Université de Montréal

Comparative analysis of corolla shape transitions in the sister genera *Gesneria* and *Rhytidophyllum* (Gesneriaceae)

Par
Carolina Vergolino Martini

Département de Sciences Biologiques
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de
Maîtrise en Sciences Biologiques.

Décembre 2023

Université de Montréal

Département de Sciences biologiques / Institut de Recherche en Biologie Végétale,
Faculté des Arts et des Sciences

Ce mémoire intitulé

Comparative analysis of corolla shape transitions in the sister genera *Gesneria* and *Rhytidophyllum* (Gesneriaceae)

Présenté par
Carolina Vergolino Martini

A été évalué par un jury composé des personnes suivantes

Anne Bruneau
Président-rapporteur

Simon Joly
Directeur de recherche

David Morse
Membre du jury

Abstract

Convergence, the independent acquisition of similar phenotypes, is an important aspect of diversity that can provide valuable insights about the nature of evolutionary change. In plants, pollination syndromes - combinations of floral traits adapted to their pollinators - make good examples of convergence occurring on flowers. We used a comprehensive approach that includes cell morphology and transcriptomics to analyze the floral shape convergence of two pollination syndromes found in the sister genera *Gesneria* and *Rhytidophyllum* (Gesneriaceae), an Antillean group that contains approximately 81 species with different morphologies and pollination strategies varying in their degree of ecological specialization. Flower shape has already been found to play an important role in the evolution of this group, which shows many transitions between pollination strategies. We tested convergence in the corolla cell shapes and in gene expression for the pollination syndromes using (1) cell measurement statistical analysis (Phylogenetic Mixed Model) of mature petals and (2) a comparative transcriptomic approach that combined differential gene expression (DESEq2) and co-expression analysis (WGCNA) in genes expressed in specific regions of the petals. All analyses took the phylogenetic relationships of the species into account. We found convergent cellular anisotropy occurring in the distal regions of the petals within species of the same syndrome (form). We also found greater similarity in gene expression patterns occurring among species of the same syndromes than between more closely related species and produced a list of 203 genes potentially associated with convergent flower forms. The floral morphological convergence observed in the pollination syndromes of the investigated species is paralleled both at the cellular and expression levels. The results shown here amplify the background information of the Gesneriaceae family for future studies of convergence and floral form in the group.

Keywords: Convergence, transcriptomics, flower, petal, shape, Gesneriaceae, RNA-seq, pollination syndrome, morphology, development.

Résumé

La convergence, soit l'acquisition indépendante de phénotypes similaires, est un aspect intéressant de la diversité qui peut fournir des informations importantes sur la nature du changement évolutif. Dans les systèmes végétaux, les syndromes de pollinisation – combinaisons de traits floraux adaptés à leurs pollinisateurs – constituent de bons exemples de convergence se produisant sur les fleurs. Nous avons utilisé une approche globale incluant la morphologie cellulaire et la transcriptomique pour analyser la convergence de formes florales de deux syndromes de pollinisation trouvés dans les genres frères non *Gesneria* et *Rhytidophyllum* (Gesneriaceae), un groupe antillais qui contient environ 81 espèces avec différentes morphologies et stratégies de pollinisation variables dans leur degré de spécialisation écologique. Il a déjà été démontré que la forme des fleurs joue un rôle important dans l'évolution de ce groupe, qui présente de nombreuses transitions entre les stratégies de pollinisation. Nous avons testé la présence de convergence dans les forms de cellules de la corolle et dans l'expression des gènes de la corolle en utilisant (1) une analyse pour mesurer la forme des cellules de pétales matures à l'aide d'un modèle phylogénétique mixte et (2) une approche transcriptomique comparative combinant l'expression différentielle des gènes (DESEq2) et l'analyse de co-expression (WGCNA) de gènes exprimés dans certaines regions précises des pétales. Toutes les analyses ont pris en compte les relations phylogénétiques entre les espèces. Nous avons trouvé une anisotropie cellulaire convergente se produisant dans les régions distales des pétales au sein des espèces du même syndrome (forme). Nous avons également constaté une plus grande similarité dans les modèles d'expression génique entre les espèces d'un même syndrome qu'entre les espèces apparentées et avons produit une liste de 203 gènes potentiellement associés aux formes de fleurs convergentes. La convergence morphologique florale observée dans les syndromes de pollinisation des espèces étudiées se retrouve tant au niveau cellulaire qu'au niveau de l'expression. Les résultats présentés ici amplifient les informations de base sur la famille des Gesneriaceae pour les études futures sur la convergence et la forme florale dans le groupe.

Mots-clés: Convergence, transcriptomique, fleur, pétale, forme, Gesneriaceae, séquençage d'ARN, syndrome de pollinisation, morphologie, développement.

Table of contents

List of tables

List of figures

List of abbreviations

dds: DESeq2 data set

DIC: Deviance information criterion

DE: Differential gene expression

FDR: False discovery rate

GO: Gene ontology

MCMC: Markov chain Monte Carlo

mRNA: messenger ribonucleic acid

ORF: Open reading frame

PMM: Phylogenetic mixed model

PSRF: Potential scale reduction factor

PCA: Principal component analysis

QTL: Quantitative trait loci

RNA: Ribonucleic acid

RNA-seq: Ribonucleic acid sequencing

TF: Transcription factor

VST: Variance stabilizing transformation

WGCNA: Weighted gene co-expression network analysis

Acknowledgments

Firstly, I want to thank my research director, Simon Joly, for providing me with this enriching research experience. Thank you for all the advice, suggestions and for introducing me to the fascinating domain of phylogenetics and plant diversity, as well as to bioinformatics, the path I intend to follow professionally.

I also want to thank my partner, Alexandre Martini, for his friendship and support throughout this academic journey. And my son, Bruno, who always brings a smile to my face.

I thank the Montreal Botanical Garden greenhouses staff Janique Perrault and Josiane Boucher for their kind assistance with sample collection, and the members of Genome Québec Innovation Center for help with the sequencing.

Finally, I thank the Joly lab members (Marion Leménager, Jérôme Burkiewicz) and Céline Devaux for helpful discussions and assistance, Martin Lefrançois for technical support in the extraction laboratories, Anas Elkhettabi for helping with the cell measurements, and Evangelini Ballerini for her insightful inputs for methodology.

1. Introduction

1.1 Flower diversity

The astonishing floral diversity of angiosperms is an intriguing and much investigated topic in evolutionary biology (Stebbins, 1970; Dellinger, 2020). Efforts to elucidate the emergence, radiation and diversification of flowers gained momentum with the emergence of molecular biology techniques that have produced large-scale sequencing data, enabling phylogenetic studies to be carried out in broad collaborations (Soltis et al., 2011). Because floral diversity is an interesting subject for understanding the evolutionary process, model species have been widely used to reveal the genetics and molecular mechanisms that determine floral traits, greatly advancing the field in the last 30 years. This progress continues to pave the way for further efforts to understand how evolution produces diversity in nature, but there are still challenges to overcome. For example, there is a need for expanding the knowledge obtained in model species to non-model ones before we can have a broader and more complete understanding of the ecological and molecular processes involved in the evolution of flowers.

Flower diversity comes from the ability of angiosperms to modify genes, gene expression and genetic pathways to produce flowers with different scents, colors, symmetry, shapes, sizes and textures, to obtain optimal reproduction. This ability possibly explains why angiosperms represent the vast majority of extant plant species on earth, even though they originated relatively recently. The achievement of reproductive success is one of the main factors that can lead to transitions in mating systems and speciation. Since there are numerous factors that influence reproduction, there are also plenty of structural variations in flower reproductive systems in response to these diverse demands (Barrett, 2010). Due to their motionless state, plants must overcome challenges when it comes to producing genetic diversity through cross-pollination. Therefore, the evolution of structural variations such as the self-fertilizing systems (autogamy), and the spatial (herkogamy) and temporal (dichogamy) separation of sexual structures, have occurred and are attributed mainly as an adaptation to unreliable vectors for cross-pollination. Moreover, gender variations and the evolution of separate sexes (dioecy) from cosexuality, and from animal to wind pollination (anemophily) also produced a plethora of structural variations that contributed to diversity.

1.2 Flower convergence

One aspect of floral diversity that receives a lot of attention is convergence. Convergence is the independent acquisition of similar phenotypes, commonly explained as the evolution of similar morphological traits through natural selection, when distant species adapt to the same environment (Wake, 1991). It is interesting to study convergence because it is an evolutionary process that generates phenotypic similarities instead of diversity, which can be very informative about the genetics of new adaptations when coupled with large-scale sequencing and advances in phylogenetics, a niche that boosted the understanding of convergent evolution (Stern, 2013). Convergence is widespread in nature and found among closely related species – sometimes called parallelism (Arendt and Reznick, 2008) - and among highly divergent ones. Classic examples include echolocation in bats and dolphins (Liu et al., 2010; Parker et al., 2013) and the emergence of wings in insects, birds and bats. But convergence might range from these broad functions to very specific ones, such as when similar amino acid substitutions lead to convergent phenotypes (Castoe et al., 2009; Liu et al., 2010). Nevertheless, these phenomena clearly provide outstanding study systems and its genetic causes have been the subject of much current research (Stern, 2013).

In plant systems, pollination syndromes make good examples of convergence occurring on flowers. They are defined as combinations of floral traits adapted to the main pollinator, which arise independently among species (Fenster et al., 2004; Schiestl and Johnson, 2013; Rosas-Guerrero et al., 2014). Indeed, species from different clades often have the same pollination syndrome when pollinated by the same functional pollinators. For example, red tubular flowers are associated with hummingbird pollination. Although paradoxical, pollination syndromes, which represent floral convergence, are extremely useful for understanding the macroevolutionary mechanisms that lead to floral diversification. It allows to easily study transitions between pollination strategies at broad taxonomic scales and link these transitions to external factors. Transitions in floral traits resulting in speciation have been widely associated with pollinator shifts between pollination syndromes (Schemske and Bradshaw, 1999; Whittal and Hodges, 2007).

1.3 ABC model and the genetic bases of flower development

The genetics of flower development has been extensively investigated to unravel the biological processes that produce the varied morphology of flowers and today we have a good idea

of the main genes involved in the development of flowers. The initial establishment of the four floral organs (sepals, petals, stamens and carpels) in early flower development is defined in a conserved combinatorial framework explained by the ABC model (Coen and Meyerowitz, 1991). The model, derived from studies of mutants in *Arabidopsis* and *Antirrhinum,* explains how organ identities and their positions are initially determined in floral primordia, a process that likely occurs similarly in most angiosperms (Causier *et al*., 2010). Basically, the activity of classes of homeotic genes alone or in combination, generate the development of the different organs; A-class genes alone promote the whorl of sepals, A+B genes promote the whorl of petals, C-class genes alone promote the carpels and C+B the stamens (Figure 1).

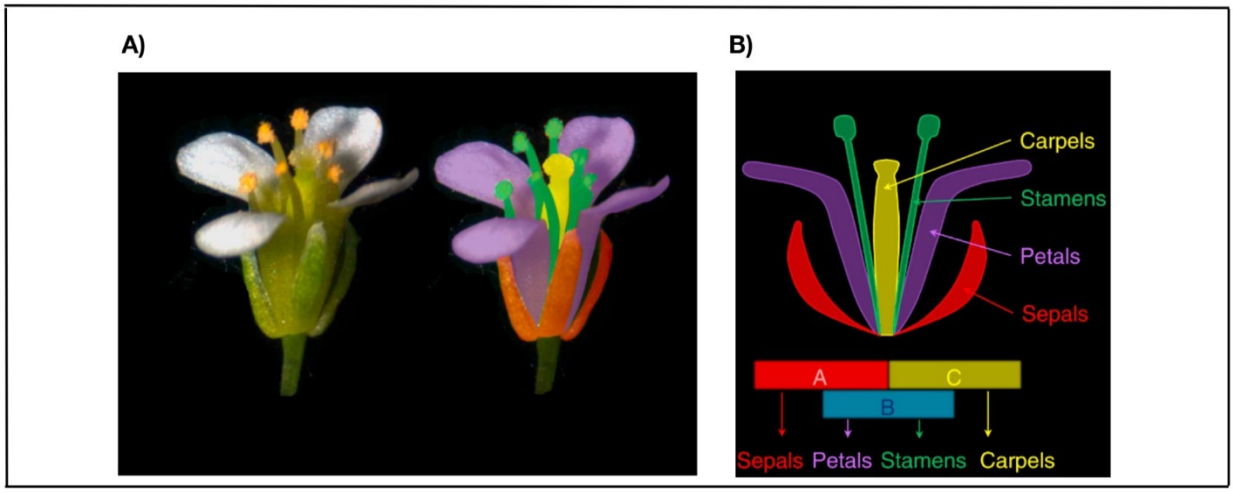

Figure 1. The ABC model. A) *Arabidopsis thaliana* in natural colors (left) and color coded (right), **B)** Illustration of the model, showing A-class genes alone promoting the sepals; A+B, the petals; C-class genes alone promoting the carpels, and B+C genes promoting the stamens. Figure modified from Irish, 2017.

Further studies in *Petunia hybrida* and *Arabidopsis thaliana* incorporated other classes of homeotic genes, D and E, to the model (Angenent and Colombo, 1996; Pelaz et al., 2000; Ditta et al., 2004). In this context, the ABC model was extended to ABCDE, whereby D-class genes are involved in the specification of ovule identity and E-class genes would mainly have functions redundant to A-class genes. The homeotic genes of the ABCDE model are mostly part of the MADS-box family that code for transcription factors (TFs). In *Arabidopsis*, A-class genes are the *APETALA1 (AP1)* and *APETALA2 (AP2)*, B-class genes are the *APETALA3 (AP3)* and

PISTILLATA (PI), C-class is the *AGAMOUS (AG)*, D-class is *SEEDSTICK (STK)*, and E-class by *SEPALLATA (SEP)*.

Given that this basic pattern is conserved in most flowering plants, the diversity we observe in flowers is largely produced in subsequent developmental stages, when taxa activate specific genetic pathways that lead to major phenotypic differences in each of the four floral organs. Such differences have been often shown to be dictated by TFs that modulate floral development at the expression level. The MYB gene family is an example of TFs that are frequently involved in the determination of floral pigmentation, scent and cell texture (Klahre *et al*., 2011; Yuan *et al.*, 2013). The gene family has expanded throughout evolution, going from a few members in the genomes of green algae to several hundred in the genomes of current flowering plants (Feller *et al*., 2011). This supports the idea that TFs play a key role in the production and expansion of floral diversity.

1.4. Flower shape

The corolla, the whorl of petals, contributes to the diversity of flowers in several important ways. It can vary in color, size, petal number, symmetry, texture and shape. Therefore, it can play an important role in adaptation to pollination pressures, in both animal and self-driven pollination strategies. In animal-pollinated species, this relevance comes from the physical potential of the corolla to attract or deter pollinators and to ensure proper pollen deposition on the pollinator and thus proper pollination. Not surprisingly, petal variations have been extensively studied. Petal shape, for instance, has been associated with restrictions of the entry that select specific pollinators as seen in the Orchidaceae. It is also important for guiding the approach of pollinators to ensure that they enter in contact with the reproductive organs (Muchhala, 2007), and has been associated with pollinator shifts in monkeyflowers and columbines (Schemske and Bradshaw, 1999; Whittall and Hodges, 2007).

Floral shape development involves a combination of growth patterns with variations in cell division, expansion and directional elongation, as well as variations in the rate, duration and location of these cellular processes (Coen, 2004; Puzey et al., 2012; Yant et al., 2015; Kierzkowski and Routier-Kierzkowska, 2019); all tools present in plants to enable the production of their final corolla shape. Even variability in the shape and size of individual cells have been shown to be used to correct noise for the fine-tuning of optimal petal shapes (Hong *et al.*, 2016). However, the

complexity of this trait makes it challenging to be investigated, especially outside model species, where studies are scarce. Information on the mechanisms of shape development collected from model systems provide an essential background for studies in non-model species, as there can be overlapping functions across taxa. However, the function of genes can also be quite distinct or even opposite in different species. For instance, the transcription factor *CINCINNATA* has been shown to promote cell proliferation in the petal of *Antirrhinum* while its *Arabidopsis* homologue, *TCP4,* inhibits growth (Crawford *et al*., 2004; Nag *et al*., 2009). These disparities point to the need for more studies on the shape of petals outside the model species.

Studies using QTL mapping have identified genes or genomic regions associated with flower shape determination (Bradshaw *et al.*, 1998; Stuurman *et al*., 2004; Galliot *et al*., 2006; Alexandre *et al.*, 2015), and some were able to link the trait to its importance in the natural environment by demonstrating that phenotypic changes in flower shape are associated with pollinator shifts (Schemske and Bradshaw, 1999). In general, however, QTL studies have limitations for the investigation of traits such as the shape of flowers because they are polygenic and determined by networks of genes that might lie within several different QTLs of small to moderate effect (Fishman *et al.*, 2002; Wessinger *et al.*, 2014; Feng *et al*., 2018). Because of this complexity, some studies choose to investigate the mechanisms that generate the shape of specific flower regions considered to be functionally important, such as the wedge-shaped fold of snapdragon flowers (Rebocho *et al.*, 2017). Live confocal imaging and new software developments have also allowed the analysis of shape and morphogenesis using cell lineage tracking, that provides information on growth rate, duration and direction (Kierzkowski *et al.,* 2019; Rambaud-Lavigne and Hay, 2020). But this approach is of limited use for most non-model species that do not have established protocols for plant transformation and transgenesis.

An interesting alternative for non-model species is the investigation of gene expression, as different phenotypes might be largely produced at the transcriptional regulatory level (Brawand *et al*., 2011; Romero et al., 2012; Uebbing *et al*., 2016). Non-model groups without previous information on gene expression, mutants or a sequenced genome can be sequenced to obtain *de novo* transcriptome assemblies that can be further analysed using reference genomes and public databases. Such gene expression approaches have revealed interesting genes associated with petal shapes in different plant systems. For instance, studies have compared and analysed transcriptional activity across developmental stages (Min et al., 2019), across species or in a combination of both

(Roberts and Roalson, 2017, 2020; Ballerini *et al.*, 2019) to find genes of interest involved in floral shape.

1.5 *Gesneria* and *Rhytidophyllum* (Gesneriaceae)

The non-model sister genera *Rhytidophyllum* and *Gesneria* (Gesneriaceae) make an interesting group to study petal shape diversity. The Antillean group contains approximately 81 species (Acevedo-Rodríguez and Strong, 2012), that radiated around 10 Ma (Roberts and Roalson, 2016) with different morphologies and pollination strategies varying in their degree of ecological specialization. Flower shape has already been found to play an important role in the evolution of this group, which shows many transitions between pollination strategies (Martén-Rodríguez *et al.*, 2010; Joly *et al.,* 2018).

The two genera have species with three main types of pollination strategies: (1) specialists pollinated by hummingbirds, (2) specialists pollinated by bats and (3) generalists pollinated by hummingbirds, bats and occasional insects (Martén-Rodríguez et al., 2009). Hummingbird specialists present elongated, tubular corollas with narrow opening and diurnal production of nectar (Figure 2, [A]); bat specialists present bell-shaped (campanulate) corollas with a wide opening and nocturnal production of nectar (Figure 2, [C]); and generalists present bell-shaped, wide-open corollas, with a basal constriction above the nectar chamber (subcampanulate) and produce nectar day and night (Figure 2, [B]). Notably, the basal constriction of the latter was shown to be the most important feature to distinguish the two types of bell-shaped strategists in a multidimensional scaling analysis that included color, corolla curvature, nectar concentration and timing of anthesis (Martén-Rodríguez *et al.*, 2009), demonstrating the importance of this specific region for the generalists of this group.

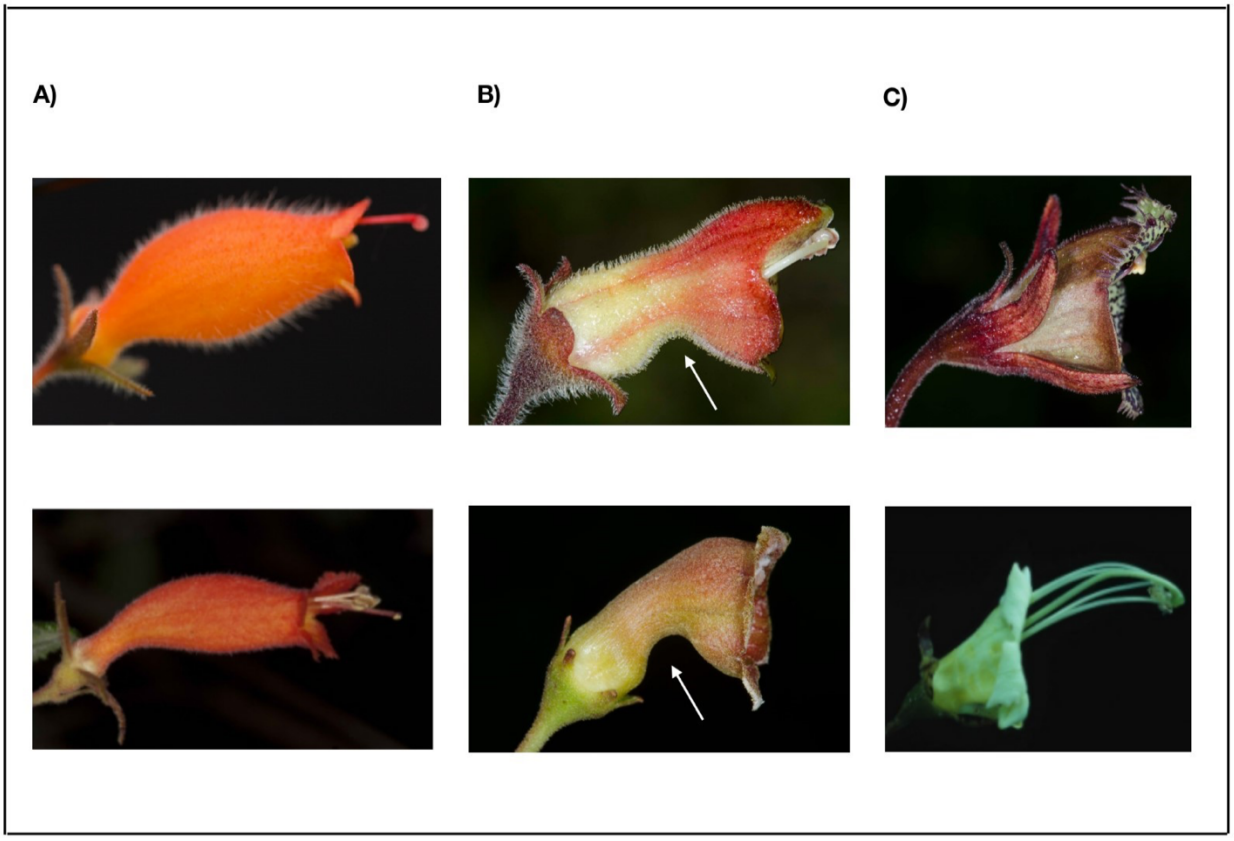

Figure 2. Examples of Gesneriaceae flower shapes. A) Tubular corollas of hummingbird specialists (*R. Rupincola*, top; *R. earlei*, bottom), **B)** Bell-shaped corollas (subcampanulate) of generalists with arrows pointing to the basal constriction (*R. Exsertum,* top; *G. bicolor*, bottom), **C)** Bell-shaped corollas of bat specialists (campanulate) (*G. fruticosa*, top; *G. pedunculosa,* bottom). Photos by J. Clark. Flower sizes on photos do not reflect their actual sizes.

Typical hummingbird syndromes have elongated corollas with narrow openings. In general, the elongation is considered necessary to increase the distance between the nectar and the pollen, which ensures pollen deposition on the head of the long-beaked bird. The narrow opening might also deter unwanted pollinators such as insects or bats, as well as forcing the bird to position itself in a manner that allows for ideal fit. Like elongation, corolla width has also been shown to be an adaptive trade-off in studies investigating hummingbird and bat pollination, whereby pollination by one class of animal imposed the loss of the other **(**Muchhala, 2007**),** demonstrating again the high level of specificity in the diversity and evolution of this trait.

Given that the generalist species of *Rhytidophyllum* and *Gesneria* (pollinated by bats and hummingbirds) have short corollas with wide openings - which would imply non-effective ornithophilous pollination - the basal constriction has been proposed as a feature that forces

hummingbirds to move upwards while approaching the flower to reach the nectar, thus touching the anthers with their heads (Martén-Rodríguez et al., 2009). A study carried out on *Gesneria* species showed that while both campanulate (*G. pendunculosa*) and subcampanulate (*G. viridiflora*) flowers were visited by hummingbirds, only the subcampanulates had their reproductive organs contacted during visits (Fenster and Martén-Rodríguez, 2007). The basal constriction of bell-shaped flowers in generalists could therefore be an ideal compromise to have legitimate pollination by hummingbirds and bats. In another Gesneriaceae genus, *Drymonia,* grooves in the corolla have been hypothesized to guide hummingbird bills into the nectaries of flowers to prevent damages to the ovary (Cronk and Ojeda, 2008) and to facilitate hummingbird pollination (Clark *et al*., 2015). Interestingly, the shapes of the Gesnerieae generalists were shown to present lower disparity at the interspecific level compared to that of hummingbird specialists, despite opposite expectations, given their generalist state (Joly *et al.,* 2018). This further suggests that the corolla shape of generalists might have converged to an optimal form, and the constriction might be an important player in this configuration.

The genetic bases of corolla morphology on species of the group have been previously investigated. Corolla shape quantitative trait loci (QTLs) of moderate to small effect were found in a study that used *R. auriculatum* (generalist) and *R. rupincola* (hummingbird specialist) F2 hybrids (Alexandre *et al.*, 2015). To further investigate the genes responsible for this variation, 23 genes thought to likely play a role in corolla shape variation were genotyped for the same hybrid population and this resulted in 3 genes that were strongly correlated to corolla shape variation: *RADIALIS, GLOBOSA* and *JAGGED* (Poulin *et al.*, 2022). Although the role of these genes in determining corolla shape awaits further confirmation, *JAGGED* is of particular interest as it co-localizes with a shape QTL associated with the basal constriction and is involved in growth anisotropy, coordination between cell cycle and cell size, and petal growth and shape determination in *Arabidopsis* (Sauret-Gueto et al., 2013; Schiessl et al., 2014). Despite these interesting results, many other genes that could explain the variation in the polygenic corolla shape of this group remain to be found.

2. Objectives and hypotheses

The overall objective of this thesis was to test whether the convergence of floral form presented in the pollination syndromes of generalists (subcampanulate) and hummingbird specialists (tubular) of the group, is reflected in the shapes of the corolla cells and in corolla gene expression levels. This objective focuses on a multifactorial analysis of the development of the form, with the intention of maximizing the power of inferences of the study. The thesis had the following hypotheses:

1. As the group presents relatively recent transitions between the two syndromes (forms) (Martén-Rodriguez et al., 2010; Joly et al., 2018), and knowing that phenotypic changes can occur more easily at the level of gene regulation than in coding regions (Stern and Orgogozo, 2008; Stern, 2013), the convergent forms of the group could have evolved through convergent gene expression changes. If this was the case, gene expression patterns would be similar for species that share similar floral forms.

2. Due to the spatiotemporal specificity in which these transcription factors act during floral development, it would be possible to obtain significant differences in gene expression patterns between different zones of the corolla.

3. Article

Comparative analysis of corolla shape transitions in the sister genera *Gesneria* and *Rhytidophyllum* (Gesneriaceae)

Carolina Vergolino and Simon Joly

In preparation for future publication

3.1 Introduction

The floral diversity of angiosperms and its emergence is a fascinating and much investigated topic in evolutionary biology (Stebbins, 1970; Dellinger, 2020). Called an abominable mystery by Darwin in 1879 in a letter to Dr. Joseph Hooker, a botanist friend, many aspects of the sudden evolution of flowers have since been clarified. Floral diversity comes from the ability of angiosperms to modify genes, gene expression and genetic pathways to produce flowers with different colors, scents, nectar composition, textures, shapes and sizes to achieve optimal reproduction (Specht and Bartlett, 2009). But the topic of diversity still presents intriguing questions, such as the ability of evolution to produce similarity, just as it produces diversity. Flower convergence is frequently observed in nature, occurring when morphologically similar flowers arise in phylogenetically distant taxa (Rosas-Guerrero et al., 2014; Lagomarsino et al., 2017; Dellinger, 2020; Bilbao et al., 2021). The independent acquisition of similar phenotypes is a remarkable aspect of evolution that has generated much taxonomic misclassification when classifications were based uniquely on similar morphologies (Oyston et al., 2022). The topic remains interesting because it addresses the predictability of evolutionary change. Knowing the mechanisms behind the occurrence of convergence can greatly deepen our understanding of how changes are generated in the evolutionary process.

Convergence is often explained as adaptation through natural selection, whereby distant species need to adapt to the same environmental pressures, favouring the emergence of similar morphological traits. The theme maintains central and frequently investigated questions, such as whether it is generated by similar or distinct molecular processes, and if it is the result of adaptation or developmental constraints (Christin et al., 2010; Losos, 2011). Studies have been conducted in

various systems, demonstrating that the mechanisms leading to the independent generation of similar traits might range from strictly similar molecular processes such as mutations on the same genes, to developmental constraints biasing traits into a given direction (Donoghue and Ree, 2000; Fernald, 2006; Protas et al., 2006; Arendt and Reznick, 2008; Liu et al., 2010; Barrett et al., 2019).

In plant science, the concept of pollination syndromes is a good example of the evolutionary convergence occurring in flowers. It refers to combinations of floral traits adapted to the main pollinator, which arise independently between species (Fenster et al., 2004; Schiestl and Johnson, 2013; Rosas-Guerrero et al., 2014). In each of these syndromes, the corolla, in particular, often shows a high level of convergence between unrelated species. The whorl of petals can vary in size, color, smell, texture, symmetry and shape, therefore having a wide range of variables to be used in the optimization of flower reproduction. Not surprisingly, petal variations have been extensively studied in model systems like *Petunia, Mimulus, Aquilegia* and *Antirrhinum* (Bradshaw and Schemske, 1999; Stuurman et al., 2004; Crawford et al., 2004; Whittal and Hodges, 2007). Petal shape, in particular, is known to be closely related to plant fitness (Galen, 1989; Muchhala, 2007; Wester and Bockhoff, 2007) and has been linked to pollinator shifts in numerous plant clades (Schemske and Bradshaw, 1999; Whittal and Hodges, 2007). However, the study of shape can be complex. The dynamics of gene expression in flowers undergoes constant changes throughout development (Vincent and Coen, 2004; Puzey et al., 2012; Yant et al., 2015), affecting the shape of corollas via variations in cell division, expansion, and directional elongation, as well as variations in the rate, duration, and location of these cellular processes. Even variability in individual cell shape and size has been shown to correct for noise to fine-tune optimal organ shapes (Hong et al., 2016).

Molecular genetic studies of model species have provided a better understanding of flower development. We know about the determination of petal symmetry (Luo et al., 1996; Almeida et al., 1997), elongation (Puzey et al., 2012; Rebocho et al., 2017; Conway et al., 2021; Edwards et al., 2022), texture (Whitney et al., 2009; Yuan et al., 2013) and color (Hoballah et al., 2007; Yuan et al., 2013). Likewise, some studies on the determination of corolla shapes have collected information on transcription factors (TFs) that provide an essential background for the study of non-model species, in which overlapping functions have been shown among taxa. However, the function of TFs can also be quite distinct or even opposite between species; for example, the *CINCINNATA* transcription factor has been shown to promote cell proliferation in the *Antirrhinum*

petal, while its Arabidopsis homologue, *TCP4*, inhibits growth (Crawford et al., 2004; Nag et al., 2009). These disparities point to the need for more studies on the mechanisms involved in petal shape determination outside model species.

Studies using quantitative trait loci (QTL) mapping have already identified genes and genomic regions associated with flower shape determination (Bradshaw et al., 1998; Stuurman et al, 2004; Galliot et al., 2006; Alexandre et al., 2015), and some were able to link the trait to its importance in the natural environment, demonstrating that phenotypic changes in flower shape can be associated with shifts in pollinators (Schemske and Bradshaw, 1999). However, in general, QTL studies have limitations in the investigation of flower shape, as shape is determined by networks of genes that can lie within several different QTLs of small to moderate effect (Fishman et al., 2002; Wessinger et al., 2014; Feng et al., 2019). An interesting alternative for studies of flower shape in non-model species is the investigation of gene expression, as different flower phenotypes can be largely produced at the transcriptional regulatory level (Brawand et al., 2011; Romero et al., 2012; Uebbing et al. al., 2016). This gene expression approach revealed interesting genes associated with petal shapes in different plant systems. For example, studies have compared and analyzed transcriptional activity across developmental stages (Min et al., 2019), across species, or in a combination of both (Roberts and Roalson, 2017, 2019; Ballerini et al., 2019; Edwards et al., 2022) to find candidate genes of interest. However, few studies investigate the extent to which convergence of flower shape is paralleled by a similar level of convergence at the gene expression level.

In an effort to expand knowledge about the cellular and molecular aspects that may contribute to the evolution of convergent flower shapes, this study aims to test to what extent the morphological convergence observed in pollination syndromes arising independently are due to similarities at the level of cellular shape and at the gene expression level. We use species from the non-model sister genera *Gesneria* and *Rhytidophyllum* (Gesneriaceae), an Antillean group that has approximately 81 species (Acevedo-Rodríguez and Strong, 2012) and radiated around 10 Ma (Roberts and Roalson, 2016) with different morphologies and pollination strategies. The two genera have species with three main types of pollination syndromes: (1) specialists pollinated by hummingbirds, (2) specialists pollinated by bats and (3) generalists pollinated by hummingbirds, bats and occasional insects (Martén-Rodríguez et al., 2009). Hummingbird specialists present elongated, tubular corollas with a narrow opening and diurnal activity; bat specialists present bell-

shaped (campanulate) corollas with a wide opening and nocturnal activity; and generalists present bell-shaped, wide-open corollas, with a basal constriction above the nectar chamber (subcampanulate) and day/night activities. Notably, the basal constriction of the latter proved to be the most important feature to distinguish the two types of bell-shaped strategists in a multidimensional scale analysis that included color, corolla curvature, nectar concentration, and anthesis time (Martén-Rodriguez et al., 2009), demonstrating the importance of this specific aspect of the form for the generalist syndrome. Studies using corolla shape analysis have shown multiple evolutionary origins for the three pollination syndromes in the group, and specifically for the generalized state from hummingbird specialist ancestors (Martén-Rodriguez et al., 2010; Joly et al. al., 2018), confirming the occurrence of floral convergence in the group. The multiple origins of pollination syndromes make this group a great candidate to test if the convergent evolution of corolla shape is paralleled at the levels of cellular shape and gene expression.

We measured the cellular shapes and quantified the gene expression of convergent forms of several species of generalists and hummingbird specialists. We then tested if cell morphologies and gene expression correlate with corolla shapes in the group and found convergence occurring both at the cellular and gene expression levels.

3.2 Materials and methods

3.2.1 Plant material

Petal tissue samples for both cell measurement and RNA-seq analysis were obtained from mature plants in the green-house of the Montreal Botanical Garden, in Montreal, Canada (Appendix, Table 1). The species are grown in 25° Celsius and 45% humidity. Cell measurements were done on the mature petals of 11 species, belonging to the genera *Gesneria* and *Rhytidophyllum*: *Gesneria acaulis, Gesneria cuneifolia, Gesneria quisqueyana, Gesneria ventricosa, Gesneria bicolor, Rhytidophyllum auriculatum, Rhytidophyllum rupincola, Rhytidophyllum exsertum, Rhytidophyllum intermedium, Rhytidophyllum tomentosum, Rhytidophyllum vernicosum* (Figure 3, [B]).

RNA-seq was done on three generalists (*Gesneria bicolor, Rhytidophyllum auriculatum and Rhytidophyllum exsertum)* and three hummingbird specialists (*Gesneria acaulis, Gesneria*

ventricosa and Rhytidophyllum rupincola) (Figure 3, [C]). Species for gene expression analysis were chosen based on the availability of fresh material, the possibility of having genetically distinct individuals, and selecting species that were evolutionary distant and that most likely had distinct origins of the pollination strategies (Figure 3, [A]). *Rhytidophyllum exertum and Gesneria acaulis* were sampled from three genetically different individuals, while *Gesneria ventricosa, Gesneria bicolor, Rhytidophyllum auriculatum* and *Rhytidophyllum rupincola* were sampled from three plants obtained by vegetative propagation (clones).

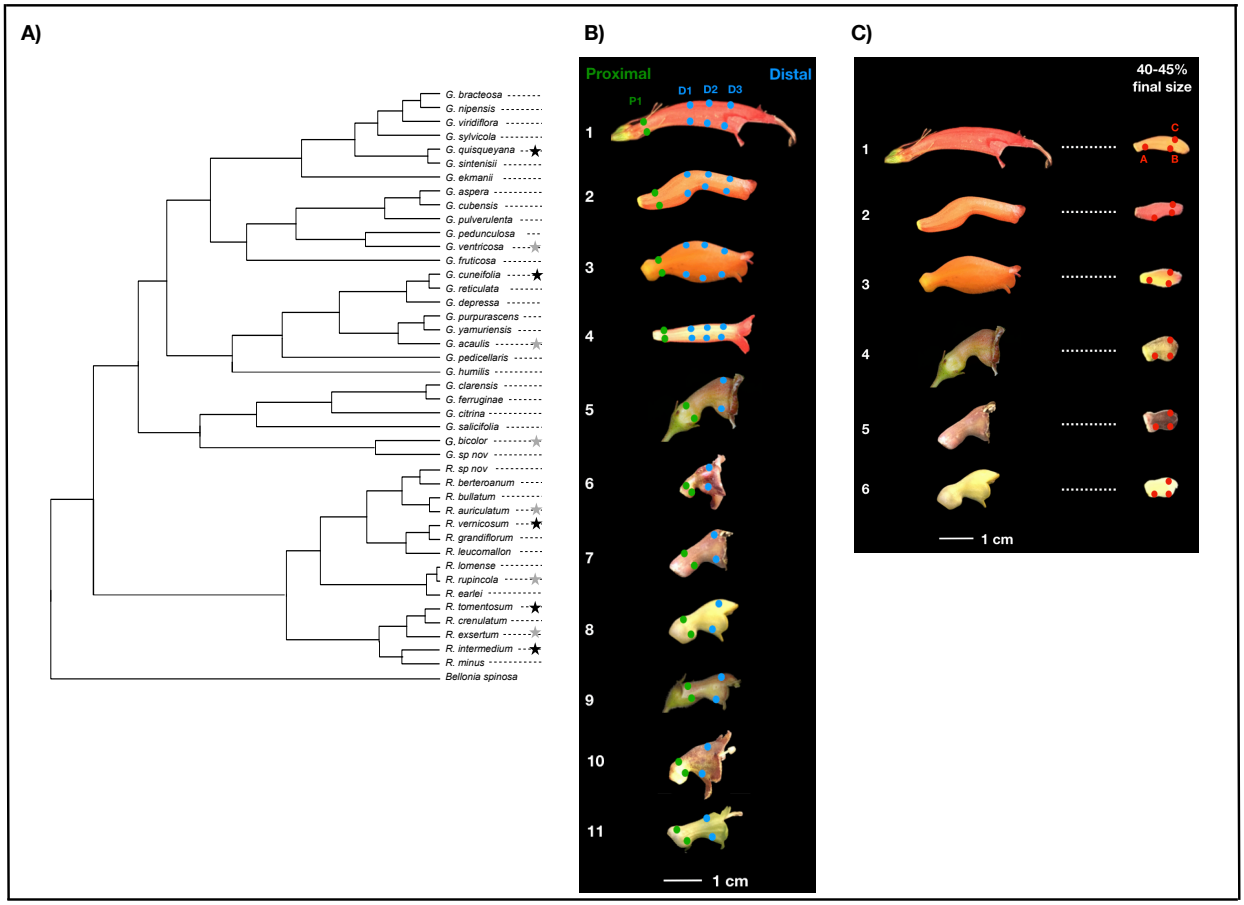

Figure 3. Species selected for the study and their sampling scheme. A) Phylogeny of *Gesneria* (top) and *Rhytidophyllum* (bottom) modified from Joly *et al.* (2018), showing the species selected for the cell measurement analysis (black stars) and the species selected for both the cell measurement and RNA-seq analysis (grey stars). **B)** Positions of the regions selected for the imaging of cell measurement analysis at proximal/distal petals (1. *G. ventricosa*, 2. *G. acaulis*, 3. *R. rupincola*, 4. *G. cuneifolia*, 5. *G. bicolor*, 6. *G. quisqueyana*, 7 *R. exsertum*, 8. *R. auriculatum, 9. R. tomentosum, 10. R. vernicosum, 11. R. intermedium*). **C)** Species selected for RNA-seq analysis and their buds between 40-45% of final flower size (1. *G. ventricosa*, 2. *G. acaulis*, 3. *R. rupincola* , 4. *G. bicolor*, 5. *R. exsertum*, 6. *R. auriculatum*), showing the three zones (A, B, C) sequenced in red.

Samples for this study were collected with the intent to capture the gene expression mostly associated with the cell differentiation, expansion and anisotropic growth of late developmental stages of the corolla - which can be highly determinant to form - while eliminating the noise that could be caused by the machinery present in early phases of corolla development, when boundary formation and establishment of organ identities take place. Therefore, samples were collected at a specific developmental stage (40-45% of final petal size)**,** which represents the beginning of the elongation phase, assessed through the observation of the development of each species. Between 2 to 3 months, buds were photographed and measured at the greenhouse every two days from the most initial phase (when buds become visible) until anthesis. Then the total length of development was calculated and compared between the species.

3.2.2 Cell measurement

Petals were fixed in 70% ethanol for 72h at 4°C. They were then separated into dorsal and ventral sections and mounted onto microscope slides to be imaged on a Axio imager.M2 Zeiss, using brightfield microscopy with 20x magnification. Cell measurements were taken at several potentially homologous regions on the dorsal and ventral petals (Figure 3, [B]) of corollas of the different species selected, representing the proximal and distal petal regions. For generalist species, the proximal petals were imaged between the base of the corolla and the ventral constriction, while the distal petals were imaged between the ventral constriction and the tip. For hummingbird specialists, the proximal petals were imaged at approximately the same distance from the base as generalists (most hummingbirds do not have a ventral constriction to serve as reference), while the distal petals, which have an increased length compared to the generalist syndrome, were imaged at three zones to have it more completely covered.

Thirty abaxial cells were measured at each region for their width and length, along the proximodistal flower axis, using ImageJ/Fiji (Schneider et al., 2012.). To test whether the specialists and generalists have distinct cell morphology properties at the different regions investigated, a Phylogenetic Mixed Model (PMM) that considers the phylogenetic relationships of the species was used. The analysis is conducted using two models: the first (m0) accounts for fixed effects and has no phylogenetic structure, the second (m1) accounts for both fixed and random effects. In the present study, the fixed effects correspond to the pollination strategies, while the

random effects correspond to the phylogenetic relatedness of species. The two models were applied for each petal region investigated and then compared using the Deviance Information Criterion (DIC). This was performed using the MCMCglmm R package (Hadfield JD, 2010) and the phylogeny of Joly et al. (2018). The Bayesian Markov Chain Monte Carlo (MCMC) analysis was run with the following settings: 1) number of generations for the MCMC (nitt = 55000), 2) interval in which the parameters were sampled (thin = 20), 3) number of generations to discard at the beginning of the search (burnin = 5000). The model was applied to test for convergence of cell length and the ratio (cell length/width) at the potentially homologous petal regions, within the strategies.

3.2.3 RNA-seq

RNA-seq was used to study gene expression in three zones of the corolla that are morphologically distinct between the syndromes and that were also found to have different cell shapes in mature petals (see results). The zones represent the ventral constrictions present in the proximal petal regions and the ventral and dorsal distal petal regions (Figure 3, [C]). The ventral constriction is highly pronounced in the generalists compared to the hummingbird specialists; the distal petal regions are subcampanulate and tubular in the generalists and hummingbird specialists, respectively.

Fresh material from the selected six species was harvested over a period of six months. After collecting, petals were immediately dissected for the isolation of the three corolla zones (Figure 3, [C]) and flash-frozen in liquid nitrogen. Multiple tissue samples were pooled for each biological replicate. This resulted in a total of 54 samples to be sequenced: triplicates of six species, three zones each. Samples were stored at -80°C until RNA extraction. Tissues were then ground to a fine powder and RNA was extracted using the Quick CTAB RNA extraction protocol (Gambino *et al.*, 2008). RNA concentration, integrity and quality were assessed with an Agilent Bioanalyser (Agilent, Santa Clara, CA, USA). Libraries were prepared using the NEB mRNA stranded Library preparation kit and Illumina adapters, and assessed with LabChipGX Touch, and RNA samples were sequenced by Illumina paired-ends sequencing (NOVAseq 6000 PE 100 - 25M reads) at the Genome Québec Innovation Centre (Montreal, Canada).

Raw RNA-seq reads were processed to remove low-quality readings using Trimmomatic (version 0.39; Bolger, Lohse and Usadel, 2014), with settings: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN: 36, HEADCROP:2. *De novo* transcriptome assembly was performed for all species with minimum contig lengths of 200 bp using the Trinity pipeline (version 2.14.0; Grabherr *et al*., 2011; Haas *et al.*, 2013) and transcript abundance was quantified by aligning the reads to the newly created transcriptomes using Kallisto (version 0.46.1; Bray *et al.,* 2016).

3.2.4 Orthology inference, annotation and gene ontology enrichment

Transcripts in different isoforms obtained from Trinity fasta files were converted to candidate coding sequences using the function *TransDecoder.LongOrfs* of TransDecoder (version 5.5.0; Haas et al., 2013), in order to predict open reading frames (ORF) and proteins. Orthologous genes between species were determined using OrthoMCL (version 2.0.9; Chen *et al.*, 2006). Transcripts were clustered into orthologous groups if present in at least two species.

Annotation of orthologous transcripts was performed using Trinotate (version 4.0.0; https://github.com/Trinotate/Trinotate.github.io/wiki) for each transcriptome. Samples were blasted against the SwissProt database (E-value cutoff of 10^{-4}) and gene ontology enrichment was performed with the R packages GOseq and qvalue (bioconductor version 3.17; 2023) for GO terms enrichment and to estimate false discovery rates (FDR), respectively. Gene ontology terms with FDR below or equal to 0.01 were considered enriched.

3.2.5 Differential expression analysis

Differential expression (DE) analysis was performed using the R package DESeq2 (Love et al., 2014) to determine the significant changes in gene expression levels and assess potential convergence in expression for the two strategies at the three selected zones. We noticed that the inter-individual variance was lower for the four species for which the biological replicates consist of clones in our study (data not shown). To reduce the impact of the absence of true genetic variation on the gene expression analysis, we used the mean expression of each transcript across all replicates of each species for the statistical analysis, reducing the dataset to 18 samples (six

species, three zones each). *Inter-species* DE analysis was then performed between the six selected species, considering the three species with the same pollination syndromes as replicates in an analysis contrasting the two syndromes. *Intra-species* DE analysis was also performed, but only on the two species for which we had three genetically distinct biological replicates (*G. acaulis* and *R. exsertum*) to assess whether there were differences in gene expression among the different regions. In the latter, counts from the three biological replicates were used in the analysis and the three zones were contrasted within species (ie. Zone A vs Zone B, in *R. exsertum*).

Transcript isoform abundances were imported with TXimport (version 3.2.3; Soneson et al., 2015), and converted to gene abundances. To allow a comparison between species, a matrix of raw counts for the homologous genes was created and converted into a *DEseqDataSet* (dds) object using the command *DESeqDataSetFromMatrix* (Supplementary material, S2). A matrix of gene length was also added to the object to be used as a normalization factor to correct for the variation between species (Supplementary material, S3). Although there was significant length variation for the assembled genes, the variation among orthologs was minimal. Genes that were interpreted to be differentially expressed were those with adjusted *p*-values below or equal to 0.01, obtained from Wald test using the Benjamini and Hochberg method. The *DESeqDataSet* object was created with a design formula expressing the variables to be used in modelling, which were *region* (zone A, zone B, zone C) and *strategy* (generalist, hummingbird specialist) for the contrasts between strategies. Because this method does not allow the phylogenetic relationships between the species to be accounted for, we also tested the quantitative changes in gene expression levels between *Gesneria* and *Rhytidophyllum*. The idea was to see if there were more differentially expressed genes between strategies than between more closely related species. Finally, to visualize gene expression between species, we performed a principal component analysis (PCA) on the homologous gene count data normalized via variance stabilizing transformation (Durbin et al., 2002).

3.2.6 Weighted gene correlation network analysis

Weighted gene correlation network analysis (WGCNA) was used as another way to test for convergence in gene expression among species. This clustering method was used to group

genes with similar patterns of expression across all species and zones into co-expression modules and then test these modules for correlation with the pollination strategies and genera.

Co-expression modules were obtained using the R package WGCNA (version 1.72-1; Langfelder and Horvath, 2008). Outliers were removed and a filter was applied to keep only the putative homologous genes across species that had counts greater or equal to 1 in 33% of the cluster samples. The dataset of genes was then normalized using variance stabilizing transformation (VST) with the *vst()* function from the DESEq2 package, and the orthologs were used to construct co-expression networks, using a soft-thresholding power of 12. The network was constructed with TOMtype = 'signed' and mergeCutHeight = 0.25. Module eigengenes were calculated and correlations between the modules and the traits 'strategy' and 'genera' were obtained.

3.3 Results

3.3.1 Cell measurement analysis

The observation of development in the investigated species demonstrated that flowers take approximately eight weeks to develop, growing slowly in the first seven weeks to reach 40-45% of their final size and then much more rapidly in the last week, to complete their development. Buds of all species were very similar at the 40-45% stage (corolla and cell shapes) and this suggests that this stage was a good candidate for looking at gene expression since differentiation begins at this phase.

A Phylogenetic Mixed Model was used to test for the potential convergence of cell morphology. The convergence of the MCMC runs was assessed using the Potential Scale Reduction Factors (PSRF) to compare the variance between and within runs. PSRF values for both fixed and random effects were very close to 1, confirming the convergence of the chains. For all petal regions, the model including the phylogeny as random effects had much lower DIC values than the model without random effects, indicating that it is important to account for the phylogeny in the analyses. The mixed model with the phylogeny was thus used in all analyses.

Results of the Phylogenetic Mixed Model (model m1) found a significant convergence (pMCMC < 0.05) of the pollination syndromes for cell length in all regions, both proximal and distal, with specialists having longer cells than generalists (Figure 4, [A]). Because the cell length alone could be correlated with flower size and the flowers of hummingbird specialists are all relatively larger than that of generalists, convergence was also tested for the anisotropy present in cell profiles, represented by the cell length-to-width ratio. The latter shows significant convergence at two regions of the distal petals where cells of hummingbird specialists have greater anisotropy (greater length-to-width ratios) than generalists (Figure 4, [B]).

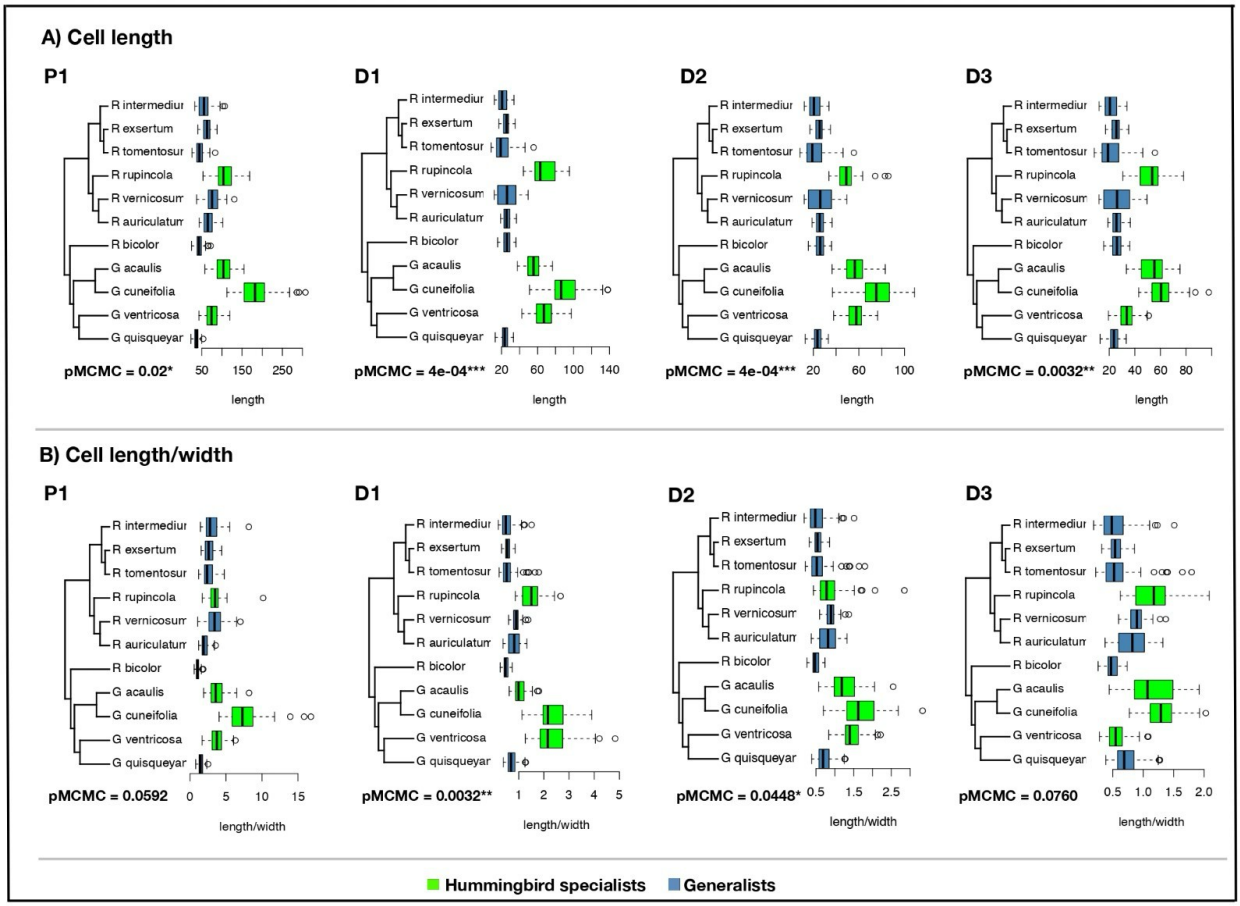

Figure 4. Cell measurement plots and *p*-values for MCMCglmm convergence. A) Cell length along the proximodistal axis in the proximal petal region (P1) and three distal petal regions (D1, D2, D3). **B)** Cell length-to-width ratio along the proximodistal axis in the proximal petal region (P1) and three distal petal regions (D1, D2, D3). Measures are combining the dorsal and ventral petals.

3.3.2 De novo transcriptomes assembly, annotation and orthology inference

The RNA extracted from the petal tissues of the six species selected for RNA-seq, collected at 40-45% of their final floral size at three potentially homologous zones was sequenced and yielded over 300M reads per species (Table 1). For each species, the reads from all samples were pooled to produce a single transcriptome. Assemblies produced between 139,490 (*R. auriculatum*) and 283,707 (*R. rupincola*) transcripts for each transcriptome (Table 1), with median contig lengths between 388 (*R. rupincola*) and 1394 (*R. auriculatum*) and N50 between 2244 (*G. bicolor*) and 2772 (*G. ventricosa*).

	G. acaulis	*G.ventricosa*	*G. bicolor*	*R. rupincola*	*R. exsertum*	*R.auriculatum*
Total sequencing reads	357,462,753	364,933,210	365,528,625	312,048,219	340,303,493	313,153,593
Total Trinity transcipts	164,044	155,744	245,493	283,707	226,388	139,490
Median contig length	1014	1205	701	388	776	1394
N50	2606	2772	2244	2452	2250	2768
Total assembled bases	248,752,311	260,582,331	303,575,498	310,214,048	289,408,751	241,279,435
Total Trinity genes	78,470	69,109	125,410	194,704	105,141	67,198

Table 1. Sequencing and statistics for the assemblies of the six newly created petal transcriptomes.

The transcript and protein sequences obtained from Transdecoder were used for the annotation. The search for putative orthologous genes between species with OrthoMCL resulted in a total of 19735 putative homologous genes present in at least two species, with 21.4% present in all six species, 12.7% present in five species, 13.2% present in four species, 18.4% present in three species and 34.3% present in two species (supplementary material, S1).

3.3.3 Gene expression analyses

To initially examine the patterns of petal gene expression of the selected species across the three zones investigated, a principal component analysis (PCA) was conducted with the entire set of read counts normalized with Variance Stabilizing Transformation (Figure 5, [A]). The first

principal component represents 26% of the variance and separates most species by the pollination strategy, except for *R. auriculatum* that groups with the hummingbird specialists, particularly close to its generalist relative *R. rupincola*. PC2 has 21% of the variance and tends to represent phylogenetic relationships, with the *Gesneria* species having positive scores and *Rhytidophyllum* species negative scores. The two generalists *G. bicolor* and *R. exsertum*, despite having fairly distant phylogenetic origins, group quite close to each other. The PCA does not show important differences between the three petal zones investigated; all samples from each species group very close to each other. The next principal components are mainly explained by differences between species.

3.3.3.1 Differential gene expression

An *Intra-species* differential expression (DE) analysis was first performed on the two species for which there were genetically distinct replicates (*R. exsertum* and *G. acaulis*). This analysis allowed us to test whether there were genes differentially expressed between zones A, B and C. We found 123 DE genes between the zones A and C, 9 DE genes between zones B and C and 45 DE genes between zones A and B, in *R. exsertum*. As for *G. acaulis*, we found 423 (A vs C), 3 (B vs C) and 333 (A vs B) DE genes between zones (Figure 5, [B]).

Inter-species DE analysis was conducted on all six species to quantify and test the potential convergence of gene expression between species sharing the same pollination strategy. In this analysis, we found a total of 959 significantly DE genes between the two pollination strategies (generalists vs hummingbird specialists) for zone A, 949 DE genes for zone B and 930 DE genes for zone C. To offer a point of comparison to quantify convergence in gene expression, we also quantified and tested the DE genes observed between the genera *Gesneria* and *Rhytidophyllum* as a means to account for the phylogenetic relationships. This analysis resulted in a total of 799 DE genes for zone A, 809 DE genes for zone B and 807 DE genes for zone C (Figure 5, [C]). In both analyses, most DE genes are present in all three zones (Figure 5, [D]).

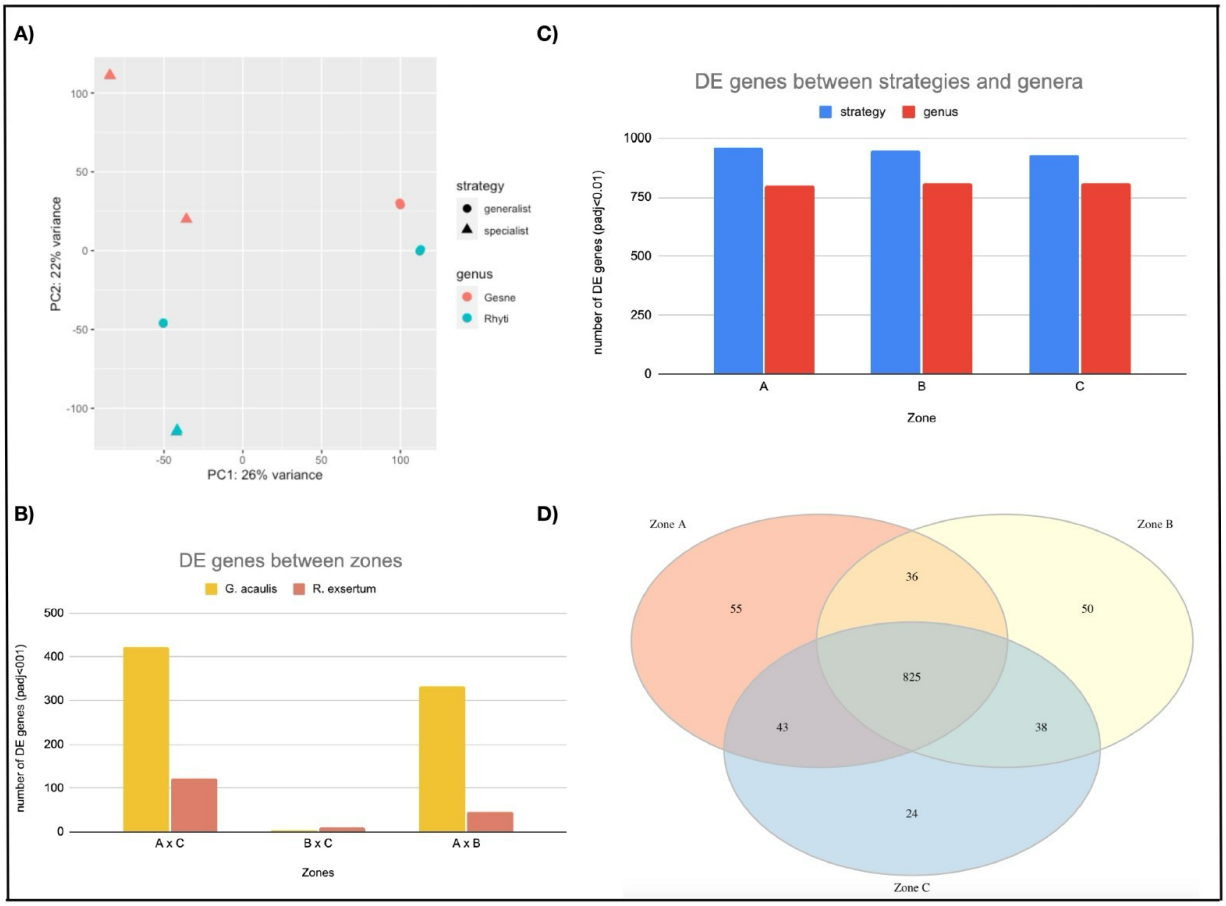

Figure 5. Summary of the differential gene expression analyses. A) Principal component analysis of expression data. **B)** Intra-species analysis of genes differentially expressed between zones, on *G. acaulis* and *R. exsertum*. **C)** Inter-species analysis of genes differentially expressed between strategies and between genera at zones A, B and C. **D)** Venn diagram of genes differentially expressed between strategies at zones A, B and C.

Gene ontology (GO) term analysis was performed on the list of 1071 genes differentially expressed between the pollination strategies and the 907 genes DE between genera. The background used to test for GO enrichment was a list of all homologous genes expressed in the RNA-seq data set. Gene ontology top terms (FDR =< 0.05) for genes DE between strategies and DE between genera are mostly of parental terms of cellular components. Overall, genes found to be DE had diverse GO terms, including biological processes associated with cell wall growth and cell division, which could influence flower shape.

3.3.3.2 Weighted gene co-expression network analysis (WGCNA)

To investigate convergence in gene expression with a different approach, we looked for genes with similar patterns of expression that were grouped into modules presenting high positive correlation with the pollination strategy. A total of 15,006 orthologs were used to construct co-expression modules and 33 modules were obtained (Figure 6), varying in size from 69 to 1255 genes.

Five modules (dark grey, n = 365; cyan, n = 481; yellow, n = 746; brown, n = 752; light cyan, n = 440) were significantly positively correlated with hummingbird pollination specialists and six (purple, n = 514; turquoise, n = 1255; black, n = 559; pale turquoise, n = 212; royal blue, n = 397; salmon, n = 503) were significantly positively correlated with pollination generalists. The genus *Rhytidophyllum* was positively correlated with six modules (magenta, n = 520; midnight blue, n = 477; purple, n = 514; blue, n = 789; salmon, n = 503; grey, n = 69), while *Gesneria* was positively correlated with six modules (tan, n = 508; green yellow, n = 509; dark grey, n = 365; red, n = 595; brown, n = 752; green, n = 634).

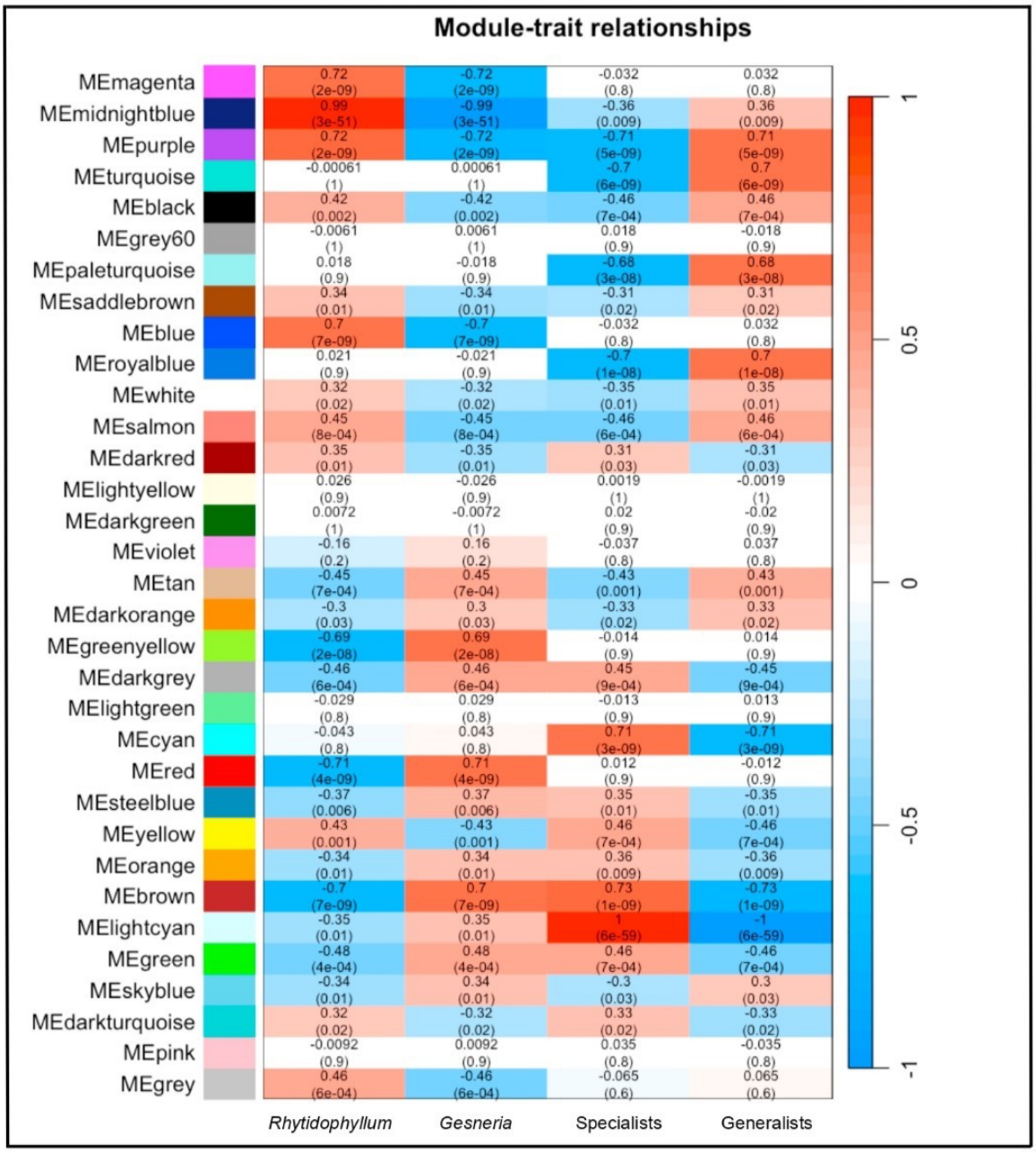

Figure 6. Heatmap. 33 modules with correlations and *p*-values corresponding to the traits tested. Modules are in rows and traits in columns. A color scale shows negative correlations in blue and positive correlations in red.

3.3.3.3 Convergence between approaches

To determine whether the same genes were found to be correlated to pollination syndromes between the differential gene expression and the WGCNA approaches, the intersections between the lists of differentially expressed genes and the genes in the modules with the lowest *p*-values

for the traits tested in WGCNA (strategy and genera) were obtained. The light cyan module (p = 6e-59; n = 440) had 46.1% (203) genes in common with the list of differentially expressed genes between the two pollination strategies. Similarly, the midnight blue module (p = 3e-51; n = 477) had 46.5% (222) genes in common with the list of differentially expressed genes between the two genera. GO enrichment analysis was performed on modules light cyan, midnight blue and on the list of genes which overlapped between approaches, in both traits. The list of genes related to strategy had no GO terms enriched, while the list of genes which overlapped for genera had top terms enriched for negative regulation of peroxidase and oxidoreductase activities. Modules light cyan and midnight blue had no GO enriched terms.

3.4 Discussion

Phenotypic convergence among species can occur through various processes. To better understand these processes, this study investigated the morphological convergence of flowers of non-model species with a comprehensive approach to investigate whether the flower shape convergence is also observed at two other levels: cellular morphology and gene expression. The results revealed that the morphological convergence we observe in flowers is also present at smaller biological scales.

The analysis of the abaxial cellular profile of mature petals shows that the species with tubular shape have longer cells and cells with a greater length to width ratio than the species with subcampanulate shape in the distal region of the dorsal and ventral petals. Cell length can be correlated with size and indeed hummingbird pollinated species studied have bigger flowers than generalist species. But the different cell length-to-width ratio between specialists and generalists clearly supports that convergence of flower shape is reflected in the cell shape of the corolla. Because all species have similar cell morphologies at the 40-45% bud stage, the cell morphologies of the fully developed petals suggest that the cell shape differentiation at later stages of development occurs in the proximodistal axis in tubular flowers and in the transverse axis in subcampanulate flowers. This ratio distinction is not observed in proximal petal regions, where the cell morphology differences between tubular and subcampanulate flowers are not as important. Our results suggest that convergent flower shapes in *Gesneria* and *Rhitydophyllum* species are caused by convergent cell directional growth strategies in the distal petals, that distinguish the

tubular and subcampanulate flower morphologies. Anisotropy also seems to play an important role in flower shape determination in other taxa, such as *Antirrhinum* (Rolland-Lagan et al., 2003) and *Aquilegia.* In the latter, it is responsible for 99% of spur-length variation between species (Puzey et al., 2012).

In addition to the convergence detected at the cell morphology level on the convergent flower shapes, there are also smaller idiosyncratic variations on the cell morphology analysis within certain species that support a correlation between the length-to-width ratios and individual flower shapes. For example, the corolla tube of *G. ventricosa*, despite being tubular, gets wider with increasing distance along the proximodistal axis and its length-to-width ratio gradually decreases from proximal to distal regions of the petal, being smaller at the most distal point, where the flower is widest. In *R. rupincola*, the corolla tube is widest in the center of the petal and the length-to-width ratio is also lowest at this central region. Finally, *G. cuneifolia*, which has a more consistent cylindrical shape, with no noticeable enlargements, presents a constant length-to-width ratio along the corolla tube. These observations within species reinforce that in this group the length-to-width ratio of cells plays a role in determining corolla shape. This is observed both in the individual analysis and in the analysis among species, where convergence is detected among species of the same shape.

We also investigated whether we could detect evidence of convergence at a lower biological level, that is, the gene expression level. Other studies have investigated if floral convergence is paralleled at the gene expression level, but most of these studies have looked at gene expression of the whole flower, including its four organs and sometimes more than one developmental stage (Serrano-Serrano *et al*, 2017; Roberts and Roalson, 2017, 2019). Although these studies have the advantage of gathering a more complete list of the genes putatively involved in flower development, they lack specificity regarding the functions that these genes may play in flower development. We chose to sample tissues for the gene expression study at the stage of development when the morphological differentiation between the two flower shapes starts to be noticeable and in specific zones of the petals that differ between the two distinct shapes in the mature flowers.

In the first analysis of differential gene expression, performed within species, we found differentially expressed genes between petal regions for *G. acaulis* and *R. exsertum,* the two

species for which we had genetically distinct biological replicates, suggesting that it is important to be as specific as possible in selecting the tissues to study when investigating correlative analyses across species. This is not surprising as it is well known from spatial transcriptomics studies that large differences in gene expression can occur along short physical distances (Moses and Pachter, 2022). In the intra-species analysis, the number of differentially expressed genes between the ventral and dorsal distal regions of the petal (zones B vs C) is much smaller than for the comparison between the distal and proximal regions (A vs B and A vs C). This is expected and corresponds to the previous cellular morphology analysis, which demonstrated that in the distal region of the petals, cell shapes are more similar than between the proximal and distal regions.

We also found evidence of convergence at the gene expression level that paralleled the convergence in corolla shape. Because no robust phylogenetic comparative methods exist to study gene expression, it is not possible to use the same approach as the one used for the cell morphological analysis and control for the phylogeny in the statistical tests. Instead, we used a different approach and compared the results of differential gene expression analysis between flower types and between the monophyletic genera *Gesneria* and *Rhytidophyllum* (Joly et al. 2018). We found that the number of differentially expressed genes between the flower types was greater than between the two genera, which we interpret as evidence for convergence in gene expression. In other words, the shape factor (pollination syndrome), within the scope of this analysis, has a greater influence on the genetic expression by species than their phylogenetic proximity.

A total of 1071 genes were differentially expressed between tubular and subcampanulate flower forms when all three zones were considered. Among these, 825 genes are common to the three zones. To get a better sense of the genes that could be involved in shape determination, we performed the gene ontology analysis of these genes, but only broad GO terms were found to be significant. The 825 genes common to all zones may be linked to aspects other than shape that distinguish the two syndromes. For example, they may be associated with color production, as the tubular species studied are all red while the subcampanulate species vary from pale yellow to brown. Genes active exclusively in one of the three zones, however, may play roles related to the shape of these zones. For instance, the 50 and 24 genes unique to the distal ventral and distal dorsal regions (zones B and C), respectively, could be associated with production of the distal curvatures on the distinct tubular and subcampanulate corollas. The list of 50 genes includes a homologue to

the *Arabidopsis ASA1*, which plays an important regulatory role in auxin production. As for the 55 genes unique to the proximal region of the petal (zone A), these could be associated with production of the basal constriction of generalist species, which possibly have the important functional aspect of allowing the generalist syndrome to be pollinated by different classes of pollinators (Fenster and Martén-Rodriguez, 2008). Indeed, this list includes genes homologous to the *Arabidopsis* and *Soybean CRK9,* which have GO terms for programmed cell death, and to the *Arabidopsis* and *Oryza sativa CESA5,* involved in the production of cellulose and cell wall organization/formation (Appendix, Table 2). These are interesting candidates to be further investigated since programmed cell death or cell wall modifications could play a role in shaping the basal constriction and cell wall expansion are important determinants of plant tissue growth (Cosgrove 2005).

Convergence in gene expression is also generally supported by the Principal Component Analysis. The first principal component indeed generally discriminates species according to their pollination strategy and thus flower shape, whereas the second principal component separates species according to their genus. To further detect convergence at the gene expression level and to reinforce the differential gene expression analyses, we also performed an analysis based on gene coexpression. Eleven modules were significantly correlated with a pollination strategy or corolla shape, further supporting convergence at the gene expression level. Among these, the module with the highest correlation to strategy has 50.7% of its genes also present in the list of genes differentially expressed between flower shapes. As this clustering approach focuses on networks of genes that are potentially functionally associated, these 203 genes make interesting candidates to be linked to shape determination. Among the 203 genes, we find genes homologous to the *Soybean A10A5*, an auxin-induced protein, and to the *Arabidopsis TTG1*, which is a protein involved in epidermal cell fate commitment, response to auxin and anthocyanin biosynthesis. Other interesting genes found in this list are *MUC70*, involved in pectin synthesis, *NAT6*, a cell wall component, and the *Arabidopsis P2C10,* involved in myosin phosphatase (Appendix, Table 3). In plants, myosin movements are necessary for cell growth and polarity and can be involved in cellular expansion (Madison and Nebenfuhr, 2013). The module most correlated to the phylogeny of the species (midnight blue: n = 477; correlation = 0.99, *p*-value = 3e-51) also had a little over 50% of its genes present in the differential expression analysis, corroborating the percentage of intersection found in the analysis of the strategy. As DESeq2 and WGCNA analyze expression

data with different approaches, it is notable that more than half of the genes found in the light cyan and midnight blue modules are also genes given as differentially expressed by strategy and genera, respectively. This reinforces the results and indicates that the two methods can complement each other, suggesting an interesting strategy for transcriptional profiling studies.

Gene expression dynamics change throughout development (Vincent and Coen, 2004; Puzey et al., 2012; Yant et al., 2015), promoting different cell behaviours, such as localized cell divisions in an initial phase, followed by later phases of cell expansion and directional growth. Thus, the genes identified in this study, which was performed at a late developmental phase (40-45% final size), might be associated with the distinct anisotropic cell expansion found by the cell measurement analysis. Therefore, genes linked to cell expansion could be further investigated in the Gesneriaceae, such as those associated with brassinosteroids and auxin signaling, responsible for cell elongation in *Aquilegia* flowers (Conway et al., 2021; Edwards et al., 2021). In addition, it is important to associate these with structural cell properties such as the cell wall, as cell expansion depends on the degree of extensibility of cell walls and on how the cellulose and pectin fibers of the walls are positioned (Cosgrove DJ, 2016; Zhang et al., 2019). It is therefore important to investigate proteins such as α-expansin and Cel12A, which promote such extensibility (Cosgrove, 1999), as well as genes associated with the polarity fields that might promote directional elongation.

3.5 Conclusion

We performed statistical analysis of cell measurements and RNA-seq analysis on species with two distinct flower shapes to investigate convergence at the cell and gene expression levels. We found that there is convergence in the cellular shapes of cells in the distal region of the petals, where species with tubular flowers have cells elongated in the proximodistal direction and subcampanulate species have cells elongated in the transverse direction. In addition, we found convergence in gene expression in three critical regions of the corolla for species of the same pollination syndrome that have similar flower shapes.

3.4 Supplementary material

Supplemental information is accessible via this link:

https://figshare.com/articles/dataset/Comparative_analysis_of_corolla_shape_transitions_in_the_sister_genera_i_Gesneria_i_and_i_Rhytidophyllum_i_Gesneriaceae_/25555725

4. General discussion

The global objective of this study was to investigate the origin of the phenotypic convergence in flowers of non-model species. As evolution uses different ways to adapt, it is important to promote studies across taxa and uncover this diversity. One outstanding question surrounding convergence is whether it occurs through similar or distinct molecular processes. This question is difficult to address because one first must determine the relevant molecular level that should be considered, which could be an amino acid, a gene or several genes at different loci independently recruited to produce a given phenotype. All these situations have already been found in studies (Zhang and Kumar, 1997; Parker et al., 2013) and certainly occur widely in nature. But as phenotypes are often determined through the combined action of several genes, it is interesting to think about molecular convergence in a broader way, as demonstrated by genome-scale studies (Hu et al., 2023). Therefore, this study sought to investigate convergence in the shape of flowers in a comprehensive manner. The goal was to find convergence in gene expression patterns, rather than trying to identify specific candidate genes. Likewise, and with the aim of studying convergence at the cellular level, the project also tested cellular morphology patterns between the species that converge.

The idea of investigating molecular convergence at the level of gene expression came from different points. From a practical point of view, this is an approach that is methodologically viable in studies of non-model flowers due to better sequencing technologies, and that has previously demonstrated interesting results in plant systems. Furthermore, it has been widely demonstrated that changes in cellular expression patterns and levels is correlated with the evolution of specific morphological phenotypes (Wray, 2007; Stern and Orgogozo, 2008; Romero et al, 2012). Also, the role of transcription factors in determining floral phenotypes has already been widely demonstrated. Another aspect of our methodology that is worth highlighting is that we investigate

tissue-specific (petal) gene regulation. It has already been demonstrated that evolution in gene expression can occur under tissue-specific selection pressures (Blekhman et al., 2008), which is not surprising given that it is a relatively less restricted process that can act differently temporally and spatially.

At the gene expression scale, we found evidence of convergence in the analysis of differential gene expression, where the number of differentially expressed genes is greater among tubular species versus subcampanulate species than between species of the genus *Rhytidophyllum* versus *Gesneria*, making the shape factor more determinant than their phylogenetic relationship. At the cellular scale, we also found convergence in the length-to-width ratio of the cells in the distal regions of the petals, where tubular species (hummingbird specialists) have a higher ratio than that of subcampanulate species (generalists). This cell shape convergence may be an important factor contributing to the formation of the elongated tubular corollas of hummingbird specialists and the more open or bell-shaped corollas of generalists. These two findings bring a new contribution to future shape studies in species of the group.

4.1 Cell measurement analysis

The cell morphology analysis was done first and had two purposes. The first was to make the investigation of convergence more complete, including the cellular level in addition to the molecular. The second was to obtain indications about where there were significant differences between the tubular and subcampanulate flowers at the cellular level, which would help us define the sampling for the second gene expression analysis with regard to the zones that should be collected for sequencing, because if there were clear distinctions at the cellular level, we could potentially find them at the molecular level as well. This analysis indicated that there were significant distinctions at the dorsal and ventral distal petal regions, hence these regions were further analyzed by RNA-seq.

4.1.1 Convergence of directional cell growth (anisotropy) in distal petal regions

The statistical analysis model used to compare the distinct flowers shapes revealed that the species with a tubular shape have cells with a length-to-width ratio significantly greater than the

species with a subcampanulate shape in the distal petal regions. This suggests that in this region, the directional growth of the tubular flowers occurs in the proximodistal axis, while that of the subcampanulate flowers occurs in the transverse axis. Also, as no significant differences in cell shape were observed in buds of 40-45% of the final flower size, this result suggests that this differentiation in cell shape occurs in later stages of development of these species. This aspect of cell growth may be an important determining factor in the final shape of the petals. According to Rebocho *et al*. (2017), the heterogeneity of cell growth in different regions of a tissue sheet, that is, cell growth at different rates and directions, is one of the two mechanisms that can generate deformations and curvatures, which lead to shape. In this analysis, the rate of growth of the cells was not evaluated, but the observation of the shape of cells in the fully developed petals suggests the direction of growth.

This cell shape distinction in tubular versus subcampanulate flowers is not observed in the proximal regions of the petals, where the distinction between the two shapes is not morphologically striking. If we observe the tested species only in their proximal regions, more specifically in the region proximal to the basal constriction of generalists, there is no important difference between the two syndromes. This observation corroborates the result above, in which the distinction of ratio in the distal region is an important factor in determining the two types of corollas. According to previous studies in *Aquilegia* spurs, cell directional growth (anisotropy) is responsible for 99% of spur-length variation between species (Puzey et al., 2012), indicating that cell anisotropy is largely responsible for flower shape distinctions in the genus. Our results may suggest that *Gesneria* and *Rhytidophyllum* species could rely on convergent cellular anisotropy to reach their final forms.

4.1.2 Association between cell length-to-width ratio and aspects of individual flower shapes

In addition to the ratio distinctions discussed above, there is also a correlation between the length-to-width ratio and the shape of certain species, when analyzed separately. It is observed that this ratio is lower in regions where the petal enlarges. For example, the flower of the species *G. ventricosa*, despite being tubular, shows a slight enlargement that gradually increases along the proximal-distal axis (Figure 3, [B]). In this species, the ratio decreases at the most distal point of the petal, where the flower is wider. The flower of *R. rupincola*, which enlarges in the central region of the petal and then becomes narrower towards the tip (Figure 3, [B]), also has a lower

ratio where the petal is widest. As for *G. cuneifolia*, which has a more consistent cylindrical shape (Figure 3, [B]), with no noticeable enlargements, there are no significant changes in the length-to-width ratio.

These intra-species observations were important in corroborating the length-to-width convergence discussed in the above section (4.1.1). The latter was challenging because, although the analysis assumes a general shape pattern that is common to species of the same syndrome (tubular and subcampanulate), these still present important disparities among themselves, as seen between the tubular *G. ventricosa, G. cuneifolia* and *R. rupincol*a. These disparities, which are particularly true among tubular species and have been noted by Joly *et al.* (2018), could impair statistical comparisons due to variations among species of the same experimental group (in this case, syndrome). However, this intra-species analysis demonstrates the same trend (smaller length-to-width ratios in enlarged regions of the petals) and supports the previously obtained results.

4.2 Gene expression analyses

Many comparative transcriptomics studies compare pools of RNA collected from the entire flower ensemble, including its four organs and sometimes more than one developmental stage. These studies gather more complete lists of genes involved in flower development; however, they lack specificity regarding the functions that these genes may play in flower development. This study aimed to specify its sampling with respect to 1) the stage of development (40-45% of final size), 2) the flower organ (petals) and 3) selected petal zones with potentially important functional aspects (A, B and C). These three aspects of the sampling scheme were carefully defined to maximize our chances of obtaining meaningful and relevant results.

Regarding the developmental stage, many RNA-seq studies collect samples at more than one stage and then compare differences in gene expression levels between stages to know which genes are up or down regulated. This is an interesting method, but it increases project costs as more samples need to be sequenced and could not be afforded. As our objective was to investigate convergence between species, it was important to have a certain number of species to analyze, so we already had decided to sequence at least six species. We could then choose between sequencing more than one developmental stage or more than one petal zone. And we chose the latter, for the following reasons: (1) the previous cellular analysis had already indicated to us that there were

significant differences between the tubular and subcampanulate species in certain areas of the petals, (2) the focus of the study was not to investigate up or down regulation of putative genes, but rather to test convergence in the general pattern of gene expression, (3) the final shape of the petals is largely defined at later developmental stages. So we decided that the genes we were interested in would probably not be active in early phases, but more specifically in the stage chosen for sampling. The 40-45% stage was chosen because it is the moment that occurs just before the rapid final elongation of the flowers, when the buds undergo the greatest change to reach the final shapes that distinguish tubular and subcampanulate flowers.

The species were selected according to the availability of plants in the greenhouses. Ideally we would have chosen three monophyletic pairs of tubular and subcampanulate flowers, but this sampling was not possible. Therefore, we chose three species of tubular flowers and three species of subcampanulate flowers that had different evolutionary origins.

The analyses in this section gave us interesting and somewhat unexpected results. Firstly, we did not expect to obtain more genes differentially expressed according to flower shape than according to their genera. The contrast between genera was essentially used as a reference for the contrast between forms, since similarities in the expression patterns of species belonging to the same genus, and thus due to their common ancestry, would certainly be present. My expectation was that there would be more differentially expressed genes according to genus and a relatively lower, but still significant, number of genes differentially expressed according to shape. The fact that the result was a slightly higher number in the contrast between forms confirmed the hypothesis that there is convergence of expression between forms. It was reassuring that the clustering analysis found modules associated with the shapes. And, as in the DE analysis, the module most significantly associated with shape had a lower *p*-value than the module most significantly associated with genera. In addition, it was as well unexpected that there would be so many genes in common found by these two methods, which analyze gene expression data in very different ways.

4.2.1 Genes differentially expressed between zones of the petal

The intra-species analysis of genes differentially expressed between zones of the petal, carried out in the two species that had genetically distinct replicates, *G. acaulis* and *R. exsertum,*

revealed that at this stage of development, the corolla gene expression of these species varies significantly according to location. This analysis demonstrates the relative importance of separating regions of interest before sequencing. It was noted that the greatest distinction in gene expression occurs between zones A vs B and A vs C, or between the proximal and distal regions of the petal. Therefore, future studies could focus on these two regions only. Although the analysis between zones could only have been carried out on two species, it was also valuable in the inter-species analysis, as it allowed us to obtain lists of genes uniquely differentially expressed in each of the three zones, that could be associated with the production of the dorsal and ventral distal curvatures and the basal constriction.

4.2.2 Genes differentially expressed in tubular versus subcampanulate species

The analysis of differential expression between tubular and subcampanulate species was the most important of the project and the reason why the previous analyses were carried out. For example, the analysis on cell shape, performed in the first section of the project, was primarily done to indicate the zones that should be sequenced. Also, the development observation of the species was made to find out which stage of the development of these flowers seemed the most interesting to investigate - and we decided to investigate the beginning of the elongation phase. Species were selected with this analysis in mind, as we wanted an equal number of generalists and specialists and an equal number of *Gesneria* and *Rhytidophyllum* species. Ultimately, we did our best to make the analysis possible with the resources we had.

This careful selection was also made because this type of analysis is not common. Typically, differences in gene expression levels are determined intra-species, comparing different tissues of an organism, treatment versus control, etc. That is because this analysis does not have tools to calculate for the variance between homologues. Yet, this type of inter-species analysis is gradually increasing in studies, especially those that compare closely related species. This is not surprising as studies of this type have a strong power of inference if the variation between homologues is properly accounted for. In our study, we followed a suggestion from the author of the DESeq2 package and prepared our data after read counts, including a matrix of gene length between orthologs, which was used for normalization. Another challenge for us was heterochrony, the difference in the relative timing of development between species. As our project tested a

developing organ, we needed to determine a criteria to make their stage of development comparable and adopted a method used in other similar studies, which is the percentage of final size. We determined a relatively broad window between 40-45% of final size to collect the buds of the different species in order to ensure comparability. Finally, there was some sampling challenge, as numerous buds needed to be collected at a specific stage of development and dissected in three specific zones, across the six species tested. As these buds were relatively small (around 1 cm), dozens of samples, in triplicates, were collected to enable a minimum amount of tissue for sequencing. All these challenges are inherent to gene expression analyses, which, unlike genomic analyses, have spatial and temporal fluctuations. However, comparative transcriptomics is a powerful tool for providing insights into the evolution of diverse phenotypes.

We obtained exciting results regarding genes that might play a role in shape determination. We have different lists of candidate genes which are differentially expressed between tubular and subcampanulate flowers, in specific zones of the petals. Furthermore, we have lists of genes that were detected by more than one RNA-seq analysis method. In these lists, there are putative genes linked to the response to auxin, myosin movements, cell wall formation, programmed cell death, flower development and epidermal cell fate commitment, all functions highly related to shape determination.

4.3 Perspectives

The study brought interesting results on cellular and gene expression patterns associated with the morphological convergence of the group's flowers, which demonstrates that this is a good approach to be used in non-model species as a first step. The results obtained here, both in the broad sense of confirming convergence at these scales, and in more specific aspects such as the difference in ratios between the two shapes and the candidate genes to act in shape determination, can be widely explored in future studies in the species tested and in other species of the group. Future studies may add other species of the group that also present the generalist and hummingbird specialist syndromes. For example, of the eleven species that were tested in the cellular analysis, only six were sequenced. Also, it would be interesting to add species that present the bat specialist syndrome. The flowers of the latter resemble generalist flowers as they also are bell-shaped, but the basal constriction is absent in them. Therefore, the zone A of bat specialists could be sequenced

and compared with the analyzes in this study to determine which genes would be active/inactive in bell-shaped flowers without constriction. There is also the possibility of sequencing a different developmental stage and testing whether candidate genes are up or downregulated between stages. Something that could be changed in future investigations is that zones B and C could be pooled, given that differences in expression levels were less important when comparing these zones. Finally, we did not detect any of the genes previously found in the shape QTL study between *R. rupincola* and *R. auriculatum* (ie. *GLOBOSA, RADIALIS* and *JAGGED*) to be candidates for shape determination.

A next step could be chromatin immunoprecipitation sequencing (ChIP-seq) analysis, to investigate epigenetic aspects of the process or the correlation between differential gene expression and regulatory mechanisms. To move beyond the comparison of expression levels would be to perform functional experiments to characterize the genes of interest revealing how differences in gene expression levels affect phenotypes and to possibly link these findings with their adaptive ecological function.

5. Conclusion

Understanding the mechanisms and processes involved in the evolutionary process is necessary to protect and safeguard the natural environment, as well as to predict adaptive changes and prepare for future events. An interesting subject to be used in this learning process is the floral diversity generated through evolution, which has several facets and can occur for different reasons and mechanisms in different plant clades. Considering the extensive angiosperms, it is necessary to expand investigations of such processes beyond model species, so that we can achieve a more complete understanding of the different mechanisms used by nature for adaptation. As studies of non-model species have limitations, it is important to take advantage of new technologies available for large-scale data analysis, thus increasing the range of resources to be made available for future studies and advancing the field. Furthermore, as there is normally a correlation between the different biological scales, it is also interesting to seek holistic analyzes that investigate subjects at their different scales (morphological, cellular, molecular, functional). Our study demonstrated that the phenotypic convergence of a group of flowers from the Gesneriaceae family is paralleled

by convergence at the cell and gene expression levels, and detected putative genes that might be associated with flower shape determination.

References

Abraham, Mary C., et al. "Natural Variation Identifies Multiple Loci Controlling Petal Shape and Size in Arabidopsis Thaliana." *PLOS ONE*, vol. 8, no. 2, Feb. 2013, p. e56743. *PLoS Journals*, https://doi.org/10.1371/journal.pone.0056743.

Acevedo-Rodríguez, Pedro, and Mark T. Strong. *Catalogue of Seed Plants of the West Indies*. 2012. *repository.si.edu*, http://repository.si.edu/xmlui/handle/10088/17551.

Alexandre, Hermine, et al. "Genetic Architecture of Pollination Syndrome Transition between Hummingbird-Specialist and Generalist Species in the Genus Rhytidophyllum (Gesneriaceae)." *PeerJ*, vol. 3, June 2015, p. e1028. *peerj.com*, https://doi.org/10.7717/peerj.1028.

Almeida, Jorge, et al. *Genetic Control of FLower Shape in Antirrhinum Majus*. p. 6.

Angenent, Gerco C., and Lucia Colombo. "Molecular Control of Ovule Development." *Trends in Plant Science*, vol. 1, no. 7, July 1996, pp. 228–32. *ScienceDirect*, https://doi.org/10.1016/1360-1385(96)86900-7.

Arendt, Jeff, and David Reznick. "Convergence and Parallelism Reconsidered: What Have We Learned about the Genetics of Adaptation?" *Trends in Ecology & Evolution*, vol. 23, no. 1, Jan. 2008, pp. 26–32. *ScienceDirect*, https://doi.org/10.1016/j.tree.2007.09.011.

Ballerini, Evangeline S., et al. "Comparative Transcriptomics of Early Petal Development across Four Diverse Species of Aquilegia Reveal Few Genes Consistently Associated with Nectar Spur Development." *BMC Genomics*, vol. 20, no. 1, Aug. 2019, p. 668. *BioMed Central*, https://doi.org/10.1186/s12864-019-6002-9.

Barrett, Rowan D. H., et al. "Linking a Mutation to Survival in Wild Mice." *Science*, vol. 363, no. 6426, Feb. 2019, pp. 499–504. *science.sciencemag.org*, https://doi.org/10.1126/science.aav3824.

Barrett, Spencer C. H. "Darwin's Legacy: The Forms, Function and Sexual Diversity of Flowers." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 365, no. 1539, Feb. 2010, pp. 351–68. *PubMed*, https://doi.org/10.1098/rstb.2009.0212.

Bilbao, Gonzalo, et al. "Judge It by Its Shape: A Pollinator-Blind Approach Reveals Convergence in Petal Shape and Infers Pollination Modes in the Genus Erythrina." *American Journal of Botany*, vol. 108, no. 9, Sept. 2021, pp. 1716–30. *PubMed*, https://doi.org/10.1002/ajb2.1735.

Blekhman, Ran, et al. "Gene Regulation in Primates Evolves under Tissue-Specific Selection Pressures." *PLoS Genetics*, vol. 4, no. 11, Nov. 2008, p. e1000271. *PubMed*, https://doi.org/10.1371/journal.pgen.1000271.

Bradshaw, H. D., et al. "Quantitative Trait Loci Affecting Differences in Floral Morphology between Two Species of Monkeyflower (Mimulus)." *Genetics*, vol. 149, no. 1, May 1998, pp. 367–82.

Bradshaw, H. D., and Douglas W. Schemske. "Allele Substitution at a Flower Colour Locus Produces a Pollinator Shift in Monkeyflowers." *Nature*, vol. 426, no. 6963, Nov. 2003, pp. 176–78. *PubMed*, https://doi.org/10.1038/nature02106.

Brawand, David, et al. "The Evolution of Gene Expression Levels in Mammalian Organs." *Nature*, vol. 478, no. 7369, Oct. 2011, pp. 343–48. *PubMed*, https://doi.org/10.1038/nature10532.

Bray, Nicolas L., et al. "Near-Optimal Probabilistic RNA-Seq Quantification." *Nature Biotechnology*, vol. 34, no. 5, May 2016, pp. 525–27. *www.nature.com*, https://doi.org/10.1038/nbt.3519.

Castoe, Todd A., et al. "Evidence for an Ancient Adaptive Episode of Convergent Molecular Evolution." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 22, June 2009, pp. 8986–91. *PubMed*, https://doi.org/10.1073/pnas.0900233106.

Causier, Barry, et al. "Floral Organ Identity: 20 Years of ABCs." *Seminars in Cell & Developmental Biology*, vol. 21, no. 1, Feb. 2010, pp. 73–79. *PubMed*, https://doi.org/10.1016/j.semcdb.2009.10.005.

Chen, Feng, et al. "OrthoMCL-DB: Querying a Comprehensive Multi-Species Collection of Ortholog Groups." *Nucleic Acids Research*, vol. 34, no. Database issue, Jan. 2006, pp. D363-368. *PubMed*, https://doi.org/10.1093/nar/gkj123.

Christin, Pascal-Antoine, et al. "Causes and Evolutionary Significance of Genetic Convergence." *Trends in Genetics*, vol. 26, no. 9, Sept. 2010, pp. 400–05. *ScienceDirect*, https://doi.org/10.1016/j.tig.2010.06.005.

Clark, John, et al. "Convergence of Anti-Bee Pollination Mechanisms in the Neotropical Plant Genus Drymonia (Gesneriaceae)." *Evolutionary Ecology*, vol. 29, May 2015, pp. 355–77. *ResearchGate*, https://doi.org/10.1007/s10682-014-9729-4.

Coen, Enrico S., and Elliot M. Meyerowitz. "The War of the Whorls: Genetic Interactions Controlling Flower Development." *Nature*, vol. 353, no. 6339, 6339, Sept. 1991, pp. 31–37. *www.nature.com*, https://doi.org/10.1038/353031a0.

Conway, Stephanie J., et al. *Brassinosteroids Regulate Petal Spur Length in Aquilegia by Controlling Cell Elongation*. p. 12.

Cosgrove, D. J. "Enzymes and Other Agents That Enhance Cell Wall Extensibility." *Annual Review of Plant Physiology and Plant Molecular Biology*, vol. 50, 1999, pp. 391–417. *PubMed*, https://doi.org/10.1146/annurev.arplant.50.1.391.

Cosgrove, Daniel J. "Growth of the Plant Cell Wall." *Nature Reviews Molecular Cell Biology*, vol. 6, no. 11, 11, Nov. 2005, pp. 850–61. *www.nature.com*, https://doi.org/10.1038/nrm1746.

---. "Plant Cell Wall Extensibility: Connecting Plant Cell Growth with Cell Wall Structure, Mechanics, and the Action of Wall-Modifying Enzymes." *Journal of Experimental Botany*, vol. 67, no. 2, Jan. 2016, pp. 463–76. *Silverchair*, https://doi.org/10.1093/jxb/erv511.

Crawford, Brian C. W., et al. "CINCINNATA Controls Both Cell Differentiation and Growth in Petal Lobes and Leaves of Antirrhinum." *Plant Physiology*, vol. 135, no. 1, May 2004, pp. 244–53. *PubMed Central*, https://doi.org/10.1104/pp.103.036368.

Cronk, Quentin, and Isidro Ojeda. "Bird-Pollinated Flowers in an Evolutionary and Molecular Context." *Journal of Experimental Botany*, vol. 59, no. 4, 2008, pp. 715–27. *PubMed*, https://doi.org/10.1093/jxb/ern009.

Dellinger, Agnes S. "Pollination Syndromes in the 21st Century: Where Do We Stand and Where May We Go?" *The New Phytologist*, vol. 228, no. 4, Nov. 2020, pp. 1193–213. *PubMed*, https://doi.org/10.1111/nph.16793.

Ditta, Gary, et al. "The SEP4 Gene of Arabidopsis Thaliana Functions in Floral Organ and Meristem Identity." *Current Biology: CB*, vol. 14, no. 21, Nov. 2004, pp. 1935–40. *PubMed*, https://doi.org/10.1016/j.cub.2004.10.028.

Donoghue, Michael J., and Richard H. Ree. "Homoplasy and Developmental Constraint: A Model and an Example from Plants1." *American Zoologist*, vol. 40, no. 5, Oct. 2000, pp. 759–69. *Silverchair*, https://doi.org/10.1093/icb/40.5.759.

Durbin, Blythe, et al. "A Variance-Stabilizing Transformation for Gene-Expression Microarray Data." *Bioinformatics*, vol. 18, 2002. *escholarship.org*, https://escholarship.org/uc/item/37j5b4z7.

Edwards, Molly B., Evangeline S. Ballerini, et al. "Complex Developmental and Transcriptional Dynamics Underlie Pollinator-Driven Evolutionary Transitions in Nectar Spur Morphology in Aquilegia (Columbine)." *American Journal of Botany*, vol. 109, no. 9, 2022, pp. 1360–81. *Wiley Online Library*, https://doi.org/10.1002/ajb2.16046.

Edwards, Molly B., Gary P. T. Choi, et al. "Genetic Architecture of Floral Traits in Bee- and Hummingbird-Pollinated Sister Species of Aquilegia (Columbine)." *Evolution*, vol. 75, no. 9, Sept. 2021, pp. 2197–216. *Silverchair*, https://doi.org/10.1111/evo.14313.

Feller, Antje, et al. "Evolutionary and Comparative Analysis of MYB and bHLH Plant Transcription Factors." *The Plant Journal: For Cell and Molecular Biology*, vol. 66, no. 1, Apr. 2011, pp. 94–116. *PubMed*, https://doi.org/10.1111/j.1365-313X.2010.04459.x.

Feng, Chen, et al. "Genetic Architecture of Quantitative Flower and Leaf Traits in a Pair of Sympatric Sister Species of Primulina." *Heredity*, vol. 122, no. 6, June 2019, pp. 864–76. *PubMed*, https://doi.org/10.1038/s41437-018-0170-2.

Fenster, Charles B., et al. "Pollination Syndromes and Floral Specialization." *Annual Review of Ecology, Evolution, and Systematics*, vol. 35, no. 1, 2004, pp. 375–403. *Annual Reviews*, https://doi.org/10.1146/annurev.ecolsys.34.011802.132347.

Fenster, Charles B., and Silvana Martén-Rodríguez. "Reproductive Assurance and the Evolution of Pollination Specialization." *International Journal of Plant Sciences*, vol. 168, no. 2, Feb. 2007, pp. 215–28. *journals.uchicago.edu (Atypon)*, https://doi.org/10.1086/509647.

Fernald, Russell D. "Casting a Genetic Light on the Evolution of Eyes." *Science*, vol. 313, no. 5795, Sept. 2006, pp. 1914–18. *science.org (Atypon)*, https://doi.org/10.1126/science.1127889.

Fishman, Lila, et al. "Minor Quantitative Trait Loci Underlie Floral Traits Associated with Mating System Divergence in Mimulus." *Evolution*, vol. 56, no. 11, 2002, pp. 2138–55.

Galen, Candace. "MEASURING POLLINATOR-MEDIATED SELECTION ON MORPHOMETRIC FLORAL TRAITS: BUMBLEBEES AND THE ALPINE SKY PILOT, POLEMONIUM VISCOSUM." *Evolution; International Journal of Organic Evolution*, vol. 43, no. 4, July 1989, pp. 882–90. *PubMed*, https://doi.org/10.1111/j.1558-5646.1989.tb05185.x.

Galliot, Céline, et al. "Genetics of Flower Size and Nectar Volume in Petunia Pollination Syndromes." *Planta*, vol. 225, no. 1, Dec. 2006, pp. 203–12. *PubMed*, https://doi.org/10.1007/s00425-006-0342-9.

Grabherr, Manfred G., et al. "Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome." *Nature Biotechnology*, vol. 29, no. 7, July 2011, pp. 644–52. *www.nature.com*, https://doi.org/10.1038/nbt.1883.

Haas, Brian J., et al. "De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis." *Nature Protocols*, vol. 8, no. 8, Aug. 2013, pp. 1494–512. *www.nature.com*, https://doi.org/10.1038/nprot.2013.084.

Hoballah, Maria Elena, et al. "Single Gene–Mediated Shift in Pollinator Attraction in Petunia." *The Plant Cell*, vol. 19, no. 3, Mar. 2007, pp. 779–90. *PubMed Central*, https://doi.org/10.1105/tpc.106.048694.

Hong, Lilan, et al. "Variable Cell Growth Yields Reproducible OrganDevelopment through Spatiotemporal Averaging." *Developmental Cell*, vol. 38, no. 1, July 2016, pp. 15–32. *PubMed*, https://doi.org/10.1016/j.devcel.2016.06.016.

Hu, Yibo, et al. "Molecular Mechanisms of Adaptive Evolution in Wild Animals and Plants." *Science China Life Sciences*, vol. 66, no. 3, Mar. 2023, pp. 453–95. *Springer Link*, https://doi.org/10.1007/s11427-022-2233-x.

Irish, Vivian. "The ABC Model of Floral Development." *Current Biology*, vol. 27, no. 17, Sept. 2017, pp. R887–90. *ScienceDirect*, https://doi.org/10.1016/j.cub.2017.03.045.

Joly, Simon, et al. "Greater Pollination Generalization Is Not Associated with Reduced Constraints on Corolla Shape in Antillean Plants." *Evolution*, vol. 72, no. 2, 2018, pp. 244–60. *Wiley Online Library*, https://doi.org/10.1111/evo.13410.

Kierzkowski, Daniel, et al. "A Growth-Based Framework for Leaf Shape Development and Diversity." *Cell*, vol. 177, no. 6, May 2019, pp. 1405-1418.e17. *PubMed*, https://doi.org/10.1016/j.cell.2019.05.011.

Kierzkowski, Daniel, and Anne-Lise Routier-Kierzkowska. "Cellular Basis of Growth in Plants: Geometry Matters." *Current Opinion in Plant Biology*, vol. 47, Feb. 2019, pp. 56–63. *ScienceDirect*, https://doi.org/10.1016/j.pbi.2018.09.008.

Klahre, Ulrich, et al. "Pollinator Choice in Petunia Depends on Two Major Genetic Loci for Floral Scent Production." *Current Biology: CB*, vol. 21, no. 9, May 2011, pp. 730–39. *PubMed*, https://doi.org/10.1016/j.cub.2011.03.059.

Lagomarsino, Laura P., et al. "Repeated Evolution of Vertebrate Pollination Syndromes in a Recently Diverged Andean Plant Clade." *Evolution*, vol. 71, no. 8, 2017, pp. 1970–85. *Wiley Online Library*, https://doi.org/10.1111/evo.13297.

Langfelder, Peter, and Steve Horvath. "WGCNA: An R Package for Weighted Correlation Network Analysis." *BMC Bioinformatics*, vol. 9, Dec. 2008, p. 559. *PubMed*, https://doi.org/10.1186/1471-2105-9-559.

Liu, Liang, et al. "Estimating Species Phylogenies Using Coalescence Times among Sequences." *Systematic Biology*, vol. 58, no. 5, Oct. 2009, pp. 468–77. *Silverchair*, https://doi.org/10.1093/sysbio/syp031.

Liu, Yang, et al. "Convergent Sequence Evolution between Echolocating Bats and Dolphins." *Current Biology: CB*, vol. 20, no. 2, Jan. 2010, pp. R53-54. *PubMed*, https://doi.org/10.1016/j.cub.2009.11.058.

Losos, Jonathan B. "CONVERGENCE, ADAPTATION, AND CONSTRAINT." *Evolution*, vol. 65, no. 7, July 2011, pp. 1827–40. *Silverchair*, https://doi.org/10.1111/j.1558-5646.2011.01289.x.

Love, Michael I., et al. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology*, vol. 15, no. 12, 2014, p. 550. *PubMed*, https://doi.org/10.1186/s13059-014-0550-8.

Luo, D., et al. "Origin of Floral Asymmetry in Antirrhinum." *Nature*, vol. 383, no. 6603, Oct. 1996, pp. 794–99. *PubMed*, https://doi.org/10.1038/383794a0.

Madison, Stephanie L., and Andreas Nebenführ. "Understanding Myosin Functions in Plants: Are We There Yet?" *Current Opinion in Plant Biology*, vol. 16, no. 6, Dec. 2013, pp. 710–17. *PubMed*, https://doi.org/10.1016/j.pbi.2013.10.004.

Martén-Rodríguez, Silvana, Mauricio Quesada, et al. "A Comparison of Reproductive Strategies between Island and Mainland Caribbean Gesneriaceae." *Journal of Ecology*, vol. 103, no. 5, 2015, pp. 1190–204. *Wiley Online Library*, https://doi.org/10.1111/1365-2745.12457.

Martén-Rodríguez, Silvana, Abel Almarales-Castro, et al. "Evaluation of Pollination Syndromes in Antillean Gesneriaceae: Evidence for Bat, Hummingbird and Generalized Flowers." *Journal of Ecology*, vol. 97, no. 2, 2009, pp. 348–59. *Wiley Online Library*, https://doi.org/10.1111/j.1365-2745.2008.01465.x.

Martén-Rodríguez, Silvana, Charles B. Fenster, et al. "Evolutionary Breakdown of Pollination Specialization in a Caribbean Plant Radiation." *New Phytologist*, vol. 188, no. 2, 2010, pp. 403–17. *Wiley Online Library*, https://doi.org/10.1111/j.1469-8137.2010.03330.x.

---. "Pollination Ecology and Breeding Systems of Five Gesneria Species from Puerto Rico." *Annals of Botany*, vol. 102, no. 1, July 2008, pp. 23–30. *PubMed Central*, https://doi.org/10.1093/aob/mcn056.

Min, Ya, et al. "Homologs of the STYLISH Gene Family Control Nectary Development in Aquilegia." *The New Phytologist*, vol. 221, no. 2, Jan. 2019, pp. 1090–100. *PubMed*, https://doi.org/10.1111/nph.15406.

Moses, Lambda, and Lior Pachter. "Museum of Spatial Transcriptomics." *Nature Methods*, vol. 19, no. 5, 5, May 2022, pp. 534–46. *www.nature.com*, https://doi.org/10.1038/s41592-022-01409-2.

Muchhala, Nathan. "Adaptive Trade-off in Floral Morphology Mediates Specialization for Flowers Pollinated by Bats and Hummingbirds." *The American Naturalist*, vol. 169, no. 4, Apr. 2007, pp. 494–504. *PubMed*, https://doi.org/10.1086/512047.

Nag, Anwesha, et al. "miR319a Targeting of TCP4 Is Critical for Petal Growth and Development in Arabidopsis." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 52, Dec. 2009, pp. 22534–39. *PubMed*, https://doi.org/10.1073/pnas.0908718106.

Oyston, Jack W., et al. "Molecular Phylogenies Map to Biogeography Better than Morphological Ones." *Communications Biology*, vol. 5, no. 1, 1, May 2022, pp. 1–12. *www.nature.com*, https://doi.org/10.1038/s42003-022-03482-x.

Papadopulos, Alexander S. T., et al. "Convergent Evolution of Floral Signals Underlies the Success of Neotropical Orchids." *Proceedings of the Royal Society B: Biological Sciences*,

vol. 280, no. 1765, Aug. 2013, p. 20130960. *PubMed Central*, https://doi.org/10.1098/rspb.2013.0960.

Parker, Joe, et al. "Genome-Wide Signatures of Convergent Evolution in Echolocating Mammals." *Nature*, vol. 502, no. 7470, 7470, Oct. 2013, pp. 228–31. *www.nature.com*, https://doi.org/10.1038/nature12511.

Pelaz, S., et al. "B and C Floral Organ Identity Functions Require SEPALLATA MADS-Box Genes." *Nature*, vol. 405, no. 6783, May 2000, pp. 200–03. *PubMed*, https://doi.org/10.1038/35012103.

Poulin, V., D. Amesefe, E. Gonzalez, H. Alexandre, and S. Joly. 2022. Testing candidate genes linked to corolla shape variation of a pollinator shift in Rhytidophyllum (Gesneriaceae). *PLOS ONE* 17:e0267540.

Protas, Meredith E., et al. "Genetic Analysis of Cavefish Reveals Molecular Convergence in the Evolution of Albinism." *Nature Genetics*, vol. 38, no. 1, 1, Jan. 2006, pp. 107–11. *www.nature.com*, https://doi.org/10.1038/ng1700.

Puzey, Joshua R., et al. "Evolution of Spur-Length Diversity in Aquilegia Petals Is Achieved Solely through Cell-Shape Anisotropy." *Proceedings. Biological Sciences*, vol. 279, no. 1733, Apr. 2012, pp. 1640–45. *PubMed*, https://doi.org/10.1098/rspb.2011.1873.

Rambaud-Lavigne, Léa, and Angela Hay. "Floral Organ Development Goes Live." *Journal of Experimental Botany*, vol. 71, no. 9, May 2020, pp. 2472–78. *PubMed*, https://doi.org/10.1093/jxb/eraa038.

Rebocho, Alexandra B., et al. "Generation of Shape Complexity through Tissue Conflict Resolution." *eLife*, vol. 6, Feb. 2017. *PubMed*, https://doi.org/10.7554/eLife.20156.

Roalson, Eric H., and Wade R. Roberts. "Distinct Processes Drive Diversification in Different Clades of Gesneriaceae." *Systematic Biology*, vol. 65, no. 4, July 2016, pp. 662–84. *Silverchair*, https://doi.org/10.1093/sysbio/syw012.

Roberts, Wade R., and Eric H. Roalson. "Co-Expression Clustering across Flower Development Identifies Modules for Diverse Floral Forms in Achimenes (Gesneriaceae)." *PeerJ*, vol. 8, Mar. 2020, p. e8778. *peerj.com*, https://doi.org/10.7717/peerj.8778.

---. "Comparative Transcriptome Analyses of Flower Development in Four Species of Achimenes (Gesneriaceae)." *BMC Genomics*, vol. 18, no. 1, Mar. 2017, p. 240. *PubMed*, https://doi.org/10.1186/s12864-017-3623-8.

Rolland-Lagan, Anne-Gaëlle, et al. "Growth Dynamics Underlying Petal Shape and Asymmetry." *Nature*, vol. 422, no. 6928, 6928, Mar. 2003, pp. 161–63. *www.nature.com*, https://doi.org/10.1038/nature01443.

Romero, Irene Gallego, et al. "Comparative Studies of Gene Expression and the Evolution of Gene Regulation." *Nature Reviews Genetics*, vol. 13, no. 7, 7, July 2012, pp. 505–16. *www.nature.com*, https://doi.org/10.1038/nrg3229.

Rosas-Guerrero, Víctor, et al. "A Quantitative Review of Pollination Syndromes: Do Floral Traits Predict Effective Pollinators?" *Ecology Letters*, vol. 17, no. 3, Mar. 2014, pp. 388–400. *PubMed*, https://doi.org/10.1111/ele.12224.

Sauret-Güeto, Susanna, et al. "JAGGED Controls Arabidopsis Petal Growth and Shape by Interacting with a Divergent Polarity Field." *PLOS Biology*, vol. 11, no. 4, Apr. 2013, p. e1001550. *PLoS Journals*, https://doi.org/10.1371/journal.pbio.1001550.

Schemske, D. W., and H. D. Bradshaw. "Pollinator Preference and the Evolution of Floral Traits in Monkeyflowers (Mimulus)." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 21, Oct. 1999, pp. 11910–15. *PubMed*, https://doi.org/10.1073/pnas.96.21.11910.

Schiessl, Katharina, et al. "Arabidopsis JAGGED Links Floral Organ Patterning to Tissue Growth by Repressing Kip-Related Cell Cycle Inhibitors." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 7, Feb. 2014, pp. 2830–35. *PubMed*, https://doi.org/10.1073/pnas.1320457111.

Schiestl, Florian P., and Steven D. Johnson. "Pollinator-Mediated Evolution of Floral Signals." *Trends in Ecology & Evolution*, vol. 28, no. 5, May 2013, pp. 307–15. *PubMed*, https://doi.org/10.1016/j.tree.2013.01.019.

Serrano-Serrano, Martha L., et al. "Transcriptomic Resources for an Endemic Neotropical Plant Lineage (Gesneriaceae)." *Applications in Plant Sciences*, vol. 5, no. 4, Apr. 2017, p. apps.1600135. *PubMed*, https://doi.org/10.3732/apps.1600135.

Soltis, Douglas E., et al. "Angiosperm Phylogeny: 17 Genes, 640 Taxa." *American Journal of Botany*, vol. 98, no. 4, Apr. 2011, pp. 704–30. *PubMed*, https://doi.org/10.3732/ajb.1000404.

Soneson, Charlotte, et al. "Differential Analyses for RNA-Seq: Transcript-Level Estimates Improve Gene-Level Inferences." *F1000Research*, vol. 4, 2015, p. 1521. *PubMed*, https://doi.org/10.12688/f1000research.7563.2.

Specht, Chelsea, and Madelaine Bartlett. "Flower Evolution: The Origin and Subsequent Diversification of the Angiosperm Flower." *Annu Rev Ecol Syst*, vol. 40, Feb. 2009, pp. 217–43. *ResearchGate*, https://doi.org/10.1146/annurev.ecolsys.110308.120203.

Stebbins, G. Ledyard. "Adaptive Radiation of Reproductive Characteristics in Angiosperms, I: Pollination Mechanisms." *Annual Review of Ecology and Systematics*, vol. 1, no. 1, 1970, pp. 307–26. *Annual Reviews*, https://doi.org/10.1146/annurev.es.01.110170.001515.

Stern, David L. "The Genetic Causes of Convergent Evolution." *Nature Reviews Genetics*, vol. 14, no. 11, 11, Nov. 2013, pp. 751–64. *www.nature.com*, https://doi.org/10.1038/nrg3483.

Stern, David L., and Virginie Orgogozo. "The Loci of Evolution: How Predictable Is Genetic Evolution?" *Evolution; International Journal of Organic Evolution*, vol. 62, no. 9, Sept. 2008, pp. 2155–77. *PubMed Central*, https://doi.org/10.1111/j.1558-5646.2008.00450.x.

Stuurman, Jeroen, et al. "Dissection of Floral Pollination Syndromes in Petunia." *Genetics*, vol. 168, no. 3, Nov. 2004, pp. 1585–99. *www.genetics.org*, https://doi.org/10.1534/genetics.104.031138.

Uebbing, Severin, et al. "Divergence in Gene Expression within and between Two Closely Related Flycatcher Species." *Molecular Ecology*, vol. 25, no. 9, 2016, pp. 2015–28. *Wiley Online Library*, https://doi.org/10.1111/mec.13596.

Vincent, Coral A., and Enrico S. Coen. "A Temporal and Morphological Framework for Flower Development in *Antirrhinum Majus*." *Canadian Journal of Botany*, vol. 82, no. 5, May 2004, pp. 681–90. *DOI.org (Crossref)*, https://doi.org/10.1139/b04-042.

Wake, David B. "Homoplasy: The Result of Natural Selection, or Evidence of Design Limitations?" *The American Naturalist*, vol. 138, no. 3, 1991, pp. 543–67.

Wessinger, Carolyn A., et al. "Identification of Major Quantitative Trait Loci Underlying Floral Pollination Syndrome Divergence in Penstemon." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 369, no. 1648, Aug. 2014, p. 20130349. *PubMed*, https://doi.org/10.1098/rstb.2013.0349.

Wester, Petra, and Regine Claßen-Bockhoff. "Floral Diversity and Pollen Transfer Mechanisms in Bird-Pollinated Salvia Species." *Annals of Botany*, vol. 100, no. 2, Aug. 2007, pp. 401–21. *PubMed Central*, https://doi.org/10.1093/aob/mcm036.

Whitney, Heather M., et al. "Conical Epidermal Cells Allow Bees to Grip Flowers and Increase Foraging Efficiency." *Current Biology*, vol. 19, no. 11, June 2009, pp. 948–53. *ScienceDirect*, https://doi.org/10.1016/j.cub.2009.04.051.

Whittall, Justen B., and Scott A. Hodges. "Pollinator Shifts Drive Increasingly Long Nectar Spurs in Columbine Flowers." *Nature*, vol. 447, no. 7145, 7145, June 2007, pp. 706–09. *www.nature.com*, https://doi.org/10.1038/nature05857.

Wray, Gregory A. "The Evolutionary Significance of Cis-Regulatory Mutations." *Nature Reviews. Genetics*, vol. 8, no. 3, Mar. 2007, pp. 206–16. *PubMed*, https://doi.org/10.1038/nrg2063.

Yant, Levi, et al. "Molecular Basis for Three-Dimensional Elaboration of the Aquilegia Petal Spur." *Proceedings. Biological Sciences*, vol. 282, no. 1803, Mar. 2015, p. 20142778. *PubMed*, https://doi.org/10.1098/rspb.2014.2778.

Yuan, Yao-Wu, et al. "Genetic Dissection of a Major Anthocyanin QTL Contributing to Pollinator-Mediated Reproductive Isolation between Sister Species of Mimulus." *Genetics*, vol. 194, no. 1, May 2013, pp. 255–63. *PubMed*, https://doi.org/10.1534/genetics.112.146852.

Zhang, J., and S. Kumar. "Detection of Convergent and Parallel Evolution at the Amino Acid Sequence Level." *Molecular Biology and Evolution*, vol. 14, no. 5, May 1997, pp. 527–36. *PubMed*, https://doi.org/10.1093/oxfordjournals.molbev.a025789.

Zhang, Tian, et al. "Disentangling Loosening from Softening: Insights into Primary Cell Wall Structure." *The Plant Journal: For Cell and Molecular Biology*, vol. 100, no. 6, Dec. 2019, pp. 1101–17. *PubMed*, https://doi.org/10.1111/tpj.14519.

Appendix

Table 1. Identification of specimens used for cell measurement and RNA-seq analysis.

Species	Collection number of Montreal Botanical Garden	Specimen voucher collector and number at Marie-Victorin Herbarium
G. quisqueyana	891-2015	no voucher
G. ventricosa	7778-1968	Léveillé-Bourret, G4
G. cuneifolia	976-2017	no voucher
G. acaulis	1328-2021	Joly 1149
G. bicolor	587-2019	no voucher
R. auriculatum	937-1971	no voucher
R. vernicosum	1267-1966	Léveillé-Bourret, G3
R. rupincola	113-1991	Léveillé-Bourret, G5
R. tomentosum	943-1971	Léveillé-Bourret, G2
R. exsertum	1073-2010	Léveillé-Bourret, G1
R. intermedium	1450-2021	Joly 1101

Table 2. Selected Gene Ontology terms of genes uniquely expressed in Zone A

Homologous cluster ID	Protein name	Species of origin	GO-terms
cluster13289	ROMT	*A. thaliana*	GO:0008171^molecular_function^O-methyltransferase activity`GO:0046983^molecular_function^protein dimerization activity`GO:0102303^molecular_function^resveratrol 3,5-O-dimethyltransferase activity`GO:0008757^molecular_function^S-adenosylmethionine-dependent methyltransferase activity`GO:0019438^biological_process^aromatic compound biosynthetic process`GO:0032259^biological_process^methylation
cluster14970	LIP4	*A. thaliana*	GO:0005576^cellular_component^extracellular region`GO:0016788^molecular_function^hydrolase activity, acting on ester bonds`GO:0016042^biological_process^lipid catabolic process
	GDL2	*A. thaliana*	GO:0009570^cellular_component^chloroplast stroma`GO:0005576^cellular_component^extracellular region`GO:0016788^molecular_function^hydrolase activity, acting on ester bonds`GO:0016042^biological_process^lipid catabolic process

cluster18893	AGL16	*A. thaliana*	GO:0005634^cellular_component^nucleus`GO:0003700^molecular_function^DNA-binding transcription factor activity`GO:0000981^molecular_function^DNA-binding transcription factor activity, RNA polymerase II-specific`GO:0042803^molecular_function^protein homodimerization activity`GO:0000978^molecular_function^RNA polymerase II cis-regulatory region sequence-specific DNA binding`GO:0000976^molecular_function^transcription cis-regulatory region binding`GO:0009908^biological_process^flower development`GO:0045944^biological_process^positive regulation of transcription by RNA polymerase II`GO:0006357^biological_process^regulation of transcription by RNA polymerase II`GO:0010440^biological_process^stomatal lineage progression
	MAD27	*O. sativa*	GO:0005634^cellular_component^nucleus`GO:0000981^molecular_function^DNA-binding transcription factor activity, RNA polymerase II-specific`GO:0046983^molecular_function^protein dimerization activity`GO:0000978^molecular_function^RNA polymerase II cis-regulatory region sequence-specific DNA binding`GO:0045944^biological_process^positive regulation of transcription by RNA polymerase II`GO:0006357^biological_process^regulation of transcription by RNA polymerase II
cluster19380	P2C39	*A. thaliana*	GO:0005829^cellular_component^cytosol`GO:0046872^molecular_function^metal ion binding`GO:0017018^molecular_function^myosin phosphatase activity`GO:0006470^biological_process^protein dephosphorylation
cluster25053	CESA5	*O. sativa*	GO:0005886^cellular_component^plasma membrane`GO:0016760^molecular_function^cellulose synthase (UDP-forming) activity`GO:0016759^molecular_function^cellulose synthase activity`GO:0046872^molecular_function^metal ion binding`GO:0071555^biological_process^cell wall organization`GO:0030244^biological_process^cellulose biosynthetic process`GO:0009833^biological_process^plant-type primary cell wall biogenesis
	CESA5	*A. thaliana*	GO:0005886^cellular_component^plasma membrane`GO:0016760^molecular_function^cellulose synthase (UDP-forming) activity`GO:0016759^molecular_function^cellulose synthase activity`GO:0046872^molecular_function^metal ion binding`GO:0071555^biological_process^cell wall organization`GO:0030244^biological_process^cellulose biosynthetic process`GO:0010192^biological_process^mucilage biosynthetic process`GO:0009833^biological_process^plant-type primary cell wall biogenesis
cluster25326	RHA4A	*A. thaliana*	GO:0016020^cellular_component^membrane`GO:0046872^molecular_function^metal ion binding`GO:0061630^molecular_function^ubiquitin protein ligase activity`GO:0016567^biological_process^protein ubiquitination`GO:0006511^biological_process^ubiquitin-dependent protein catabolic process
cluster25572	GNT10	*A. thaliana*	GO:0009507^cellular_component^chloroplast`GO:0031415^cellular_component^NatA complex`GO:0008080^molecular_function^N-acetyltransferase activity`GO:0007064^biological_process^mitotic sister chromatid cohesion`GO:0006474^biological_process^N-terminal protein amino acid acetylation`GO:0018394^biological_process^peptidyl-lysine acetylation

cluster25801	PXL2C	*A. thaliana*	GO:0009507^cellular_component^chloroplast
cluster26313	CRK9	*A. thaliana*	GO:0048046^cellular_component^apoplast`GO:0099503^cellular_component^secretory vesicle`GO:0012501^biological_process^programmed cell death`GO:0009751^biological_process^response to salicylic acid`GO:0009627^biological_process^systemic acquired resistance

Table 3. Selected Gene Ontology terms of the 203 genes found to be involved in the determination of form by DESeq2 and WGCNA.

Homologous cluster ID	Protein name	Species of origin	GO-terms
cluster12857	DNJ15	*A. thaliana*	GO:0005856^cellular_component^cytoskeleton`GO:0005789^cellular_component^endoplasmic reticulum membrane`GO:0000139^cellular_component^Golgi membrane`GO:0008092^molecular_function^cytoskeletal protein binding`GO:0009958^biological_process^positive gravitropism
cluster14456	PSB6	*A. thaliana*	GO:0005737^cellular_component^cytoplasm`GO:0005634^cellular_component^nucleus`GO:0005839^cellular_component^proteasome core complex`GO:0019774^cellular_component^proteasome core complex, beta-subunit complex`GO:0004175^molecular_function^endopeptidase activity`GO:0004298^molecular_function^threonine-type endopeptidase activity`GO:0010498^biological_process^proteasomal protein catabolic process
cluster14970	LIP4	*A. thaliana*	GO:0005576^cellular_component^extracellular region`GO:0016788^molecular_function^hydrolase activity, acting on ester bonds`GO:0016042^biological_process^lipid catabolic process
cluster14994	GDL88	*A. thaliana*	GO:0005783^cellular_component^endoplasmic reticulum`GO:0005634^cellular_component^nucleus`GO:0016787^molecular_function^hydrolase activity`GO:0016042^biological_process^lipid catabolic process
cluster15072	NDUS4	*A. thaliana*	GO:0005747^cellular_component^mitochondrial respiratory chain complex I`GO:0005739^cellular_component^mitochondrion`GO:0009536^cellular_component^plastid`GO:0050897^molecular_function^cobalt ion binding`GO:0008137^molecular_function^NADH dehydrogenase (ubiquinone) activity`GO:0009631^biological_process^cold acclimation`GO:0006970^biological_process^response to osmotic stress
cluster15098	NP214	*A. thaliana*	GO:0005739^cellular_component^mitochondrion`GO:0005643^cellular_component^nuclear pore`GO:0008139^molecular_function^nuclear localization sequence binding`GO:0017056^molecular_function^structural constituent of nuclear pore`GO:0009793^biological_process^embryo development ending in seed dormancy`GO:0051028^biological_process^mRNA transport`GO:0006606^biological_process^protein import into nucleus`GO:0006405^biological_process^RNA export from nucleus`GO:0010070^biological_process^zygote asymmetric cell division

cluster15257	MUC70	*A. thaliana*	GO:0005794^cellular_component^Golgi apparatus`GO:0000139^cellular_component^Golgi membrane`GO:0016757^molecular_function^glycosyltransferase activity`GO:0080001^biological_process^mucilage extrusion from seed coat`GO:0048358^biological_process^mucilage pectin biosynthetic process`GO:0010246^biological_process^rhamnogalacturonan I biosynthetic process`GO:0045491^biological_process^xylan metabolic process
cluster15319	A10A5	*Soybean*	GO:0009733: response to auxin; GO:0009734: auxin-activated signaling pathway
cluster15324	TTG1	*A. thaliana*	GO:0005737^cellular_component^cytoplasm`GO:0005634^cellular_component^nucleus`GO:0003677^molecular_function^DNA binding`GO:0045165^biological_process^cell fate commitment`GO:0009957^biological_process^epidermal cell fate specification`GO:0032880^biological_process^regulation of protein localization`GO:0009733^biological_process^response to auxin`GO:0009723^biological_process^response to ethylene`GO:0010026^biological_process^trichome differentiation
cluster15361	TBL11	*A. thaliana*	GO:0005794^cellular_component^Golgi apparatus`GO:0016020^cellular_component^membrane`GO:0016413^molecular_function^O-acetyltransferase activity
cluster15365	NUD23	*A. thaliana*	GO:0009507^cellular_component^chloroplast`GO:0047631^molecular_function^ADP-ribose diphosphatase activity`GO:0047884^molecular_function^FAD diphosphatase activity`GO:0046872^molecular_function^metal ion binding`GO:0042726^biological_process^flavin-containing compound metabolic process`GO:0009416^biological_process^response to light stimulus
cluster16283	PTN2A	*A. thaliana*	GO:0005829^cellular_component^cytosol`GO:0070300^molecular_function^phosphatidic acid binding`GO:0052866^molecular_function^phosphatidylinositol phosphate phosphatase activity`GO:0016314^molecular_function^phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase activity`GO:0004725^molecular_function^protein tyrosine phosphatase activity`GO:0016311^biological_process^dephosphorylation`GO:0035335^biological_process^peptidyl-tyrosine dephosphorylation`GO:0046856^biological_process^phosphatidylinositol dephosphorylation`GO:0006970^biological_process^response to osmotic stress`GO:0009651^biological_process^response to salt stress
cluster16458	P2C10	*A. thaliana*	GO:0005886^cellular_component^plasma membrane`GO:0046872^molecular_function^metal ion binding`GO:0017018^molecular_function^myosin phosphatase activity`GO:0006470^biological_process^protein dephosphorylation
cluster18119	GDL79	*A. thaliana*	GO:0005576^cellular_component^extracellular region`GO:0016788^molecular_function^hydrolase activity, acting on ester bonds`GO:0042335^biological_process^cuticle development`GO:0016042^biological_process^lipid catabolic process
cluster18155	HHO3	*A. thaliana*	GO:0005634^cellular_component^nucleus`GO:0003677^molecular_function^DNA binding`GO:0003700^molecular_function^DNA-binding transcription factor activity`GO:0071456^biological_process^cellular response to hypoxia`GO:0016036^biological_process^cellular response to phosphate starvation`GO:0006355^biological_process^regulation of DNA-templated transcription

cluster18407	GSTZ	*E. esula*	GO:0005737^cellular_component^cytoplasm`GO:0004364^molecular_function^glutathione transferase activity`GO:0009072^biological_process^aromatic amino acid metabolic process`GO:0042221^biological_process^response to chemical
cluster18503	NAT1	*A. thaliana*	GO:0016020^cellular_component^membrane`GO:0009506^cellular_component^plasmodesma`GO:0022857^molecular_function^transmembrane transporter activity`GO:0071702^biological_process^organic substance transport
cluster18572	RVE3	*A. thaliana*	GO:0005634^cellular_component^nucleus`GO:0003677^molecular_function^DNA binding`GO:0003700^molecular_function^DNA-binding transcription factor activity
	RVE6	*A. thaliana*	GO:0005634^cellular_component^nucleus`GO:0003677^molecular_function^DNA binding`GO:0003700^molecular_function^DNA-binding transcription factor activity`GO:0042752^biological_process^regulation of circadian rhythm
cluster18684	PAR1	*Rosa hybrid cultivar*	GO:0016491^molecular_function^oxidoreductase activity
	CAD2	*M. trucantula*	GO:0005737^cellular_component^cytoplasm`GO:0045551^molecular_function^cinnamyl-alcohol dehydrogenase activity`GO:0016616^molecular_function^oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor`GO:0052747^molecular_function^sinapyl alcohol dehydrogenase activity`GO:0009699^biological_process^phenylpropanoid biosynthetic process
cluster18812	CLT3	*A. thaliana*	GO:0031969^cellular_component^chloroplast membrane`GO:0009536^cellular_component^plastid`GO:0002229^biological_process^defense response to oomycetes`GO:0034635^biological_process^glutathione transport`GO:0046686^biological_process^response to cadmium ion
cluster18840	NAT6	*A. thaliana*	GO:0005783^cellular_component^endoplasmic reticulum`GO:0016020^cellular_component^membrane`GO:0009505^cellular_component^plant-type cell wall`GO:0009506^cellular_component^plasmodesma`GO:0005773^cellular_component^vacuole`GO:0022857^molecular_function^transmembrane transporter activity`GO:0071702^biological_process^organic substance transport
	NAT7	*A. thaliana*	GO:0005886^cellular_component^plasma membrane`GO:0009506^cellular_component^plasmodesma`GO:0022857^molecular_function^transmembrane transporter activity`GO:0071702^biological_process^organic substance transport
cluster18853	PCKA2	*A. thaliana*	GO:0005829^cellular_component^cytosol`GO:0005524^molecular_function^ATP binding`GO:0016301^molecular_function^kinase activity`GO:0046872^molecular_function^metal ion binding`GO:0004612^molecular_function^phosphoenolpyruvate carboxykinase (ATP) activity`GO:0006094^biological_process^gluconeogenesis`GO:0016310^biological_process^phosphorylation

	PCKA1	*A. thaliana*	GO:0005737^cellular_component^cytoplasm`GO:0005829^cellular_component^cytosol`GO:0005730^cellular_component^nucleolus`GO:0005524^molecular_function^ATP binding`GO:0046872^molecular_function^metal ion binding`GO:0004612^molecular_function^phosphoenolpyruvate carboxykinase (ATP) activity`GO:0016036^biological_process^cellular response to phosphate starvation`GO:0050832^biological_process^defense response to fungus`GO:0006094^biological_process^gluconeogenesis
cluster19091	Y4345	*A. thaliana*	GO:0005886^cellular_component^plasma membrane`GO:0005524^molecular_function^ATP binding`GO:0106310^molecular_function^protein serine kinase activity`GO:0004674^molecular_function^protein serine/threonine kinase activity`GO:0006468^biological_process^protein phosphorylation