	2.13	4.41
	GSMUA_Achr2G13050_001	comp13201_c0_seq1	0.785	2.075
YABBY1 (AFO) (NM_130082.3)	GSMUA_Achr7G07830_001	comp33903_c0_seq1	97.8	91.995
	GSMUA_Achr8G11580_001	comp29400_c0_seq1	11.26	4.14
YABBY2 (NM_001084021.1)	GSMUA_Achr7G01330_001	comp35601_c0_seq1	**395.8**	15.675
	GSMUA_Achr6G31080_001	comp36506_c0_seq1	66.69	65.05
YABBY5 (NM_179749.2)	GSMUA_Achr2G12850_001	comp35601_c0_seq1	14.7	15.675
	GSMUA_Achr4G22750_001	comp35601_c0_seq1	26.65	15.675
YABBY4 (INO) (NM_102191.5)	GSMUA_Achr11G17520_001	comp29400_c0_seq1	2.19	4.14
ATS (NM_001125891.1)	GSMUA_Achr4G32400_001	comp4458_c0_seq1	1.185	3.42
KAN2 (NM_102957.3)	GSMUA_Achr10G06430_001	comp30224_c0_seq1	16.44	5.495
KAN3 (NM_117878.2)	GSMUA_Achr2G18070_001	comp33811_c0_seq1	4.265	8.03

CHAPTER 5

Explaining the evolution of morphological diversity in angiosperm flowers: a critical evaluation of reliance on the ABCs

PREFACE

Chapter 5 differs from the rest of this dissertation in the sense that I take a step back and try to look into the research field of floral development and evolution with a historical and philosophical perspective. Here, I address the development of the field, as well as its shortcomings, using Lakatos a source of inspiration.

The motivation for this chapter came from my previous experience in history and philosophy of science, during my Masters Degree. But, I was also driven by the experimental work I have done along these past five years and its incompatibilities with the mainstream research practiced in this field. The results that I present in Chapters 2 through 5, somehow, led me to question the way floral development and evolution has been approached for the past 25 years or so. Here, I try to explain some of the reasons why I feel this way.

This chapter has not been published elsewhere yet, and is summarized below.

The field of plant developmental genetics experienced a renaissance with the discovery of the MADS box genes. Based on mutant phenotypes, the "ABC model" defined a molecular genetic signature for the specification of each of four floral organ identities; sepals, petals, stamens and carpels. This model has since been used as a basis for understanding molecular mechanisms of flower development in model systems.

In addition, it is used as a template for many studies in plant developmental evolution ('evo-devo'), with researchers proposing how 'shifts' or 'changes' to the ABC model could result in the diversity of phenotypes found across flowering plants. As data accumulates, however, it has become clear that the ABC genes are not sufficient for specifying organ identity, even in model systems. Likewise, the diversity of floral organ arrangements and forms cannot be explained by applying the ABC model across angiosperms, particularly when a simple and linear mapping between gene expression and floral organ morphology is assumed.

Using Lakatos' Research Programme as an analytical framework, we argue that a flower evo-devo research program has used the ABC model as its *hard core*, despite well-understood long-term challenges. We explore historical changes to the hard core's *protective belt*, made to accommodate new discoveries. We discuss if the model, with or without these changes, is sufficient to explain the evolution of morphology in angiosperm flowers, and to what extent the ABC model could be applicable to the evo-devo research community.

Finally, we propose that despite the heuristic value of the ABC model itself, this model has a limited explanatory capacity. We argue that non-linear dynamic models that couple gene regulatory networks with physical and chemical fields should be pursued to integrate and explain observed gene expression and morphogenetic patterns, and to put forward novel predictions and testable hypotheses concerning evolution of floral morphology.

INTRODUCTION

The dissection of the genetic mechanisms underlying morphogenesis in *Arabidopsis thaliana* started in the late 1980's, and is a fertile research ground that continues to expand and diversify through today. Initially, *A. thaliana* was established as an ideal model organism for studies on the molecular genetics of many aspects of plant growth and development (Haughn & Somerville 1988), focusing plant biology research towards understanding a single tractable system and thereby leading to a great wealth of knowledge about molecular mechanisms underlying plant morphogenesis. One aspect of plant morphogenesis that benefited immediately from these efforts was focused on the understanding of flower development through targeted studies of floral organ formation.

The first papers to describe the molecular underpinnings of floral organ specification in *A. thaliana,* also paralleled by studies in *Antirrhinum majus* (Sommer et al. 1990; Bowman et al. 1989; Bowman et al. 1991; Coen & Meyerowitz 1991; Tröbner et al. 1992), suggested a simple model in which combinatorial expression of three unique classes of genes leads to the specification of the four floral organs found in the eudicot flower. This model, widely known as the ABC model of floral organ identity, is based on evidence that the expression of A-class genes is necessary for sepal identity; when expressed in combination, A-class and B-class genes specify petal identity; stamen identity is specified by combinatorial expression of B- and C-class genes; and gynoecium identity is specified by expression of C-class genes alone. The ABC model provided researchers in flower development with an elegant and simple framework on which to base their genetic studies. The simplicity of such a model, coupled with its ability to establish an apparent direct causal link from the ABC genes to their expression patterns and from these to the specification of differential floral organ identity dominated the mainstream scientific community, and it has also been successfully transposed to the science classroom (Bowman et al. 2012).

Considering its success in describing floral development and organ identity specification in *Arabidopsis* and *Antirrhinum*, the ABC model has been widely used by the flower evo-devo research community in attempts to understand and explain the evolution of morphological variation across angiosperm flowers. The MADS-box genes comprising the ABC model have been used as candidate genes for flower evo-devo studies across angiosperms, with research focusing on investigating copy number and expression patterns of the A, B, and C-class genes in a diversity of floral forms and broadly assuming that these genes play a conserved role in floral development and developmental evolution. This application of the ABC model to evo-devo research has various problems that we argue have hindered the development of a sustainable evo-devo research program. First, researchers have taken the static ABC model and simplistically applied it to understand morphogenesis, with the implicit assumption that changes or modifications to the model can explain the diversity of floral patterns; this is despite the recognized fact that the ABC model does not describe an ancestral condition on which evolutionary processes may have acted. Secondly, and perhaps most importantly, the model itself assumes a linear and simplistic relationship between the expression of one (or a few) genes and a particular phenotype, which implies a reductionist logic and cannot provide

an explanation for the dynamical mechanisms by which floral morphology can evolve. Taken together, the use of the reductionist account of a static model, such as the ABC model, for evo-devo studies is not adequate, in our view, to generate a successful framework for recovering patterns or processes underlying the evolution of floral morphology or floral organ novelty across angiosperm flowers.

Here, we use Lakatos's account on Research Programmes to further understand the establishment and evolution of the 'mainstream'[1] flower evo-devo research program despite a failure of a reductionist account of the ABC model to elucidate processes of floral development and evolution. We discuss the changes made to the ABC model as ad hoc modifications that are proposed to better fit the model to the observed data. We then discuss whether these changes are indeed useful for the flower evo-devo research community in order to understand the morphological diversification of angiosperm flowers. Finally, we argue that the limitations of the mainstream flower evo-devo research program in explaining the origin and/or diversification of the angiosperm flower is a result of the way the ABC model has been applied to an understanding of floral development and its evolution, rather than due to lack of heuristics of the canonical ABC model itself.

The Establishment and Development of the Mainstream Flower 'Evo-Devo' Research Program

Lakatos' and the establishment of Research Programmes

According to Lakatos, a scientific field can be understood as a programmatic development of knowledge as the result of implications of fundamental principles (Lakatos 1980). In his view, a compilation of fundamental knowledge, along with its peripheral assumptions, implications and methodologies, composes what is called a Research Programme, in which scientists are able to solve problems related to the peripheral assumptions however different their accounts might be. According to Lakatos, a scientist can successfully contribute to the development of the Research Programme as long as challenges to the fundamental principals are avoided. The fundamental principals of a Research Programme are called the *hard core* (Lakatos 1980), and can be understood as the ultimate defining characteristic of any given Research Programme. The hard core is a set of general hypotheses that set the stage for the development of the Research Programme, augmented by supplementary explicit and implicit assumptions that allow

[1] It is important to note that the qualifier 'mainstream' is used here to distinguish flower evo-devo research that uses the ABC model as its hard core and applies the model in a unidirectional and reductionist way, from other research programs (non-mainstream). Based on a rapid assessment of papers on some aspect of flower evo-devo published in the last 20 years, it is clear that the publications on flower evo-devo that simplistically correlate ABC gene expression to flower organ morphology outnumber publications aimed at understanding flower evo-devo using different approaches; thus, this research community is qualified as 'mainstream' in the field of flower evo-devo.

hypotheses to be ultimately tested. Inconsistencies between observations and the Research Programme are, therefore, buffered by the supplementary assumptions preventing direct attacks to the hard core. The supplementary assumptions act as a *protective belt*, rendering the hard core unfalsifiable by "methodological decisions of its protagonists" (Lakatos 1970, p.133). While the supplementary assumptions could be modified to incorporate any inconsistencies between the Research Programme and experimental observation, the hard core would ultimately remain intact.

Another important characteristic of Research Programmes, according to Lakatos (1970; 1980), is the use of heuristics. Heuristics refers to a set of rules that guide the scientific activity, aiding the development of a particular Research Programme. *Positive heuristics* guide scientists on how the protective belt should be modified to better provide explanations to new observations without challenging the hard core. In turn, *negative heuristics* refers to the set of 'prohibited' questions or venues that should not be undertaken by scientists, if they desire to remain loyal to their Research Programmes. The strength of the positive heuristics of a Research Programme provides scientists with clear hypothesis and methods, and together with the protective belt, help the advancement of knowledge and the achievement of new discoveries. The value of a Research Programme, according to Lakatos (1970), is also indicated by its ability to generate novel predictions and the extent to which these predictions are later confirmed. In this sense, a *progressive* Research Programme maintains coherence through time, leading to novel discoveries confirming derived predictions, while *degenerative* Research Programmes do not lead to such advances. The overcoming of a degenerative Research Programme by a progressive Research Programme is what Lakatos considers to be a scientific revolution.

The Establishment of the Flower 'Evo-Devo' Research Program

Since it was first described (Coen & Meyerowitz 1991), the ABC model was considered to be a conserved molecular patterning mechanism underlying flower development in angiosperms, therefore justifying its applicability to the study of flower development in organisms phylogenetically distant from the established model systems. The relative conservation of the ABC gene expression patterns and their correlation with organ specification regions in the flower meristem across phylogenetically distant species (Coen & Meyerowitz 1991; Bowman 1997) set the stage for the development of the flower evo-devo research program.

The flower evo-devo research program arose as a group of scientists, their scientific knowledge and their research tools used to explain the evolution of morphological variation in angiosperm flowers. At the *hard core* of the flower evo-devo research program lies the ABC model of floral organ identity, where combinations of A-, B-, and C-class genes determine the identity of sepals, petals, stamens, and carpels (Figure 1).

For the past 20 years, the central hypothesis of this research program is that the ABC model of floral organ identity and its subsequent modifications (Figure 1-3) are capable of explaining the morphological diversity observed across angiosperm flowers (Litt & Kramer 2010; Melzer et al. 2010; Rijpkema et al 2010; Causier et al. 2010?). Because of

its dependency on the ABC model, the flower evo-devo research program has developed hand-in-hand with the improvements of our understanding of *Arabidopsis thaliana*'s flower development. This results in a feedback loop in which advancements in understanding the molecular mechanisms of the ABC model as applied to *Arabidopsis thaliana* and other established genetic models feed the hard core of the flower evo-devo research community, and are considered to be directly applicable to the understanding of the evolution of flower development.

The mainstream research methodology for the flower evo-devo research program has become a comparative, descriptive account of the expression patterns of ABC genes across distinct floral organs in different angiosperm species (reviewed in Litt & Krammer 2010). This methodology, together with its feedback loop with the core flower development research on *Arabidopsis*, is one of the main components of the *protective belt* of the mainstream flower evo-devo development research program, such that studies that follow this methodology are included in the data supporting the use of the ABC model to explain floral diversity.

This approach, however, has two implicit limitations, frequently overlooked by the evo-devo research community. First, research on the genetics of *Arabidopsis* flower development relies heavily on mutant analysis. The widely accepted use of mutants to address gene function leads to a simplistic interpretation of the mutant phenotype, in which there is considered to be a simple and linear relationship between a mutated gene (or genes) and an observed altered morphology. This simplistic interpretation detaches the gene from its environment and from other interacting agents and assumes the gene functions as an isolated unit. This simplistic account of functional genetic experiments leads to a reductionist interpretation and application of the ABC model, not only in *A. thaliana*, but also in other systems. Second, even if a linear and unidirectional relationship between one or few genes and a phenotype could be established, the transposition of the ABC model to other angiosperm lineages also implies that it is possible to make clear statements of homology between *Arabidopsis* genes and/or morphology to genes and morphologies in other angiosperm lineages. Homology statements between *Arabidopsis* and diverse angiosperm lineages are not straightforward, both at the level of the organs studied and at the level of the genes recovered. Orthologous genes can take on different functions, while paralogs can converge on identical functions in flower development (Kramer et al. 2004). Thus the study of a single gene and its function can at most provide information about that particular class of genes in that particular species, but does not have the power to explain how the floral morphologies have diverged or evolved.

Although there are clear limitations to the dependency on mutant analysis and the ability to make homology statements between genes and phenotypes, this developmental genetic methodology serves as a strong component of the *positive heuristics* of the flower evo-devo research program, as it informs the scientific community of the kind of evidence they should provide in order to maintain successful projects within the research program. By providing strong guidance (positive heuristics) and using advancements in *Arabidopsis* flower development together with ad hoc modifications to the ABC model to

actively protect its hard core, the research program has become strengthened such that is is now the mainstream research program in flower evo-devo.

It is unquestionable that much knowledge has resulted from this program (reviewed in Krizek & Fletcher 2005; Causier et al. 2010). The elegant and simple ABC model at the core as well as the strong protective belt and positive heuristics of the research program have provided scientists with clear hypotheses in order to pursue their research on diverse plant lineages. In fact, the success of this research program is so impressive that sepals, petals, stamens, and gynoecium/carpels are now defined genetically for comparative analyses. While floral organs have long been defined by their positional homology within the flower as well as their macro- and micro-morphologies, function, and ontogenesis, currently A-, B-, and C-class gene expression pattern are considered evidence for assigning organ identity in cases of questionable organ homology. In agreement with others (Jaramillo & Kramer 2007), we believe that this conceptual standpoint has important implications for studies on the evolution of floral traits, as statements of homology can be based on gene expression similarities between organs of different plants in the place of (or in addition to) classical morphological and developmental studies.

Advances to the hard core of the Mainstream Flower 'Evo-Devo' Research Program

According to the main proponents (Bowman et al. 2012), the ABC model promoted significant scientific advances, among which are the discovery that (1) the A-, B-, and C-class genes, with the exception of *APETALA-2*, are all MADS-box transcription factors; (2) another family of MADS-box genes, the *SEPALLATA* or E-class genes, are also essential for the establishment of organ identity during flower development; and (3) these MADS-box genes act as multimeric complexes to bind and regulate downstream gene expression (Tröbner et al. 1992; Pelaz et al. 2000). The protein-protein interactions necessary for adequate functioning of the MADS-domain proteins was later formalized in what has been known as the Quartet model (Theissen & Melzer 2007). Along with these advances, the ABC model of organ identity has also been widely used to "explain" the origin of the flower and the subsequent morphological diversification of flowers across angiosperms. Below, we will place each of these advances within the context of the flower evo-devo research program.

<u>The SEPALLATA (SEP) genes and the function of A-class genes in floral organ specification</u>

Almost 10 years after the publication of the ABC model, a new important component of the floral organ identity mechanism was discovered: the *SEPALLATA* (*SEP*) genes (Pelaz et al. 2000). In *A. thaliana*, SEP genes are involved in the establishment of all floral organs (Figure 2a). Once again, *Arabidopsis* mutants identified as defective in floral organ development were instrumental to this discovery. Although the proposed involvement of *SEP* genes in flower development modified the ABC model, and therefore, the hard core of the mainstream flower evo-devo research program by the addition of a new class of genes necessary for establishing floral organ identity, this change was widely accepted by the scientific community. Three main factors may

underlie this prompt acceptance. First, the evidence provided by Pelaz and collaborators (2000) was compelling, and in full agreement with the accepted methods used within the research program. Second, the addition of *SEP* genes to the hard core did not contradict the model *per se*; to the contrary, the addition of the *SEP* genes expanded the research program's hard core in full agreement with the established ABC model. Lastly, the acceptance of the SEP genes into the hard core provided strength to the heuristics of the research program, while augmenting the power of its protective belt. In this sense, a new class of genes, E-class genes, could be used to illuminate a new set of hypotheses about the molecular underpinnings of the variation in flower morphology across angiosperms.

In contrast, research on the A-class function has proven to be more problematic for the stability of the research program. While in the ABC model the A-class genes are responsible for sepal and petal identity (Figure 1), later research on the role of A-class genes in perianth identity outside of *Arabidopsis* did not fully support such claims (e.g., Litt 2007). Instead, A-class function in organ identity was found to be more complex, and A-class genes themselves are involved in a variety of distinct developmental processes outside of their role in floral organ formation (for a detailed review, see Causier et al. 2010). In addition, A-class function was demonstrated to be not required for petal identity specification in *Arabidopsis thaliana*; rather, its expression is more closely linked with the establishment of floral meristem identity (which in turn facilitates the establishment of sepals) and to promoting the boundaries of B- and C-class gene expression (Causier 2010; Dinh et al. 2012). It is interesting to note that, in the case of A-class genes, the research steaming from the flower evo-devo research program was able to inform the scientific community about limitations of the ABC model to contribute to our understanding of the diversification of angiosperm flowers.

If A-class genes were repositioned on the ABC model according to the new observations (Figure 2a), the model becomes extremely similar to the one proposed by Schwartz-Sommer and collaborators (1990). Although contemporary to the main publications that establish the ABC model in *Arabidopsis thaliana*, the BC model proposed by Schwartz-Sommer and collaborators (1990) was supplanted by the ABC model. If we take into account the new findings, "[t]his concept of the (A)-function enables the (A)BC model to regain its widespread applicability and provides a framework with which the existing mutants can be interpreted" (Causier et al. 2010, p.78). The A-class function might represent an instance in which sociological factors, more than scientific ones, decided the history of scientific advancements (Fleck 1979).

The Quartet model

With the exception of *APETALA-2*, the canonical ABC genes are MIKC-type transcription factors. Such genes encode proteins that bear a MADS-domain that binds DNA in order to regulate the transcription of downstream genes (Krizek & Meyerowitz 1996). These proteins, however, bind DNA in a protein complex form where dimers (Yang et al. 2003), and subsequently tetramers, must be properly formed for transcriptional activity to take place. Thus, protein-protein interactions (PPI) are essential to the actual function of the ABC model (Tröbner et al. 1990; Honma & Goto 2001;

Pelaz et al. 2001). These interactions comprise the main thesis of the Quartet model (Theissen & Melzer 2007; Melzer & Theissen 2009; Immink et al. 2010). In this model, the main floral homeotic genes interact to form dimers and tetramers (quartets) in order to function as transcriptional regulators (Figure 2b).

In vitro and *in vivo* studies support the idea that homeotic proteins need to interact in order to function properly (reviewed in Immink et al. 2010). The addition of the quartet model to the mainstream flower evo-devo research program proved extremely valuable, as it greatly enhanced the program's heuristics, by incorporating a distinct set of knowledge and techniques, as well as a new set of mechanistic hypotheses about the molecular basis of morphological variation. The quartet model was consistent with the ABC model and added strength to the hard core of the evo-devo research program by adding functional significance that described at the protein level the mechanisms by which these canonical genes could function to define organ identity. It promoted the advancement of the role of *SEP* proteins as mediators of quartet formation (Causier et al. 2003), as well as a mechanism by which the floral identity genes perform its functions. It also provided scientists with new hypothesis concerning the evolution of the homeotic protein-protein interactions, and an entire research field is now devoted to the understanding of how the evolution of these PPI might explain the morphological diversity of angiosperm flowers. The formation of multimeric protein complexes by the ABC model genes could illuminate the molecular mechanisms by which flowers diversified. Hence, "[m]ultimer formation in different combinations of regulatory proteins can be a mechanistic basis for the origin of novel regulatory functions and a gene regulatory mechanism for the appearance of morphological innovations" (Hernández-Hernández et al. 2007, p. 465).

Although the Quartet Model described for *Arabidopsis thaliana* requires obligate heterodimerization, it was shortly noted by the evo-devo research community that this was not the case across angiosperms (Winter et al. 2002). Obligate heterodimerization of ABC gene products evolved from a homodimerization state, where protein products from the same gene family are capable of interacting to form multimeric complexes in non-model plants. The fact that these proteins can homodimerize and are perhaps functional in this state (or may function as a dominant negative) indicates that detailed studies of protein-protein interactions and subsequent DNA binding affinities are required to determine how these products are functioning in combination in each floral organ type in order to determine their role in floral development and the evolution of floral organ morphology. This is a vast change from descriptive studies of gene expression, and demonstrates a required shift in the methodologies used to test hypotheses of floral evo-devo. We argue that gene expression data alone will not provide the necessary information to determine the regulation of morphogenesis, and only when coupled within a more sophisticated framework will one be able to explain the evolutionary patterns of flower morphogenesis.

Diversity-based modifications to the ABC model

While the ABC model was defined in *Arabidopsis*, a highly derived eudicot lineage, it was assumed that the model was so fundamental to floral development that it could be used to explain floral organ identity across angiosperms. Attempting to use the canonical ABC model to explain the diversity of floral morphology, flower evo-devo researchers proposed modifications to the ABC model that could explain the evolution of morphological diversity, assuming that floral diversity could be evolved by simply modifying the static ABC model (Figure 3). These modifications did not contradict the ABC model but rather were proposed to make the interpretation of the spatial components of *A. thaliana* ABC model flexible enough so as to encompass the diversity of floral organizations observed. Different types of organization and organ structures required different modifications to the model, such that different modifications were proposed in a lineage-specific manner to accommodate observed trends in floral developmental evolution.

One such model, the 'fading boarders' model (Buzgo et al. 2004), was proposed to explain floral morphology in which there is a gradual transition of floral organs along a spiral axis, as is the case in early diverging lineages of flowering plants ('basal angiosperms'). In this model, "[the] organ identity gene[s] [are] broadly expressed across the floral meristem but only weakly at the outer and inner limits of its expression" (Buzgo et al. 2004, p. 943). Therefore, gradual transitions in floral organ morphology were proposed to result from gradients of homeotic gene expression levels across the floral meristem (Figure 3A). This proposal extended the view that one can explain floral organ specification by mapping gene expression to phenotype (or floral morphology) in a unidirectional and linear fashion without understanding how the gene expression patterns arise.

Other modification, exemplified by the 'shifting border' (Bowman 1997) and 'sliding boundary' models (Kramer et al. 2003), was proposed to account for the morphological variation observed in certain petaloid monocots where first (sepals) and second (petals) whorl organs share a similar morphology ('tepals'), or when petaloid organs are present in the first whorl as in some basal eudicot lineages (Figure 3B). According to this proposed model, B-class gene expression would be extended into the first whorl, converting first and second whorl organs to a single organ identity. The observed consistency between ABC gene expression patterns and the morphological variation across angiosperm lineages, points to the fact that, although the ABC model is capable of generating useful correlations (and, therefore, having some predictive capacity), these correlations never constituted, in themselves, an explanation for the patterns observed. These proposed modifications to the ABC model were generated by researchers trying to explain evolution of floral forms outside the traditional model systems, using the ABC model as their foundation. These modifications to the ABC model lead to an expansion of the hard core of the flower evo-devo research program and proved both fruitful and progressive (*sensu* Lakatos) for more than two decades.

The proposed (and in some cases experimentally tested, as for example in Sharma & Kramer (2013)) spatial modifications to the canonical ABC model contributed to the perceived applicability of the model across angiosperm lineages and thus effectively

strengthened the flower evo-devo research program. Extension of the ABC model to explain a diversity of floral morphologies resulted in the inclusion of researchers working on a variety of plant lineages with a diversity of morphological variation that could be attributed to 'changes' in the ABC model (e.g., Mondragón-Palomino & Theissen 2011).

As research moves towards understanding the fundamental processes driving floral organ formation and evolution, the limitations of the mainstream flower evo-devo research program are becoming increasingly evident. One such limitation comes from the reification of the ABC model (and its modified versions) to a reductionist extreme. This reification leads to an exacerbated focus on MADS-box genes as agents of floral development and thus developmental evolution. Although MADS-box genes are undoubtedly important genetic components of flower development, the singular focus on this family of genes comes at the expense of a broader focus on the role of other gene families in flower organ specification. A second shortcoming comes from the focus on organ identity, which is pronounced in *Arabidopsis* but is much less discrete and more variable across other flowering plant lineages. Discordant patterns of gene expression and floral organ morphology have long been described in the literature (Kanno et al. 2007; Landis et al. 2012; Almeida et al. 2013) (for example, Figure 4) and might point to the fact that other morphogenetic processes, such as the establishment of organ polarity or primordial positioning, may, in fact, contribute substantially to the morphological variation observed in the angiosperms.

For most flowering plants, the identity of a floral organ is defined by both organ position within the flower and organ structure or function. In *Arabidopsis*, these features are perfectly correlated such that the position (1^{st}, 2^{nd}, 3^{rd} or 4^{th} whorl) and the structure/function (sepal, petal, stamen, carpel) together can define what has been called organ identity. While specification of organ position and organ morphology involve various gene networks that may act in parallel (review in Alvarez-Buylla et al. 2010a), the strict 1:1 correlation of position and structure/function in *Arabidopsis* gives the false impression that organ identity is a feature that can be derived from a single pattern of gene expression. In fact, outside *Arabidopsis* and in plants where organ position, structure and function vary independently (e.g. tepals of tulips; staminodes in *Passiflora*; inside-out flower of *Lacandonia;* petaloid staminodes of the Zingberales), the ABC model fails to provide a model that can explain the *evolution* of such patterns, although it could have inspired interesting studies and hypotheses concerning the role of ABC genes in these species. Even for *Arabidopsis*, many (see, for example, Alvarez-Buylla et al. 2010a) will argue the ABC model fails to provide an *explanation* of how the ABC gene expression pattern arises, and how this expression is correlated to that of other molecular components within the same or additional gene regulatory modules, as well as to physical and chemical fields that give rise to the sub-differentiation of the floral meristem and subsequent floral organ morphogenesis.

As fundamental as this might seem, this is not the only challenge. Gene and/or whole genome duplications, extensively spread across the angiosperm phylogeny, present evolutionary developmental scientists with an even greater challenge. The functional outcomes of duplicated genes are varied, and most times difficult to uncover due to

subsequent gene loses or phylogenetic and/or morphological distance to *Arabidopsis thaliana* (Kramer & Hall 2005; Irish 2006; Litt & Kramer 2010; Rijkema et al. 2010). The ABC model of organ identity in *Arabidopsis thaliana* is characterized by the fact that each family of genes in the model comprises a single gene. This single-gene model, although extremely elegant and suitable for *Arabidopsis*, reveals itself too simplistic in face on the multitude of genome duplications that have been described during the evolution of the angiosperms. These complex patterns of gene duplications followed by sub-or neo-functionalization in distinct angiosperm lineages (see, for example, Sharma & Kramer 2013) present additional challenges to the flower evo-devo research community when using the ABC model in order to understand developmental evolution (for an example of such complexity refer to Litt & Kramer 2010, Figure 3).

The Flower Evo-Devo Research Program; searching to explain the evolution of morphological variation across angiosperms

It is unquestionable that the ABC model has provided the flower evo-devo research program with a heuristic framework that stimulated floral developmental studies. It also provided a set of candidate genes that formed a foundation for evo-devo studies during the past two decades. As a result, impressive amounts of valuable data have been produced within the scope of the mainstream flower evo-devo research program that continues to facilitate comparative studies of floral development across angiosperms. Many of these experimental observations resulted in proposed modifications to the ABC model in an attempt to modify and adapt the *Arabidopsis*-based model and make it applicable to other angiosperm lineages.

Overall, however, the flower evo-devo research program has been limited in its ability to define common mechanisms that underlie floral developmental evolution (Kanno et al. 2007; Almeida et al 2013). Thus, the establishment and development of a mainstream flower evo-devo research program with the ABC model as its hard core begs at least two interesting questions: (1) can the proposed modifications to the ABC model *explain* the evolution of morphological variation across angiosperm flowers? And if so, (2) why are so many modifications proposed to the ABC model in order to account for an explanation of floral morphological variation across angiosperms?

In our view, these questions are part of a single issue that exemplifies the way in which the flower evo-devo research program has interpreted and applied the ABC model. A reductionist approach, in which one or few genes are implicitly or explicitly assumed to be sufficient causal factors to explain wild type and mutant phenotypes as a consequence of a direct, simple and linear mapping of genes to their expression patterns and to floral organ specification and morphogenesis lies at the heart of the limitations of the flower evo-devo research program. The flower evo-devo research program has approached the ABC model and its modifications through this reductionist perspective, assuming that a simple combinatorial expression of few genes is capable of generating the variety of floral organ morphogenetic patterns observed in the angiosperms. However, such a descriptive application of the ABC model (and its modifications) in non-model organisms is not sufficient either, in our view, to explain the morphological diversification of the

angiosperm flower (see, for example, Espinosa-Soto et al. 2004; Alvarez-Buylla et al. 2007; Alvarez-Buylla et al., 2010b). Using such a simplistic application of the ABC model the flower evo-devo research program assumes that changes in expression patterns of particular candidate genes can lead directly and explicitly to the evolution of morphological diversity. This, however, limits researchers to studying organs of defined homology and completely restricts research that is interested in investigating the evolution of novelty.

It is important to clarify that the ABC gene expression patterns and their correlations to the sites of floral organ specification hold for model systems and many other species. Nonetheless, this does not imply that the ABC model itself *explains* floral morphogenesis. Instead, it only describes gene expression patterns and their correlation with morphogenetic patterns. Likewise, the modifications made to the ABC model also describe true patterns in other systems, but remain short at explaining floral organ specification in such lineages. Hence, we are still lacking models with greater explanatory capacity of both the conserved patterns, as well as the variations around them.

Furthermore, we believe that the same limitations regarding the applicability of the ABC model to our understanding of flower development outside *Arabidopsis* can also be seen in the proposed modified models and their inability to describe flower evolution. In fact, the ABC model as well as its modifications do not provide an *explanation* of how different flower morphologies arise or evolve (Espinosa-Soto et al. 2004; Alvarez-Buylla et al. 2010b). In this sense, these modifications, although of localized heuristic value as that of the original ABC model, are not, in our view, capable of providing a mechanistic and dynamic explanation of development, which is a prerequisite to address the evolution of morphological variation in the angiosperm flower. Although the research in flower evo-devo uncovered an interestingly conserved pattern of ABC gene expression, at least to a certain extent, it does not explain how these patterns arise in *Arabidopsis* nor during angiosperm evolution.

We believe, along with others, that research on the evolution of plant development is best guided by a focus on how patterns like that described by the ABC model can emerge during development as a result of complex and highly non-linear gene interactions and their feedback from and to physical and chemical fields (see for example, Espinosa-Soto et al., 2004; Barrio et al., 2010) and why certain patterns of gene expression such as those characterized by the ABC model are fairly robust across angiosperm evolution (Alvarez-Buylla et al., 2010b). Such studies will contribute to a broader understanding of how and why the overall floral bauplan is widely conserved among flowering plants and how changes in such dynamic gene regulatory networks or fields, yield altered patterns.

It has long been proposed (Mendoza & Alvarez-Buylla 1998; Mendoza et al. 1999) that the ABC model for floral organ identity (and its modifications) must be incorporated into a broader, dynamic, and integrative framework that is able to consider non-linear interactions and reciprocal causation, and hence provide researchers with new hypotheses. A dynamic, generative (rather than static) model is capable of integrating

large amounts of data from different fields, and placing various signals and constraints into a temporally and spatially dynamic framework that can be used to model morphogenesis from fundamental principles (eg., Alvarez-Buylla et al. 2008; Barrio et al., 2013). In this approach, the ABC patterns and their conservation among several lineages of angiosperms is explained using dynamical complex systems of interactions that are necessary *and sufficient* to explain not only the correlation between gene expression patterns and morphological traits in *Arabidopsis,* but also the conservation of entire gene regulatory networks that is observed across angiosperms and the factors that lead to morphological variations (reviewed in Alvarez-Buylla et al. 2010b).

A complex systems approach to flower development and evolution has become an increasingly important research field (for examples of important publications in this field, view Mendoza & Alvarez-Buylla 1998; Mendoza et al. 1999; Espinosa-Soto et al. 2004; Alvarez-Buylla et al. 2010b; Barrio et al. 2010; Bruijn et al. 2012; Hernández-Hernández et al. 2012; Lian & Mahadevan 2012; Posé et al. 2012; Takeda et al. 2013). In line with these ideas, we agree that only when flower development and the evolution of floral form are placed in the context of a complex systems approach, one in which biological phenomena are intrinsically irreducible to one or few genes, will we have the necessary *and sufficient* framework to start uncovering the mechanisms that explain both the conserved, robust themes of floral development and the processes that underlie the evolution of the amazing diversity of floral forms that characterizes the angiosperms.

Epilogue

It is important to mention that Lakatos' Research Programme was used here as an epistemological framework on which to analyze the establishment and development of the flower evo-devo research field. Although inspiring, Lakatos' Research Programme has its own shortcomings and might not fully represent the complexities of the scientific community or the intricate development of the entire body of scientific knowledge. Therefore, the structure of the analysis presented here serves more as an analytical framework, rather than a rigid categorization of the development of this rich research field. The main purpose of such an analysis is, however, to dissect the way in which research has been carried out by the flower evo-devo community, as well as to understand the contributions and limitations of the ABC model to our understanding of evolutionary developmental biology in angiosperms. Based on our analysis, we argue that rather than abandoning the ABC model, the main proposition here is to put the ABC model into a broader context that expands its heuristics, perhaps resulting in a more progressive Research Programme that aims at understanding the mechanisms underlying development and hence the evolution of morphological variation in angiosperms, and that is capable of explaining why the ABC model, despite its simplicity, holds true in a wide range of cases and despite the fact the ABC genes interact with many other components. An integration of the ABC model into a complex system's approach is already formulated in the literature. Here we argue, in line with others (reviewed in Alvarez-Buylla et al. 2010b), that such integration provides a mechanistic framework with explanatory power capable of expanding the positive heuristics of the flower evo-devo research program.

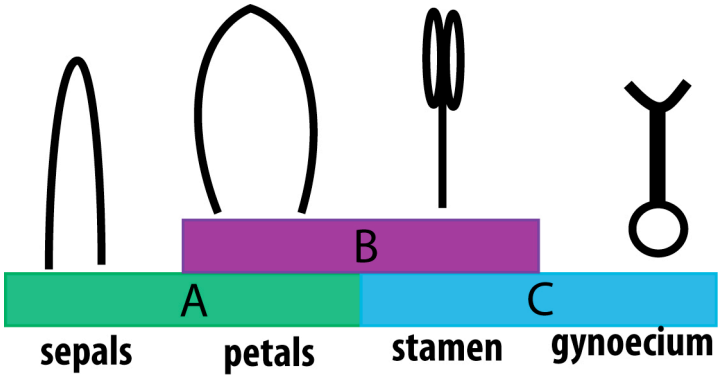

Figure 1. The ABC model that lies in the hard core of the mainstream flower evo-devo Research Programme. The ABC model of floral organ identity determines that organ identity is set during floral development as a result of combinatorial expression of three classes of genes: A-class, B-class, and C-class genes. In *Arabidopsis thaliana*, A-class genes are represented by APETALA-1 and APETALA-2; B-class genes comprise APETALA-3 and PISTALLATA; while AGAMOUS is the single C-class gene. In the ABC model, sepals are specified by isolated expression of A-class genes, petals arise at the domain of A- and B-class expression, stamens develop whenever B- and C-class genes are expressed in combination, while gynoecium identity is specified by the expression of C-class genes alone.

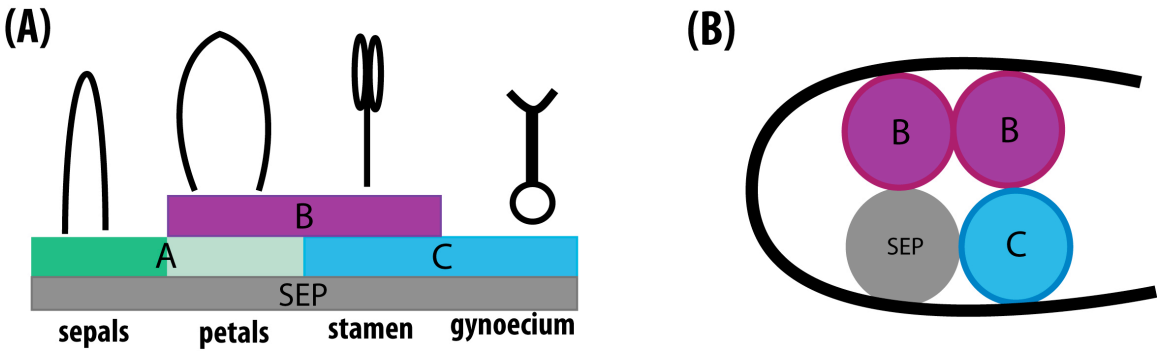

Figure 2. (A) The first modifications to the classical ABC model of floral organ identity. Two main modifications are depicted: (1) addition of the MADS-box family of SEPALLATA (SEP) genes, in grey. SEP genes are involved in specifying all floral organs in *Arabidopsis thaliana* (Pelaz et al. 2000), and since its discovery, the SEP genes has been considered master genes in floral development; (2) the discovery that A-class genes might not function to establish petal identity, but are rather related to floral organ initiation (Litt 2007; Causier et al 2010). This discovery is depicted as different green colors, where a lighter color relates to its uncertain role in establishing petal identity, while is darker green in the first whorl depicts its role in floral organ initiation. (B) The Quartet model in which A-, B-, C-, and E-class gene products interact to form a protein tetramer that binds DNA to regulate downstream genes expression (Theissen & Melzer 2007; Melzer & Theissen 2009; Immink et al. 2010). Here, we depict a hypothetical tetramer (quartet) former during stamen development, where B-, C-, and E-class proteins interact to bind DNA (looped black line).

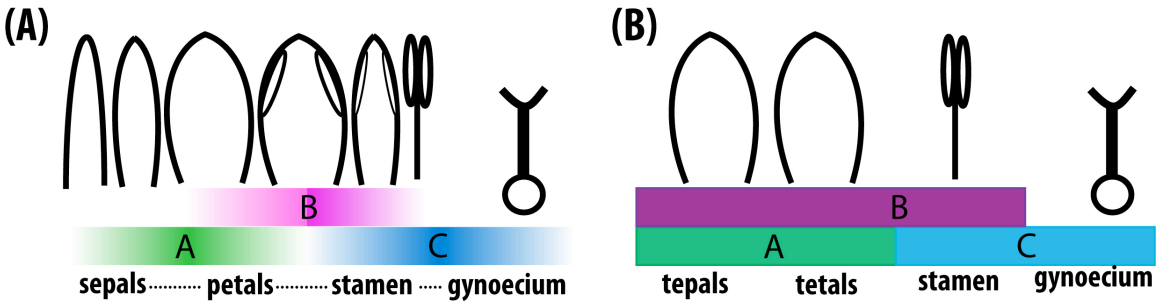

Figure 3. Other modifications of the classical ABC model of floral organ identity. (A) The fading borders model (Buzgo et al. 2004), proposed to explain gradual transitions of floral organs observed in basal angiosperm groups, suggest that expression pattern of the classical homeotic genes gradually fades away as it overlaps with expression domains of other homeotic genes. This will result in gradual expression of identity genes across the floral meristem, mirroring the morphological transitions observed in basal angiosperm flowers. (⋯⋯) depicts gradual transitions between different organ types. (B) The 'shifting border' or the 'sliding boundary' models suggest that B-class gene expression expands to the first whorl of the flower, rendering first and second whorl organs with similar morphology. Here, we exemplify this model for 'petaloid monocots' in which first and second whorl organs have similar petaloid morphology, and are usually called tepals. This model can also be applied for some basal eudicot lineages where petaloid organs are observed in the first whorl.

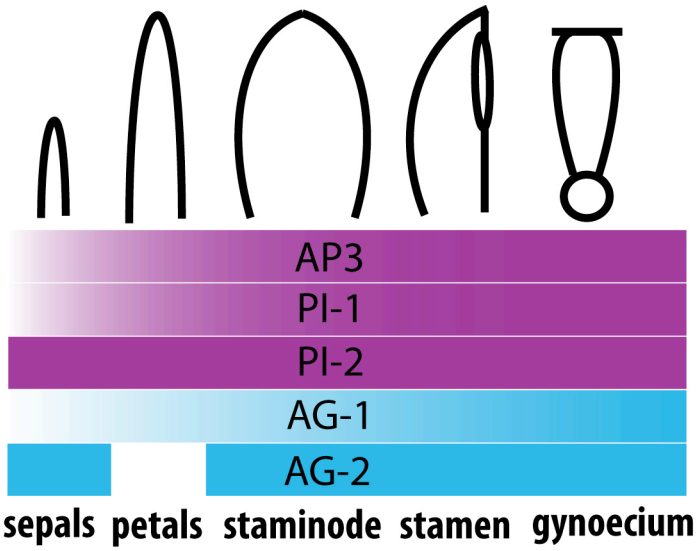

Figure 4. *Canna indica* B-, and C-class gene expression patterns in the developing flower. *Canna indica* flowers are composed by three sepals, three petals (with a sepaloid appearance), four petaloid infertile stamens (with a petaloid appearance), and one-half of a fertile stamen. Multiple gene copies of B- and C-class genes have already been identified and their expression pattern on the developing flower greatly diverges from that expected by ABC model (modified from Almeida et al. 2013). In this case, and certainly in many others, the direct application of the ABC model to understand the identity of the flower organs is clearly insufficient. In the case of *Canna indica*, we argue that more data needs to be integrated into the model, if one is to understand flower morphology in this species.

CHAPTER 6

Concluding Remarks

The dissertation presented here aimed at uncovering the molecular developmental mechanisms that shape the evolution of floral morphology in the Zingiberales. In order to achieve such task, I used different approaches, ranging from long lasting classical morphological studies, to more recent molecular techniques, to the latest sequencing technology.

Chapter 2 uses classical flower developmental series to establish the homology of the petaloid appendage of the fertile stamen in the ginger families, while pointing out the inconsistencies of the molecular data presented for *Canna indica* and the current accepted 'model' of floral organ specification. Chapter 3, in turn, digs deep into the AGAMOUS gene family evolution across Zingiberales, leading to the proposition of a mechanism that can potentially explain the evolution of morphological variation in ginger flowers. Then, organ-specific transcriptomes were generated for different species across the order, and using the latest sequencing technology, I propose, on Chapter 4, the co-option of the abaxial-adaxial polarity network for shaping filament morphology in the Zingiberales.

While performing the experiments described in this dissertation, however, it became increasingly clear to me that the epistemological framework within which you design your experiments and interpret your results has a much greater impact on your work than the methodologies used to collect the data. Although it is irrefutable that technological innovations can improve the speed with which we generate data, and hopefully, scientific knowledge, these advancements *per se* are not sufficient to improve our knowledge in a particular field. Only when we look at old problems with new epistemological frameworks can we uncover new understandings while generating more evidence but also revisiting old data. Chapter 5 in this dissertation is a result of such philosophical account. There, I decided to take a step back and look at the flower evo-devo research, trying to better understand the history of this research field for the past 25 years, as well as pointing out its shortcomings.

While using the amazing androecial petaloidy of the Zingiberales as a study case, however, I was able to uncover developmental processes that go beyond the Zingiberales order. For example, I show that the abaxial-adaxial polarity gene regulatory network is involved in shaping not only Zingiberales filament morphology but, I present evidence that this network might also be involved in shaping filament morphology across angiosperms. Also, the regulatory mechanism proposed for the AGAMOUS gene in the Zingiberales is potentially applied to all flowering plant, and has not yet been proposed in the literature.

Last, I'd like to highlight that the knowledge I generated in this dissertation steams in great part from the advantage of focusing my study on a clade with clear phylogenetic relationships. Once the evolutionary history (or at least, the best hypothesis for how the

lineages are related in evolutionary time) is clearly sorted out, we can make homology statements of any value between characters or genes. I believe that only within a clade-based approach, on which to conduct careful morphological and developmental characterizations and rigorous molecular analysis can one uncover the *evolution* of developmental molecular mechanisms. Only then can we understand how these changes can lead to the diversity of forms we see in nature.

I hope this dissertation can serve as an example of such approach.

BIBLIOGRAPHY

Airoldi, C.A.; Bergonzi, S.; Davies, B. 2010. Single amino acid change alters the ability to specify male or female organ identity. PNAS, 107(44): 18898–18902.

Almeida, A.M.R.; Brown, A.; Specht, C.D. 2013. Tracking the development of the petaloid fertile stamen in *Canna indica*: insights into the origin of androecial petaloidy in the Zingiberales. AoB PLANTS 5: plt009; doi:10.1093/aobpla/plt009.

Alvarez-Buylla, E.R.; Benítez, M.; Dávila, E.B.; Chaos, A.; Espinosa-Soto, C.; Padilla-Longoria, P. 2007. Gene regulatory network models for plant development. Current Opinion in Plant Biology, 10(1): 83-91.

Alvarez-Buylla, E.; Chaos, A.; Aldana, M.; Benítez, M.; Cortes-Poza, Y.; Espinosa-Soto, C.; Hartasánchez, D.A.; Lotto, R.B.; Malkin, D.; Sntos, G.J.E.; Padilla-Longoria, P. 2008. Floral morphogenesis: Stochastic explorations of a gene network epigenetic landscape. PLoS ONE 3: e3626.

Alvarez-Buylla, E.R.; Benítez, M.; Corvera-Poiré, A.; Chaos, A.C.; de Folter, S.; de Buen, A.G.; Garray-Arroyo, A.; García-Ponce, B.; Jaimes-Miranda, F.; Pérez-Ruiz, R.V.; Piñeyro_Nelson, A.; Sánchez-Corrales, Y.E. 2010a. Flower Development. IN: The Arabidopsis Book. American Society of Plant Biologists. doi: 10.1199/tab.0127.

Alvarez-Buylla, E.R.; Azpeitia, E.; Barrio, R.; Benítez, M.; Padilla-Longoria, P. 2010b. From ABC genes to regulatory networks, epigenetic landscapes and flower morphogenesis: Making biological sense of theoretical approaches. Seminars en Cell & Developmental Biology, 21:108-117.

Barker, S.C.; Robinson-Beers, K.; Villanueva, J.M.; Gaiser, J.C.; Gasser, C.S. 1997. Interaction among genes regulating ovule development in *Arabidopsis thaliana*. Genetics 145:1109-1124.

Barrio, R.Á.; Hernández-Machado, A.; Varea, C.; Romero-Arias, J.R.; Álvarez-Buylla. 2010. Flower Development as an Interplay between Dynamical Physical Fields and Genetic Networks . PLoS ONE 5(10): e13523. doi:10.1371/journal.pone.0013523.

Barrio, R.A.; Romero-Arias, J.R.; Noguez, M.A.; Azpeitia, E.; Ortiz-Gutiérrez, E.; Hernández-Hernández, V.; Cortez-Posa, Y.; Alvarez-Buylla, E.R. 2013. Cell Patterns Emerge from coupled chemical and physical fields with cell proliferation dynamics: The *Arabidopsis thaliana* root as a study system. PLoS Computational Biology, 9(5): e1003026. doi:10.1371/journal.pcbi.1003026.

Bartlett ME, Kirchoff BK, Specht CD. 2008. Epi-illumination microscopy coupled to in situ hybridization in non-model species. Development Genes and Evolution 218:273–279.

Bartlett, M.E.; Specht, C.D. 2010. Evidence for the involvement of GLOBOSA-like gene duplications and expression divergence in the evolution of floral morphology in the Zingiberales. New Phytologist, doi: 10.1111/j.1469-8137.2010.03279.x

Bowman J. L., Smyth D. R., Meyerowitz E. M. 1989. Genes directing flower development in *Arabidopsis*. *Plant Cell* 1, 37-52.

Bowman JL, Smyth DR, Meyerowitz EM. 1991. Genetic interactions among floral homeotic genes of Arabidopsis. Development 112:1–20.

Bowman JL. 1997. Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. Journal of Biosciences, 22:515–527.

Bowman, J.L; Smyth, D.R. 1999. CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. Development, 126(11):2387-96.

Bowman, J.L.; Smyth, D.R.; Meyerowitz, E.M. 2012. The ABC model of flower development: then and now. Development 139: 4095-4098.

Bremer B, Bremer K et al. (The Angiosperm Phylogeny Group). 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society 161:105–121.

Bruijn, S.; Angenent, G.C.; Kaufmann, K. 2012. Plant 'evo-devo' goes genomic: from candidate genes to regulatory networks. Trends in Plant Science, 17(8): 441-447.

Buzgo, M.; Soltis, D.E.; Soltis, P.S. 2004. Floral developmental morphology of Amborella trichopoda (Amborellaceae). International Journal of Plant Sciences, 165: 925–947.

Causier, B.; Cook, H.; Davies, B. 2003. An Antirrhinum ternary complex factor interacts with C-function and SEPALLATA-like MADS-box factors. Plant Molecular Biology, 52: 1051-1062.

Charlton WA, MacDonald AD, Posluszny U, Wilkins CP. 1989. Additions to the technique of epi-illumination light microscopy for the study of floral and vegetative apices. Canadian Journal of Botany 67:1739–1743.

Coen, E. Goethe and the ABC model of flower development. 2001. C.R. Acad. Sci. Paris, Sciences de la vie/Life Sciences 324: 523–530

Coen, E.S.; Meyerowitz, E.M. 1991. The war of the whorls: Genetic interactions controlling flower development. Nature, 353: 31-37.

Davies, B.; Motte, P.; Keck, E.; Saedler, H.; Sommer, H.; Schwarz-Sommer, Z. 1999.

PLENA and FARINELLI: Redundancy and regulatory interactionsbetween two *Antirrhinum* MADS-box factors controlling flower development. EMBO J., 18:4023–4034.

D'Hont, A. et al. 2012. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature, 488: 213-219.

Dihn, T.T.; Girke, T.; Liu, X.; Yant, L.; Schmid, M.; Chen, X. 2012. The floral homeotic protein APETALA2 recognizes and acts through an AT-rich sequence element. Development, 139(11):1978-86. doi: 10.1242/dev.077073.

Dreni, L.; Pilatone, A.; Yun, D.; Erreni, S.; Pajoro, A.; Caporali, I.; Zhang, D.; Kater, M.M. 2011. Functional Analysis of All AGAMOUS Subfamily Members in Rice Reveals Their Roles in Reproductive Organ Identity Determination and Meristem Determinacy. The Plant Cell, 23: 2850–2863.

Eichler AW. 1875. Blu¨thendiagramme construirt und erla¨utert, Vol. 1. Leipzig: W. Engelmann.

Eshed, Y.; Baum, S.F.; Bowman, J.L. 1999. Distinct mechanisms promote polarity establishment in carpels of Arabidopsis. Cell, 99(2): 199-209.

Espinosa-Soto, C.; Padilla-Longoria, P.; Alvarez-Buylla, E. 2004. A Gene Regulatory Network Model for Cell-Fate Determination during Arabidopsis thaliana Flower Development That Is Robust and Recovers Experimental Gene Expression Profiles. Plant Cell, 16: 2923–2939.

Fleck, L. 1979 [1935]. Genesis and Development of a scientific fact. University of Chicago Press Chicago.

Floyd, S.K.; Zalewski, C.S.; Bowman, J.L. 2006. Evolution of Class-III Homeodomain Leucine Zipper genes in Streptophytes. Genetics, 173: 373–388.

Glinos E, Cocucci AA. 2011. Pollination biology of Canna indica (Cannaceae) with particular reference to the functional morphology of the style. Plant Systematics and Evolution 291:49–58.

Goethe JW. 1790. Versuch die Metamorphose der Pflanzen zuerkla¨ren. Gotha, Germany: Ettinger.

Goto, K.; Kyozuka, J.; Bowman, J.L. 2001. Turning floral organs into leaves, leaves into floral organs. Current Opinion in Genetics & Development, 11:449–456.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length

transcriptome assembly from RNA-seq data without a reference genome. 2011. Nature Biothecnology, 29(7):644-652.

Gramzow, L.; Theissen, G. 2010. A hitchhicker's guide to the MADS world of plants. Genome Biology, 11:214. http://genomebiology.com/2010/11/6/214

Grimm, S. and Pflugfelder, G.O. 1996. Control of the gene optomotor-blind in Drosophila wing development by decapentaplegic and wingless. Science, 271: 1601-1604.
Gustafson-Brown, C.; Savidge, B.; Yanofsky, M.F. 1994. Regulation of the Arabidopsis Floral homeotic gene APETALA1. Cell, 76: 131-143.

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

Haughn, G.W.; Somerville, C.R. 1988. Genetic Control of Morphogenesis in *Arabidopsis*. *Developmental Genetics*, 9:73-89.

Heijmans, K.; Ament, K.; Rijpkema, A.S.; Zethof, J.; Wolters-Arts, M.; Gerats, T.; Vandenbussche, M. 2012. Redifining C and D in Petunia ABC. The Plant Cell, 24: 2305–2317.

Heijmans, K.; Morel, P.; Vandenbussche, M. 2012. MADS-box genes and floral development: the dark side. Journal of Experimental Biology, 63(15): 5397-5404.

Hernández-Hernández, T.; Martínez-Castilla, L.P.; Alvarez-Buylla, E.R. 2007. Funciontal Diversification of B MADS-box homeotic regulators of floral development: adaptative evolution in protein-protein interactions domains after major gene duplication events. Mol. Biol. Evol., 24(2): 465-481.

Hernández-Hernández, V.; Niklas, K.J.; Newman, S.A.; Benitez, M. 2012. Dynamical Patterning Molecules in plant development and evolution. International Journal of Developmental Biology, 56(9): 661-674.

Hong, R.L.; Hamaguchi, L.; Busch, M.A.; Weigel, D. 2003. Regulatory Elements of the Floral Homeotic gene AGAMOUS identified by Phylogenetic Footprinting and Shadowing. The Plant Cell, 15:1296-1309.

Honma, T.; Goto, K. 2001. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature, 409(6819): 525-9.

Immink, RG; Tonaco, I.A.N.; de Folter, S.; Shchennikova, A.; van Dijk, A.D.J.; Busscher-Lange, J.; Borst, J.W.; Angenent, G.C. 2009. SEPALLATA3: The 'glue' for MADS box transcription factor complex formation. Genome Biol, 10:article R24.

Immink, R.G.H.; Kaufmann, K.; Angenent, G.C. 2010. Athe 'ABC' of MADS domain protein behaviour and interactions. Seminars in Cel & Developmental Biology, 21: 87-93.

Irish, V. 2006. Duplication, diversification, and comparative genetics of angiosperm MADS-box genes. Advances in Botanical Research: Developmental Genetics of the Flower, 44: 129-161. DOI: 10.1016/S0065-2296(06)44003-9.

Jack T, Brockman L, Meyerowitz E. 1992. The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell 68:683–697.

Jaramillo, M.A.; Kramer, E.M. 2007. The role of developmental genetics in understanding homology and morphological evolution in plants. International Journal of Plant Sciences, 168(1): 61-72. DOI: 10.1086/509078.

Kanno, A.; Nakada, M.; Akita, Y.; Hirai, M. 2007. Class B Gene Expression and the Modified ABC Model in Nongrass Monocots. The ScientificWorld Journal, 7: 268-279.

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

Kidner, K.A.; Timmermans, M.C.P. Mixing and matching pathways in leaf polarity. 2007. *Current Opinion in Plant Biology*, 10:13–20
Kirchoff B. 1983. Floral organogenesis in five genera of the Marantaceae and in Canna (Cannaceae). American Journal of Botany 70:508.

Kim S, Soltis PS, Wall K, Soltis DE. 2006. Phylogeny and domain evolution in the APETALA2-like gene family. Molecular Biology and Evolution 23:107–120.

Kirchoff B. 1988. Inflorescence and flower development in Costus scaber (Costaceae). Canadian Journal of Botany 66:339–345.

Kirchoff B. 1991. Homeosis in the flowers of the Zingiberales. American Journal of Botany 78:833.

Kirchoff B, LagomarsinoLP, NewmanWH, BartlettME, Specht CD. 2009. Early floral development of Heliconia latispatha (Heliconiaceae), a key taxon for understanding the evolution of flower development in the Zingiberales. American Journal of Botany 96:580–593.

Kishino, H., Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol., 29: 170-179.

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

Kramer, E.M.; Jaramillo, M.A.; Di Stilio, V.S. 2004. Patterns of Gene Duplication and Functional Evolution During the Diversification of the *AGAMOUS* Subfamily of MADS Box Genes in Angiosperms. Genetics, 166: 1011–1023.

Kramer, E.; Hall, J.C. 2005. Evolutionary dynamics of genes controlling floral development. Current Opinion in Plant Biology, 8:13-18.

Kress J. 1990. The phylogeny and classification of the Zingiberales. Annals of the Missouri Botanical Garden 77:698–721.

Kress, W.J.; Prince, L.M.; Hahn, W.J.; Zimmer, E.A. 2001. Unraveling the Evolutionary Radiation of the Families of the Zingiberales Using Morphological and Molecular Evidence. Systematic Biology, 50(6): 926–944.

Krizek, B.A.; Meyerowitz, E.M. 1996. Mapping the protein regions responsible for the functional specificities of the Arabidopsis MADS-domain organ-identity proteins. PNAS, 93(9): 4063-70.

Krizek, B.A.; Fletcher, J.C. 2005. Molecular mechanisms of flower development: an armchair guide. Nature Reviews Genetics, 6: 688-698.

Lakatos, I. 1970. Falsification and the Metodology of Scientific Research Programmes. In: Criticism and the growth of knowledge. I. Lakatos and A. Musgrave. (eds.) p. 91-196. Cambridge University Press: Cambridge.

Lakatos, I. 1980. The Methodology of Scientific Research Programmes: Philosophical Papers Volume 1. J. Worral & G. Currie (eds). Cambridge University Press: Cambridge.

Landis, J.B.; Barnett, L.L.; Hileman, L.C. 2012. Evolution of petaloid sepals independent of shifts in B-class MADS box gene expression. Development Genes and Evolution, 222(1): 19-28.

Leinfellner W. 1956. Die blattartig flachen Staubbla¨tter und ihre gestaltliche Beziehung zum Bautypus des Angiospermenstaubblattes. Oesterreichische Botanische Zeitschrift 103: 267–290.

Léon-Kloosterziel, K.M.; Keijzer, C.J.; Koornneef, M. 1994. A seed shape mutant of Arabidopsis that is affected by integument development. Plant Cell 6: 385-392.

Liang, H.Y; Mahadevan, L. 2011. Growth, geometry, and mechanics of a blooming lily. PNAS, 108(14): 5516-5521.

Litt, A. 2007. An evaluation of A-function: Evidence from the APETALA1 and APETALA2 gene lineages. International Journal of Plant Sciences, 168(1): 73-91. DOI: 10.1086/509662.

Litt, A.; Kramer, E.M. 2010. The ABC model and the diversification of floral organ identity. Seminars in Cell & Developmental Biology, 21: 129–137.

Lönnig, W.E. Goethe, sex, and flower genes. 1994. Plant Cell, 6:574-577.

Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, North America, 17, May 2011. Available at: http://journal.embnet.org/index.php/embnetjournal/article/view/200 .

McAbee, J.M.; Hill, T.A.; Skinner, D.J.; Izhaki, A.; Hauser, B.A.; Meister, R.J.; Reddy, G.V.; Meyerowitz, E.M.; Bowman, J.L.; Gasser, C.S. 2006. ABERRANT TESTA SHAPE encodes a KANADI family member, linking polarity determination to separation and growth of Arabidopsis ovule integuments. The Plant Journal 46: 522–531

McConnell; J.R.; Barton, K.M. Leaf polarity and meristem formation in *Arabidopsis*. 1998. *Development*, 125: 2935-2942.

Melzer, R.; Theissen, G. 2009. Reconstitution of 'floral quartets' *in vitro* involving class B and class E floral homeotic proteins. Nucleic Acids Res., 37(8): 2723–2736.

Melzer, R.; Verelst, W.; Theissen, G. 2009. The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes in vitro. Nucleic Acids Res., 37:144–157

Mendoza, L.; Alvarez-Buylla, E. 1998. Dynamics of the genetic regulatory network for Arabidopsis flower morphogenesis. Journal of Theoretical Biology, 193: 307-319.

Mendoza, L.; Thieffry, D.; Alvarez-Buylla, E. 1999. Genetic control of flower morphogenesis in Arabidopsis thaliana: a logical analysis. Bioinformatics, 15(7/8): 593-606.

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 Molecular Biology, 28: 767-784.

Mondragón-Paomino, M.; Theissen, G. 2011. Conserved differential expression of paralogous DEFICIENS- and GLOBOSA-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'. The Plant Journal, 66: 1008-1019.

Nakayama, H.; Yamaguchi, T.; Tsukaya, H. 2010. Expression patterns of AADL , A CRABS CLAW ortholog in *Asparagus asparagoides* (Asparagaceae), demonstrate a

stepwise evolution of CRC / DL subfamily of YABBY genes.

Nakayama, H.; Yamaguchi, T.; Tsukaya, H. 2012. Acquisition and Diversification of Cladodes: Leaf-Like Organs in the Genus Asparagus. The Plant Cell, 24: 929–940.

Nardmann, J., Zimmermann, R., Durantini, D., Kranz, E. and Werr, W. 2007. WOX gene phylogeny in Poaceae: a comparative approach addressing leaf and embryo development. Mol. Biol. Evol. 24: 2474–2484.

Nishiyaman, T. et al. 2003. Comparative genomics of Physcomatrella patens gametophytic transcriptome and Arabidopsis thaliana: Impications for land plant evolution. PNAS, 100(13): 8007-8012.

Nylander, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Pai, RM 1963. The floral anatomy of Canna indica L. Bulletin of the Botanical Society, College of Science, Nagpur 4:45–53.

Pekker, I., Alvarez, J.P., and Eshed, Y. 2005. Auxin response factors mediate Arabidopsis organ asymmetry via modulation of KANADI activity. Plant Cell 17: 2899–2910.

Pelaz, S.; Ditta, G.S.; Baumann, E.; Wisman, E.; Yanofsky, M.F. 2000. B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature, May 11;405(6783):200-3.

Pelaz, S.; Gustafson-Brown, C.; Kohalmi, S.E.; Crosby, W.L.; Yanofsky, M.F. 2001. APETALA1 and SEPALLATA3 interact to promote flower development. Plant Journal, 26(4):385-94.

Pelaz, S.; Tapia-López, R.; Alvarez-Buylla, E.R.; Yanofsky, M.F. 2001. Conversion of leaves into petals in Arabidopsis. Current Biology, 11:182–184.

Pond, S.L.K.; Frost, S.D.W.; Muse, S.V. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics, 21(5): 676-679.

Posé, D.; Yant, L.; Schmid, M. 2012. The end of innocence: flowering networks exlode in complexity. Current Opinion in Plant Biology, 15: 45-50.

Posluszny UM, Scott MG, Sattler R. 1980. Revisions in the technique of epi-illumination light microscopy for the study of floral and vegetative apices. Canadian Journal of Botany 54:2491–2495.

Rao VS, Donde N. 1955. The floral anatomy of Canna flaccida. Journal of the University of Bombay 24:1–10.

Rijpkema, A.; Vandenbussche, M.; Koes, R.; Heijmans, K.; Gerats, T. 2010. Variations on a theme: Changes in the floral ABCs in angiosperms. Seminars in Cell & Developmental Biology, 21: 100-107.

Roberts, A.; Trapnell, C.; Donaghey, J.; Rinn, J.L.; Pachter, L. 2011. Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biology, 12(3): R22

Sarojam, R.; Sappl, P.G.; Goldschmidt, A.; Efroni, I.; Floyd, S.K.; Eshed, Y.; Bowman J.L. 2010. Differentiating Arabidopsis Shoots from Leaves by Combined YABBY Activities. *The Plant Cell*, 22: 2113–2130.

Sattler R. 1968. A technique for the study of floral development. Canadian Journal of Botany 46:720–722.

Sawa, S., Ito, T., Shimura, Y., and Okada, K. 1999. FILAMENTOUS FLOWER controls the formation and development of Arabidopsis inflorescences and floral meristems. Plant Cell 11: 69–86.

Schichnes D, Nemson JA, Slohberg L, Ruzin S. 1999. Microwave protocols for paraffin microtechinique and in situ localization in plants. Microscopy and Microanalysis 4:491–496.

Schwartz-Sommer, Z.; Huijser, P.; Nacken, W.; Saedler, H.; Sommer, H. 1990. Genetic control of flower development by homeotic genes in *Antirrhinum majus*. Science, 250: 931-936.

Sharma, B.; Kramer, E. 2012. Sub- and neo-functionalization of APETALA3 paralogs have contributed to the evolution of novel organ identity in Aquilegia (columbine, Rnnunculales). New Phytologist, 197(3): 949-957.

Sieber, K.; Gheyselinck, J.; Gross-Hardt, R.; Laux, T.; Grossniklaus, U.; Schneitz, K. 2004. Pattern formation during early ovule development in *Arabidopsis thaliana*. Developmental Biology 273: 321-334.

Soltis, D.; Chanderbali, A.; Kim, S.; Buzgo, M.; Soltis, P.S. 2007. The ABC Model and its Applicability to Basal Angiosperms. Annals of Botany, 100: 155–163.

Sommer, H.; Beltrán, J.P.; Huijser, P.; Pape, H.; Lönnig, W. E.; Saedler, H.; Schwarz-Sommer, Z. 1990. *Deficiens*, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. EMBO J., 9: 605–613.

Swofford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (and other methods) 4.0 Beta. Sinauer Associates.

Talbert, P.; Adler, H-T.; Parks, D.W.; Comai, L. 1995. The *REVOLUTA* gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of *Arabidopsis thaliana*. *Development* 121, 2723-2735.

Takeda, S.; Iwasaki, A.; Matsumoto, N.; Uemura, T.; Tatematsu, K.; Okada, K. 2013. Physical Interaction of Floral Organs Controls Petal Morphogenesis in Arabidopsis. Plant Physiology, 161: 1242-1250.

Tang M, Li G, Chen M. 2007. The phylogeny and expression pattern of APETALA2-like genes in rice. Journal of Genetics and Genomics 34:930–938.

Theissen, G.; Melzer, R. 2007. Molecular mechanisms underlying origin and diversification of the angiosperm flower. Annals of Botany, 100(3): 603-619 DOI: 10.1093/aob/mcm143.

Tomescu, A. 2008. Megaphylls, microphylls and the evolution of leaf development. Trends in Plant Science, 14(1): 5-12.

Toriba, T., Suzaki, T., Yamaguchi, T., Ohmori, Y., Tsukaya, H. and Hirano, H.Y. 2010. Distinct regulation of adaxial–abaxial polarity in anther patterning in rice. Plant Cell 22: 1452–1462.

Tröbner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P. Lönnig, W, Saedler H., Sommer, H., and Schwarz-Sommer, Z. 1992. GLOBOSA: a homeotic gene which interacts with DEFICIENS in the control of Antirrhinum floral organogenesis. EMBO J, 11: 4693-4704

True, J.R.; Carroll, S.B. 2002. Gene co-option in physiological and morphological evolution. Annual Review of Cell and Developmental Biology, 18: 53-80.

Villanueva, J.M.; Broadhvest, J.; Hauser, A.; Meister, R.J.; Schneitz, K.; Gasser, C.S. 1999. INNER NO OUTER regulates abaxial-adaxial patterning in *Arabidopsis* ovules. Genes & Development 13: 3160-3169.

Waites, R. and Hudson, A. 1995. *phantastica*: a gene required for dorsiventrality of leaves in *Antirrhinum majus*. *Development* 121, 2143-2154.

Weigel D, Meyerowitz EM. 1994. The ABCs of floral homeotic genes—a review. Cell 78:203–209.

Winter, K-U.; Weiser, C.; Kaufman, K.; Bohne, A.; Kirchner, C.; Kanno, A.; Saedler, H.; Theissen, G. 2002. Evolution of class B floral homeotic protein: Obligate heterodimerization originated from homodimerization. Mol. Biol. Evol., 19(5): 587-596.

Wu, T.D; Nacu, S. 2010. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics, 26: 873-881

Yamada, T.; Yokota, I.; Hirayama, Y.; Imaichi, R.; Kato, M.; Gasser, C.S. 2011. Ancestral expression patterns and evolutionary diversification of YABBY genes in angiosperms. The Plant Journal, 67: 26–36

Yamaguchi, T.; Nukazuka, A.; Tsukaya, H. Leaf adaxial–abaxial polarity specification and lamina outgrowth: evolution and development. 2012. Plant and Cell Physiology, 53(7): 1180–1194.

Yamaguchi, T.; Lee, D.Y.; Miyao, A.; Hirochika, H.; An, G.; Hirano, H.Y. 2006. Functional diversification of the two C-class MADS box genes *OSMADS3* and *OSMADS58* in *Oryza sativa*. Plant Cell, 18(1): 15-28.

Yang, Y.; Fanning, L.; Jack, T. 2003. The K domain mediates heterodimerization of the Arabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. Plant J., 33:47–59.

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

Yang, Z. 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution, 24: 1586-1591.

Yanofsky, M.F.; Ma, H.; Bowman, J.L.; Drews, G.N.; Feldman, K.A.; Meyerowitz, E.M. 1990. The protein encoded by the Arabidopsis homeotic gene AGAMOUS resembles transcription factors. Nature, 346(6279): 35-39.

Zahn, L.M.; Leebens-Mack, J.H.; Arrington, J.M.; Hu, Y.; Landherr, L.L.; Soltis, D.E.; dePamphilis, C.W.; Becker, A.; Theissen, G.; Ma, H. 2006. Conservation and divergence in the AGAMOUS subfamily of MADS-box genes: evidence of independent sub- and neofunctionalization events. Evol. Dev. 8(1): 30-45.